THE RANDOMIZATION MODEL FOR THE

ANALYSIS OF COVARIANCE

By

ROBERT S. WILSON

Bachelor of Arts
Grinnell College
Grinnell, Iowa
1970

Master of Science
Iowa State University
Ames, Iowa
1980

Submitted to the Faculty of the
Graduate College
of the
Oklahoma State University
in partial fulfillment of
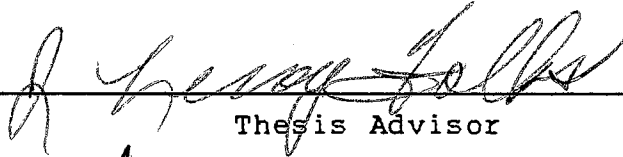the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 1993

THE RANDOMIZATION MODEL FOR THE
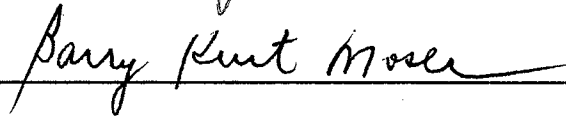
ANALYSIS OF COVARIANCE
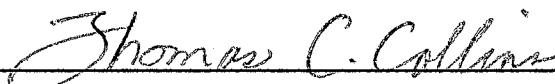
Thesis Approved:

_____
Thesis Advisor

_____

_____

_____

_____
Dean of the Graduate College

PREFACE

This dissertation is a wide-ranging examination of the randomization model as applied to the analysis of covariance. The focus is on the randomized block design with one blocking factor, one treatment factor, and one covariate. The study contrasts two sources of randomness, the usual source, normally distributed random errors, and unit errors arising from the random assigning of treatments to experimental units. Examined are two types of unit errors, experimental unit errors and experimental unit-treatment interaction errors.

We replace the $\Sigma\ Y_{ijk}\ s_{ij}^{k}$ technique with a new permutation matrix technique, which is easily adapted to other models. We replace the probability of a unit receiving a given treatment with the probability of selecting a given randomization. This facilitates studying unequal selection probabilities. Errors generated for numerical simulations are linearly independent of the covariate. The F-distribution analogue of normal probability paper facilitates comparison with the case of normally distributed random errors.

iv

We prove that the randomization distribution of several of the usual anova-table terms and estimators consists of clusters of values. For experimental unit errors, each cluster collapses to a single value for several terms and estimators. In the single covariate case, the noncentrality parameters for the tests of the usual hypotheses of adjusted treatment effects and of adjusted covariate are linked. The linkage of the noncentrality parameters permits avoiding, for both tests, randomizations with low power.

Also derived is the relationship between the Kempthorne and the Neyman definitions of experimental unit-treatment interaction errors and the relationship required for two matrices to have identical projectors. An open question is the need for, or suitability of, the F-distribution as derived from a singular normal distribution. Appendix E lists several possible extensions.

A numerical simulation examines the three types of error, each with equal and unequal within-block variances, and with zero and one nonzero set of treatment effects. Computer programs compute all possible randomizations for each of the 3x2x2=12 cases. While the experimental unit-treatment interaction errors provide the F-ratios with a distribution close to the F distribution, the experimental unit errors do not. Nonequality of the within-block errors lowers the mean of the

F-ratios and their variances. A weighting of the selection probabilities appears of little advantage. Restricting randomization to the cluster with the largest noncentrality parameter for the adjusted covariate proves advantageous. For this model, this cluster is the most balanced on the covariate.

This work is designed to rapidly introduce one familiar with the general linear model to the terminology and topics of the randomization model. To this end, it presents concepts from diverse sources in a uniform notation, reviews aspects of the general linear model, presses the bibliography into service as an index to its author's cited pages, indexes notation and key terms, and has numerous internal references. The trained statistician will find the style slow, even verbose and repetitious. We hope his/her graduate student will find such helpful.

The discoveries of this dissertation could not have been made without the extensive work of others, especially Drs. Oscar Kempthorne and S. R. Searle, both of whom built upon the work of many others. Software (SAS 6.07) and mainframe (IBM 3090-200S) available in the 1990's permitted simulations difficult, if not impossible, fifteen or more years ago. My thanks to the anonymous doners of the Version 6 SAS manuals. Earlier versions of parts of this work have been presented at the Annual

Meetings of The American Statistical Association of 1988, 1989, and 1990.

Hypotheses are like nets:

only he who casts can catch.

G. R. Dolby

BRIEF TABLE OF CONTENTS

DETAILED TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

xix

LIST OF SYMBOLS

Guide: Subscripts - i, j, k, s
Scalers - μ, b, τ, β
Vectors - τ, β, $\mathbb{O}$, $\mathbb{1}$, Y, $\mathbb{Z}$
Estimators - $\hat{b}$, $\hat{\tau}$, $\hat{\beta}$
Matrices - $\mathbb{1}$, $\mathbb{I}$, $\mathbb{O}$, Y, X, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{M}$, $\mathbb{A}$
Sets - $\mathbb{U}$, $\mathbb{L}$, $\mathbb{G}$, $\mathbb{B}$, $\mathbb{E}$, $\mathbb{S}$, $\mathbb{T}$
Number of elements in a set - $N(\mathbb{G})$, $N(\mathbb{B})$, $N(\mathbb{E})$

Symbol   Description (page of first use)

$\mathbb{O}$ a vector or matrix of zeros (16)

$\mathbb{1}$ vector of ones (20)

$\propto$ Type I error probability (171)

@s   the randomization s used to conduct the actual experiment (169)

b   number of blocks (20)
vector of block effects (19)

b_i effect of block i (22, 123)

$\hat{b}_i$ estimate of effect of block i = mean response

$(bt)_{ik}$ block i x treatment k interaction (123, 126)

β^+ all model equation parameters, τ and β (16)

β_h elements of β^+ being tested (42)

β_r elements of β^+ other than β_h (42)

β covariate parameter(s) (19, 22)
Type II error probability

$\hat{\beta}$ unique estimator of β (18)

$\hat{\beta}^{+o}$ a solution to the least squares equations (17)

$\hat{\beta}_{\mu,b,\tau,\beta}$ $\hat{\beta}$ estimate of β as adjusted for μ, b, & τ (30)

xx

Symbol   Description   (page of first use)

$\hat{F}_{s;\mu,b,\tau,\beta}$ $\hat{F}_{\mu,b,\tau,\beta}$ for randomization s (207)

$\hat{\beta}_R$ randomization probability-Type II error (173)

$\hat{\beta}_{R(F)}$ simulation probability-Type II error (176-177)

$\mathcal{B}$ set of block-wise randomizations (108)

$\mathcal{B}^n$ n-dimensional σ-field of Borel sets (16, 157)

$\mathcal{C}_{j1:k}$ set of block-wise randomizations with unit jl
                           having treatment k; inclusion probability (111)

d   denominator degrees of freedom for F-ratio (39)

e_{ij}^K Dr. Kempthorne's experimental unit error (124)

e_{ij}^{N1} Dr. Neyman's experimental unit error (127)

e_{ij}^P a postulated experimental unit error (125)

e   vector of experimental unit errors (198)

$E_N(\cdot)$ normal (Gaussian) expectation of $(\cdot)$ (25)

$E_R(\cdot)$ randomization expectation of $(\cdot)$ (114)

$\mathcal{E}$ set of experiment-wise randomizations (108)

$\mathbb{E}$ matrix of experimental unit errors (154)

ε vector of random errors (16)

ε_{ik} error for unit i receiving treatment k (22)

$\hat{\varepsilon}$ vector of residuals (17, 18)

$\mathbb{E}$ matrix of columns=experimental unit error (154)

$\mathbb{E}_s$ column of $\mathbb{E}$ for randomization s (157)

F   F-ratio (39)

$\dot{F}_s$ F-ratio for randomization s (202)

Symbol   Description   (page of first use)

$F(\beta|\mu,b,\tau)$ F-ratio to test $\beta=0$ as adjusted by μ,b,τ (63)

$F(\tau|\mu,b,\beta)$ F-ratio to test $\tau=0$ as adjusted by μ,b,β (64)

g the label of a g-group (194)

g-group   randomizations having the same value of g (195)

$\mathbb{G}$ any generalized inverse (83)

$\mathcal{G}$ the set of distinguishable groupings of randomizations (106-107)

$H_{0;\mu,b,\beta}\tau=0$ a null hypothesis test of $\tau=0$ when μ, b, and β are in the model equation (63-64)

$\mathbb{H}$ matrix of experimental unit-treatment interaction errors (155)

$\mathbb{H}_s$ column of $\mathbb{H}$ for randomization s (157)

η_{ijk}^{K} Dr. Kempthorne's experimental unit-treatment interaction error (128)

η_{ijk}^{N1} Dr. Neyman's experimental unit-treatment interaction error (127, 132)

η_{ijk}^{N2} η_{ijk}^{N1} for a model equation with blocks (127)

η_{ijk}^{P} A postulated experimental unit-treatment interaction error (128, 132)

η_{ij_l} form of η_{ijk} (234)

$\eta_{i\cdot k}$ form of η_{ijk} (235)

η_s vector of experimental unit-treatment interaction errors for randomization s (199)

i block label subscript (22)

$\mathbb{I}$ identity matrix

$\mathcal{Q}$ power set of $\mathbb{U}$ (142, 157)

j experimental unit label subscript (123)

$\mathbb{J}$ a matrix of ones

Symbol	Description (page of first use)

k — treatment label subscript (22)

L_s — left hand side of an inequality (217-219)

$\mathbb{L}$ — a matrix, usually $\mathbb{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ (18)

λ — noncentrality parameter (39)

$\lambda_{\beta|\mu,b,\tau,\beta}$ — λ for the adjusted covariates (63)

$\lambda_{s,\beta|\mu,b,\tau,\beta}$ — $\lambda_{\beta|\mu,b,\tau,\beta}$ for randomization s (212)

$\lambda_{f,\beta}$ — fixed part of $\lambda_{s,\beta|\mu,b,\tau,\beta}$ (213)

$\lambda_{g,\beta}$ — random part of $\lambda_{s,\beta|\mu,b,\tau,\beta}$, constant within each g-group (213)

$\lambda_{\tau|\mu,b,\tau,\beta}$ — λ for the adjusted treatments (64)

$\lambda_{s,\tau|\mu,b,\tau,\beta}$ — $\lambda_{\tau|\mu,b,\tau,\beta}$ for randomization s (213-214)

$\lambda_{f,\tau}$ — fixed part of $\lambda_{s,\tau|\mu,b,\tau,\beta}$ (215)

$\lambda_{s,\tau}$ — random part of $\lambda_{s,\tau|\mu,b,\tau,\beta}$ (215)

$\mathbb{M}$ — any matrix such that $\mathbf{X}_a=\mathbf{X}_b\mathbb{M}$ (77, 78, 87)

$\mathbb{M}_s$ — permutation matrix with elements (0,1) used generate $\mathbf{X}_s$ for randomization s (185-186)

$\mathbb{M}_{s,i}$ — diagonal block i of $\mathbb{M}_s$ (186)

$\mathfrak{m}$ — set of all permutation matrices (187)

μ — model equation mean for matrix (19), summation (23), identity (123)

μ^* — mean in usual summation model equation (22)

n — number of experimental units (16) numerator degrees of freedom for F ratio (39)

n_s — vector of normally distributed random errors for randomization s (199)

$N(\mathbb{B})$ — number of block-wise randomizations (108, 115)

Symbol   Description   (page of first use)

$N(\mathcal{E}_{j1:k})$ number of sets in $\mathcal{E}_{j1:k}$ (111)

$N(\mathcal{E})$ number of experiment-wise randomizations (108)

$N(\mathcal{G})$ number of groupings of experimental units into
             t groups of r experimental units each (107, 115)

$N_n(\cdot,\cdot)$ n-dimensional normal distribution (16)

$\mathbb{N}$ matrix, normally distributed random errors (155)

$\mathbb{N}_s$ column of $\mathbb{N}$ for randomization s (157)

OSL   observed significance value, the p-value (170)

$OSL_{R,@s}$ randomization OSL for experiment s (170-171)

$OSL_{R(F)}$ randomization OSL for simulation (175)

$OSL_{R,@s}$ randomization OSL for experiment s (170)

$p(ij1:k)$ probability of unit ij1 receiving
             treatment k – the assignment probability (104)

$p(s)$ probability of selecting randomization s –
             the selection probability (109)

p   rank of $\mathbf{X}$, the design matrix (16)

$\mathbb{P}$ $E_R(\mathcal{G})$, the $p(s)$ in column vector form (157)

$\mathbb{P}_{\cdot,\cdot}$ the $\cdot,\cdot$ block of the projector $\mathbf{X}(\mathbf{X'X})^{-}\mathbf{X'}$ (195)

q   rank of $\mathbb{Z}$, the covariates (18)

R_s coefficient of determination R^2
             for randomization s (223)

$R(\cdot|\cdot,\cdot)$ reduction in sum of squares (Appendix B

$R_s(\cdot|\cdot)$ $R(\cdot|\cdot)$ for randomization s (203)

$\mathbb{R}$ type one elementary matrix (187)

$\mathbb{R}^n$ n-dimensional set of real numbers (16, 142, 157)

$\mathbb{R}_X$ $\mathbb{I}-\mathbf{X}(\mathbf{X'X})^{-}\mathbf{X}$, $\mathbf{X}$ has μ, b, and τ columns (26)

$\mathbb{R}_Z$ $\mathbb{R}_X\mathbb{Z}$ when $\mathbf{X}$ has mean, blocks and treatments (26)

Symbol	Description (page of first use)
$\mathbb{R}_3$	$\mathbb{R}_Z$ when Σ has only the mean and blocks (31)
S_{ij_1}	stage one response (233)
SSB_{yy}	sum of squares for blocks for Y (35)
SSM_{yz}	sum of cross products for the mean for YZ (35)
SSR_{yy}	summation sum of squares for residuals (35)
SSR	matrix sum of squares for residual (36)
SSR_s	SSR for randomization s (204)
SST_{yy}	sum of squares for treatments (35)
$SStot$	total sum of squares (35)
$\mathbb{S}$	selection vector - one element=1, rest=0 (156)
$\mathbb{S}$	set over which to compute the OSL, $\hat{p}_R$ or $\hat{p}_{R(F)}$ (170, 173, 175, 176, 177)
σ^2	variance (16)
t	number of treatment levels (20)
$t_{@s,s}$	test statistics for randomization s in the randomization test conducted @s (170)
$\mathcal{T}_{@s}$	set of all $N(\Xi)$ $t_{@s,s}$ (170, 172)
$\mathcal{T}$	set of all $N(\Xi)$ test statistics for a simulated randomization test (174)
$\mathbb{T}$	matrix of model parameter values (152)
τ	vector of parameters: blocks, treatments (19)
τ_k	effect of treatment k (22), averaged over all units when algebraically defined (123)
τ_k^K	treatment effect equal for all units (126)
$\hat{\tau}_{\mu,b,\tau,\beta}$	$\hat{\tau}$ when u, b, τ, and β are in the model (30)
$\hat{\tau}_g$	covariate g's estimate of τ - model Σ (27)

Symbol	Description (page of first use)
$\mathcal{U}_1$	set of all bpt unit labels (142)
$\mathcal{U}$	the Cartesian product $\mathcal{U}_1$ x ... x $\mathcal{U}_{bp}$ (142, 157)
$Var_R(\cdot)$	randomization variance (251)
Ψ	variance-covariance matrix (16)
$V_{ij k}$	$S_{ij\_1} - \beta Z_{ij}$, the response adjusted for the covariate (234)
ψ	element of $\mathcal{U}$, one randomization (142-143)
W_{ijkl}	stage two model response (235)
$\mathbb{W}$	matrix $\equiv (1/t)\{diag[\mathbb{I}_{txt}]\}_{bxb}$ (46, 56)
X^-	any generalized inverse of X (83)
X^+	model matrix including covariate(s) (16), the Moore-Penrose inverse of X (83)
X	design matrix (19); a matrix of all $N(\Xi)$ possible X for all randomizations (153)
X_μ	the μ part of X, a vector of ones (19)
X_b	the block part of X (19)
X_τ	the treatments part of X (19)
X_c	coding version of X (80-81)
X_e	effects version of X (80-81)
X_h	elements of X^+ being tested (42)
X_r	elements of X^+ other than those of X_h (42)
X_s	X with X_τ arranged as per randomization s (185)
$X(X'X)^- X'$	the projector (matrix) of X (75)

Symbol	Description (page of first use)
Y_{ik}	response of unit i with treatment k (22)
Y_{ijk}	response when unit ij receives treatment k (123)
$Y...$	Y_{ijk} averaged over all i, j, k (123)
$\mathbf{Y}$	vector of observations (16), the data (144) matrix of all possible observations(152-152,156)
$\mathbf{Y}_s$	column of the matrix $\mathbf{Y}$ for randomization s (157)
$\tilde{\mathbf{Y}}$	vector of all possible responses Y_{ijk} (144)
Z_{ik}	covariate value of unit i and treatment k (22)
$\mathbf{Z}$	matrix of continuous covariates (19)

CHAPTER I

INTRODUCTION

Analysis of covariance combines elements of two
classical data analysis techniques, analysis of variance,
and linear regression. The general linear model builds
from this synthesis. Typically, the errors of such
models are assumed to be normally distributed, and
detached from the experimental units used to conduct the
experiment. Errors attached to the experimental units
are possible and arise in experimental units which are
biological in origin. With such errors, the assignment
of experimental units to treatments becomes important.
In extreme cases, the assignment chosen may determine the
outcome of the analysis of the experimental results.

Some results are known for models without a
covariate whose errors are attached to the experimental
units. The randomization test is one example. Another
is that in the randomized block design, when the treat-
ments are zero, the mean of the F-ratio, as averaged
over all randomizations, equals the first moment of the
central F distribution. Some asymptotic and simulation
results are in the literature, as are many papers on
regression models without a design component.

Results available for models with attached errors combining design and regression components, such as the analysis of covariance, are few and limited to asymptotic results.

This work extends what is known about such models by examining the randomized block design with a single covariate. The work proceeds in three steps.

The first, Chapter Two, reviews the typical linear model, laying out the assumptions, and the estimators and tests in matrix and in summation notations. Some new results are derived on the noncentrality parameters of the hypotheses of interest and on the associated projector matrices. A new proof clarifies the relationship between matrices with equal projector matrices. A brief review of the origins of the linear model concludes Chapter Two.

The second step develops the model using attached errors, the randomization model. Chapter Three derives a method to count the number of randomizations flexible enough to handle such restrictions as balancing on covariates. It replaces the traditional assignment probability and delta summation notation with selection probability and matrix notation. Two classical types of attached error terms are reviewed and related to each other. The randomization model, including its probability space, is developed, then generalized to include normally distributed random error. A flexible matrix

representation aids this generalization. A section
expands the definition of the randomization test and
randomization power computation to include unequal
selection probabilities. A second matrix notation based
on permutation matrices is developed. Results are
derived for projector matrices and noncentrality
parameters operated on by such permutation matrices. The
clustering of values of some estimates and of some
analysis of variance terms under the randomization model
is proved. A new result links the noncentrality
parameters of the usual hypotheses of interests for the
case of a single covariate.

The final step is the simulation studies of Chapter
Four. A novel method permits generating the two types of
attached errors. Means and standard deviations of
estimators and analysis of variance terms are compared,
and their histograms are displayed. The F distribution
analog of normal probability paper is developed and used
to contrast the two attached errors with the usual
normally distributed random error. The effect of unequal
within-block error variances is investigated. Examined
is a weighted randomization suggested by D. R. Cox, as is
restricting randomization to large values of the non-
centrality parameter for the adjusted covariate. Graphs
and displays illustrate results derived in Chapter Three.
Two conclusions are (1) the usual F-distribution is a
reasonable approximation to the randomization distribu-

tion for experimental unit-treatment interaction error, but not for experimental unit error and (2) restricting randomization to randomizations with the maximum value of the noncentrality parameter for the adjusted covariate improves estimates and hypothesis tests.

The permutation matrix techniques, probability space, and computer programs herein developed allow straightforward investigation of more complicated models, for example, models with multiple covariates or factor interaction in the design part, models combining types of errors, or models with missing values. Easy extensions are combining types of errors and examining different values for the treatment effects, error variances, and covariate. Appendix E collects and discusses the extensions mentioned in the text. Several computer programs, written in SAS, and reproduced in Appendix F, may be modified for further simulation work.

Principle contributions of this work include a compilation of terminology and known results presented in a common notation; bibliographies of relevant topics; development of new matrix-based methods; new results in the clustering of Anova values and estimators, and in the relationship between the noncentrality parameters of the usual hypotheses, and a new proof of a relationship of projector matrices; new methods for numerical simulation and analysis of their results; and the results themselves.

CHAPTER II

THE NORMALLY-DISTRIBUTED RANDOM ERRORS MODEL

Section 2.1

Introduction

Chapter Two presents the normally-distributed random errors model for the analysis of covariance. Section 2.2 is an overview of the technique and literature review. Sections 2.3 - 2.6 review the model, the estimators, tables of sums of squares and tests of hypotheses. A numerical example comprises Section 2.7. Section 2.8 proves a theorem on projectors and has a brief literature review of generalized inverses. Section 2.9 is a summary of the early concepts of what is today the error term in the model equation; it concludes Chapter 2.

This chapter introduces the notation and develops the "Model Σ", the projectors and the non-centrality parameters for the hypotheses of interest. Chapter Three builds upon and contrasts the results of this chapter. We begin by introducing some terminology in the context of two examples. The first appearance of a term, or when it is defined, is noted by underscoring the term and, when several pages are referenced, underscoring this page number in the index.

Section 2.2

Overview of the Analysis of Covariance

Two hypothetical experimental situations introduce
the use and terminology of the analysis of covariance in
a blocked design. Of importance are the two hypotheses
of interest. A literature review completes this section.

2.2.1 Two Motivating Examples

Scientific experiments examine the effects of
different conditions or variables upon a measurable
response of a system or outcome of an experiment. When
the conditions or variables are nominal values labeling
the types of conditions being examined, then the vari-
ables are termed factors. When the variables take values
which may be continuous along the real number line, then
the variables are termed covariates. The techniques of
analysis of covariance permit combining both types of
variables in the same theoretical framework.

A typical use of analysis of covariance in a design-
ed experiment is the study of various feed additives as
cattle growth promoters. Analysis of covariance
considers the final weight of the animal as the response
variable, the various additives as discrete levels of one
factor and initial animal weight as a continuous covar-
iate. Cattle breeds are a second factor, treated as a
blocking variable, with one level per breed. This

designed experiment focuses on the effects of the addi-
tives, a factor, upon final animal weight, the response.
Within each breed or level of the blocking variable, the
animals are assigned to each feed additive at random.
The blocking variable and the covariate are not of
interest, but are included in the analysis to remove
sources of variation that are costly or impossible to
remove prior to the experiment. As the assignment of an
animal to a feed additive is within the control of the
experimentor, this is a _designed_ _experiment_.

An example of an observational study is an exami-
nation of wages and job seniority. Within each of
several cities, the wage earners in designated sectors of
the economy are sampled at random and found to have
various years of seniority and wages. The analysis of
covariance considers the wages as the dependent variable,
sectors of the economy as levels of one factor and
seniority as a continuous covariate. Each city is a
level of a second factor, treated as a blocking variable.
This study focuses upon the effect of seniority, the
covariate, upon wages. The two factors of cities and
sector of economy are not of interest, but are included
to remove variation in the obtained sample. As the
assignment of a wage earner to a sector of the economy is
not under the control of the experimentor, this is an
observational _study_, not a designed experiment [Hocking,
1982, p. 560; Cox and McCullagh, 1982, p. 541]. The term

dependent variable, often used in place of response

variable, gives the impression that the causal variables

are known to be the blocks, factors and covariates under

study. Seldom is this the case. For this reason, we use

the term response throughout.

A treatment is a "combination of stimuli or operations

[factor levels] imposed by the experimenter" [Wilk, 1957,

p.220]. An example is that of feed additives. When there

is but one factor, each level of the factor is also a

treatment. An experimental unit is the largest group of

experimental material to which a treatment is assigned in

a single trial of the experiment [Snedecor and Cochran,

1967, Ed. 6, p. 15]. In this example, each animal is an

experimental unit.   Observational studies consider the

discovered "treatment" as having been assigned to their

randomly selected subjects (experimental units). An

example is the sector of the economy "assigned" to each

wage earner.

A typical designed experiment assigns a discrete

factor level to the experimental unit, but uses as a

covariate a continuous variable permanently affixed to

the experimental unit. However, the reverse may hold;

the continuous variable may be assigned and the discrete

factor level may be permanently affixed [Holland, 1986,

p. 945]. Throughout we assume that the continuous vari-

ables are permanently affixed to the experimental units

and measured without error and that the discrete factor
levels are assigned at random to the experimental units.

Technically speaking, a blocking variable is a
factor; like a factor, it has levels, but for blocking
variables, the levels typically are termed "blocks."
Random assigning of treatments to experimental units
takes place within each block of the blocking variable.
Hence, the experimental design is termed a completely
randomized block design, or, equivalently, a randomized
block design. A second difference between factors and
levels of the blocking variable is that the experiment
focuses on the factors, not on the blocking variable(s).
Blocking variables are intended to isolate sources of
variation which could not be eliminated from the
experiment or sampled population.

The hypothesis of interest in the first example is
that of the treatments (feed additives) adjusted for the
mean, blocks and covariate (initial animal weight). The
hypothesis of interest in the second example is that of
the covariate (seniority) adjusted for the mean, blocks,
and treatments (sector of the economy). The term
"adjusted for the blocks" replaces the more lengthly
"adjusted for the levels of the blocking variable(s)."
The more general experimental situation may have multiple
blocking variables; multiple factors, each with its own
number of levels; interactions among factors; multiple

covariates; interaction among the covariates; covariate
coefficients differing among the levels of one or more
factors; and (infrequently) interactions among the fac-
tors and the covariates. This work examines, in the main,
only the most elementary of these cases, that is, one
blocking variable, one factor and one or more covariates.

2.2.2 Uses of Analysis of Covariance

Cochran [1957, p. 262-267] further discusses this
list of the most important uses of analysis of covariance

1. To increase precision in randomized experiments.

2. To remove the effects of disturbing variables in
    observational studies.

3. To throw light on the nature of treatment effects.

4. To fit regressions in multiple classification.

5. To analyze data when some observations are missing.
A similar list is in Steel and Torrie [1980, p. 401]. The
two hypotheses of interest discussed in Section 2.2.1
illustrate, respectively, uses three and two. This paper
does not discuss uses four or five. Cox and McCullagh
[1982, p. 541-542] offer a more technical list, with com-
ponents of covariance and canonical regression analysis.

2.2.3 Literature Review - Analysis of Covariance

Most published works on the analysis of covariance
use the normally-distributed random errors model. This
model assumes an infinite population of random errors.
Older textbooks and papers use summation notation.

Kempthorne [1952, p. 98-103] shows the derivation of the
equations, in summation notation, used to analyze a
randomized block design with one factor of interest and
one covariate. Cochran [1957] provides a clearly-written
derivation of formulas and review. Ostle's numerical
examples [1963, Ch. 14, p. 410-465] detail the use of a
covariate in several experimental designs, including the
randomized block design with one covariate. Snedecor and
Cochran [1967, Ch. 14, p. 419-446] detail the numerical
calculations for one- and two-way classification designs
with one and two covariates. Wildt and Ahtola [1978, p.
58-69] provide numerical examples of several designs,
including the randomized block design with one factor of
interest and one covariate. Steel and Torrie [1980, Ch.
17, 401-437] have a good discussion of the uses and
assumptions of the analysis of covariance. Numerical
examples show the use of one and two covariates in a
randomized block design. Ott [1988, Ch. 18, p. 676-681]
casts the treatment variables as zero-one dummy variables
and discusses the computer output from such a regression-
based analysis. DeLury [1948] provides a real-data
example of a treatment affecting a covariate measurement.

More modern and advanced publications use matrix
notation instead of the older summation notation. The
introductory text by Neter, Wasserman and Kutner [1990,
Ch. 23, p. 861-906] includes numerical examples with cal-
culations for one- and two-way classifications with one

covariate. They present sufficient matrix algebra for
the calculations. The advanced treatment of Searle
[1987, Ch. 11, p. 416-456] is one of, if not the, most
extensive textbooks on this topic. He compares the sum-
mation and matrix notations, centered and noncentered co-
variates and various hypotheses. Included are the mean
model and one- and two-way classifications, both with
one and two covariates, but without numerical computa-
tions. Other chapters discuss unequal cell counts, the
mean model, interpretation of output from several compu-
ter packages and other textbooks [p. 242]. Searle's
earlier text [1971] provides an extensive theoretical
development of the general linear model, including the
overparameterized, non-full-rank case. This earlier work
discusses the analysis of covariance [p. 240-361] and
includes a numerical example of the computations for a
one-way classification with one covariate. Seber [1977,
Ch. 10] uses both summation and matrix notation for the
randomized block design. He discusses the results of Hsu
[1938] and Atiqullah [1964] applicable to quadratic forms
with vectors not normally distributed. A second work by
Seber [1980, p. 61-65] discusses the nested nature of the
hypotheses of interest. He uses projector matrices and
intersections of linear spaces for much of his theoreti-
cal development, as does Scheffe [1959]. Graybill [1976]
thoroughly develops the general linear model using matrix
notation, but his discussion of the analysis of covariance

is brief [p. 283-297]. C. R. Rao [1973, p. 288-294] has
a brief discussion of the analysis of covariance using
matrix notation and a numerical example. Morrison [1983,
p.122-131] concentrates upon regression, but also
discusses the noncentrality parameter. He includes an
example of sample size determination. C. R. Rao [1975,
p. 475-487] discusses the case of a variance-covariance
matrix which is not positive definite, as does Searle
[1971, p. 221-224].

Searle [1982] and Graybill [1983] provide matrix
algebra textbooks designed for students of statistics.
Either textbook presents a solid foundation for the
advanced texts mentioned above.

Beyond such texts there is a vast literature. Two
entire issues of *Biometrics* [Vol. 13, 1957 and Vol. 38,
1983] discuss the analysis of covariance. Cox and
McCullagh and discussants [1982] provide a review inclu-
ding a brief history of the technique. They use matrix
notation. A comment by Hocking [1982, p. 558-561] adds
the use of the hat matrix [Hoaglin and Welsch, 1978] as a
diagnostic tool and cautions about the interpretation of
computer-provided estimated treatment effects. Herein,
the hat matrix is termed the projector of the matrix X.
Henderson [1982] details the assumptions of the normally-
distributed random-errors model [p. 623-624] and
discusses the more realistic case of a mixed model; our
model equation parameters are assumed to be fixed.

Bingham and Fienberg [1982] discuss two numerator sums of squares in the F-test for adjusted treatment effects when block x treatment interaction terms are considered nonsignificant. Their hierarchical method places them in the treatment sum of squares. This paper uses the usual method and enters this interaction as the, or as part of the, error term. When there is but one experimental unit per treatment per block, the block x treatment interaction is the error term. When there are multiple experimental units per treatment per block, the block x treatment interaction may be tested for statistical significance. This paper does not consider such tests.

   We turn now to the model for the analysis of covariance, using both matrix and summation notation for the model equation.

Section 2.3

The General Linear Model for

the Analysis of Covariance

This section presents the model, that is, the model
equation, the assumptions and the probability space, for
the analysis of covariance.

The first sub-section begins with the general linear
model in matrix notation, then restricts it by stages
until achieving a model similar to the usual summation
notation model for the case of one blocking variable, one
factor and one covariate. The second sub-section
presents the model equation in summation notation.

The matrix notation is adopted from Searle [1971,
p. 340-361] and Cox and McCullagh [1982, p.543].
Henderson [1982, p. 623-624] presents a detailed list of
assumptions, most of which are incorporated in
Assumptions 1 - 8, presented in Section 2.3.

Of primary importance are the matrix-notation model
equation and partitioning of (2.3.4) and (2.3.5), (p.17),
the summation-notation model equation of (2.3.8), (p.23),
and the discussion ending Section 2.3.2.

2.3.1 The Model in Matrix Notation

Beginning with the general linear model, define

n   the number of experimental units,

Y the n x 1 vector whose elements are

yields or responses,

p'   the number of columns of X^+, p' < n,

X^+ the n x p' model matrix whose

elements are explanatory variables,

β^+ the p' x 1 vector of model equation

parameters and

ε the n x 1 vector whose elements are

random errors.

The most general model equation is

$$Y = X^+\beta^+ + \varepsilon .$$ (2.3.1)

At this stage the assumptions are Assumptions 1-4[*],

1. The model equation is correct, that is,

the model matrix X^+ contains the correct

variables and/or the proper functions of these

variables, such as logs or squares,

2. The errors are distributed normally,

each with mean 0,

3[*]. The errors have variance-covariance

matrix $\Psi\sigma^2$; with σ^2 unknown, and Ψ (nxn),

known and positive definite,

4[*]. The elements of X^+ are known constants,

measured without error.

Assumptions 3^* and 4^*, marked with an asterisk, are modified below.

The <u>probability</u> <u>space</u> for the random variable ε is the triple, $(\ \mathbb{R}^n,\ \mathbb{B}^n,\ N_n(\mathbb{0},\Psi\sigma^2)\ )$,   (2.3.2)

with $\mathbb{R}^n$ the n-dimensional set of real numbers, $\mathbb{B}^n$ the n-dimensional σ-field of Borel sets of real numbers and $N_n(.,.)$ the n-dimensional multivariate normal distribution.

When $\mathbf{X}^+$ lacks full column rank the generalized inverse $(\mathbf{X}^{+\prime}\mathbf{X}^+)^-$ is not unique; thus, neither is the <u>solution</u>, $\beta^{+\mathrm{o}}$, to $(\mathbf{X}^{+\prime}\mathbf{X}^+)^-\mathbf{X}^{+\prime}\mathbf{Y}$. When $\mathbf{X}^+$ has full column rank, the unique solution is termed an <u>estimator</u> and denoted $\hat{\beta}^{+\mathrm{o}}$. The estimator $\hat{\beta}^{+\mathrm{o}}$, or a solution $\beta^{+\mathrm{o}}$, [Searle, 1971, p. 80 and 169] defines $\hat{\varepsilon}$, the estimator for ε, which is the n x 1 vector of <u>residuals</u>

$$\hat{\varepsilon} = \mathbf{Y} - \hat{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}^+\hat{\beta}^{+\mathrm{o}} = \mathbf{Y} - \mathbf{X}^+\beta^{+\mathrm{o}}\ . \qquad (2.3.3)$$

The last equality follows from the discussion below (2.8.5), (p. 76). The distribution of the residuals, $\hat{\varepsilon}$, is the singular normal distribution. The usual estimators of ε and $\mathbf{Y}$ and a result in generalized inverses [Searle, 1971, p. 20], shows $\hat{\varepsilon}$ to have a form which characterizes the <u>singular</u> <u>normal</u> <u>distribution</u>.

The following form characterizes the singular normal

distribution. For clarity, X temporarily replaces X^+.

$$\hat{\varepsilon} \equiv Y - \hat{Y} = X\beta + \varepsilon - X\hat{\beta}$$

$$= X\beta + \varepsilon - X[(X'X)^-X'Y]$$
$$= X\beta + \varepsilon - X(X'X)^-X'[X\beta + \varepsilon]$$
$$= X\beta + \varepsilon - [X(X'X)^-X'X]\beta - [X(X'X)^-X']\varepsilon$$
$$= X\beta + \varepsilon - X\beta - [X(X'X)^-X']\varepsilon$$
$$= [I - X(X'X)^-X']\varepsilon$$
$$= 0 + [I - X(X'X)^-X']\varepsilon .$$
$$= 0 + L\varepsilon, \text{ where } L \equiv [I - X(X'X)^-X'].$$

When ε has the normal distribution with mean zero and

variance $\sigma^2 I$, and LL' lacks full rank, this form charac-

terizes the singular normal distribution, as given by

Anderson and cited by Searle [1971, p. 66-72]. Since L

is idempotent and $L \neq I$, L lacks full rank; thus, so does

$L'L$. Thus, $\hat{\varepsilon}$ has the singular normal distribution here

with mean of zero and variance-covariance matrix of $LL'\sigma^2$.

To partition the model matrix X^+ into an nxp <u>design</u>

or <u>incidence</u> matrix X, with discrete elements and an nxq

matrix of covariates Z, with continuous elements, replace

Assumption 4[*] with

<u>Assumption 4</u>

   The columns of the design matrix X are causally

   and linearly independent of the columns of the

   matrix of covariates Z; all elements of both

   matrices are known constants, measured without

   error. The matrix Z has full column rank, q.

See Henderson [1982, p. 623-624] for a discussion

of a mixed model when Assumption 4 does not hold.

This partitioning gives

$$X^+ = \begin{bmatrix} X & | & Z \end{bmatrix} , \quad \beta^+ = \begin{bmatrix} \tau \\ \beta \end{bmatrix} \text{ and the model equation}$$

$$Y = X\tau + Z\beta + \varepsilon . \tag{2.3.4}$$

The randomized block design further partitions X and τ

as $\quad X = \begin{bmatrix} X_\mu & | & X_b & | & X_\tau \end{bmatrix}$ and $\tau = \begin{bmatrix} \mu \\ b \\ \tau \end{bmatrix}$. $\tag{2.3.5}$

Note the reuse of τ as the treatments part of the vector

of design or incident coefficients. The context will

make clear the meaning of τ. The matrix X_μ is the n x 1

column vector of ones, $\mathbb{1}$. The matrices X_b and X_τ are the

design or incident matrices for, respectively, the block

and treatment effects; they are assumed to be constructed

so as to be orthogonal to each other and to X_μ.

This paper considers the treatments as having been

relabeled into one new factor with a number of levels

equal to the number of treatments. Discussion of the

original factors and their interactions is an extension.

Much of the following restricts the model to one

having full column rank in each of the partitions of X,

X_b and X_τ. As discussed in Section 2.8.2, (p. 76),

below, many such restrictions are possible. For compar-

ability with the summation notation and ease of interpre-

tation, we select the "usual constraints," which are

usual and useful only in limited cases such as this

balanced design [Searle, 1971, p.209-220]. Consideration

of unbalanced models, such as the unbalanced mean model,

is an extension. Limitation to this effects model gives

   Assumption 5,

          Restrict the model so that $\mathbb{1}'b = \mathbb{1}'\tau = 0$,

       and constrain the estimators so that

       $\mathbb{1}'\hat{b} = \mathbb{1}'\hat{\tau} = 0$. For τ and $\hat{\tau}$ these sums

       hold within each block. Require an equal

       number of observations for each treatment

       within each block.

With b levels of the blocking variable and t treatments,

Assumption 5 results in X_b being n x (b-1), X_τ being n x

(t-1) and X having rank p = 1 + (b-1) + (t-1). The

vectors b and τ have dimensions, respectively, (b-1) x 1

and (t-1) x 1. Each element of X_b and X_τ is one of 0,

1, or -1. The partitions of (2.3.5) are mutually

orthogonal; in particular, $X_b'X_\tau = \mathbb{0}$. These are the

effects versions of X_b and X_τ, as in the left-hand matrix

of (2.8.2), (p. 74).

   A further limitation on the models considered is

       Assumption 6:

          The covariate slope coefficient(s) of β

       have the same value for each treatment.

This assumption details Assumption 1 by excluding the

case of a separate slope for each level of the treatment.

The randomized block design adds Assumptions 7 and 8,

Assumption 7:

There is no interaction between the blocks
and the treatments. Any which appears in the
estimation procedures is considered to be
residual error.

This assumption provides an error term when there is but
one observation per treatment per block. When there are
multiple observations, the presence of block x treatment
interaction may be tested, but is not in this paper. If
the interaction is not statistically significant, it
typically is considered as part of the residual error.
See Bingham and Fienberg [1982] for an argument to
consider it as part of the treatment sum of squares.

Assumption 8:

Randomization occurs independently within
each level of the blocking variable.

Thornett [1982, p. 138] posits that the assignment of
experimental units to treatments should be independent of
the responses. He then proves that a procedure involving
equally likely (uniform probability) randomization is
virtually essential in obtaining such independence [his
Theorem 3, p. 140]. The only role that randomization
plays in the normally distributed random errors model is
that randomization necessarily follows from this sound
experimental practice. Hooper [1989, p. 578-579]
suggests an alternative conditional interpretation of the

usual hypotheses using a weaker condition.

The normally distributed random errors model is conditioned on the randomization used to conduct the experiment. However, this model treats the errors as detached from and unaffected by the experimental units and/or the treatments they receive. Thus, the conditional distribution of the errors, given the randomization, reduces to the distribution of the errors alone.

2.3.2 Summation Form of the Model Equation

The experimental design having one blocking variable, one factor, one covariate and one experimental unit per treatment has the model equation, in summation form, of

$$Y_{ik} = \mu^* + b_i + \tau_k + \beta(Z_{ik} - \bar{Z}_{..}) + \varepsilon_{ik} \qquad (2.3.6)$$

with $i = 1,\ldots,b$ levels (or blocks) of
                    the blocking variable,

$k = 1,\ldots,t$ levels (or treatments)
                    of the single factor,

μ^* = an overall mean,

b_i = the block effect for block i,

τ_k = the treatment effect for level k,

β = covariate slope coefficient,
          constant for all treatments τ_k,

Z_{ik} = covariate of the experimental unit
          in block i assigned to treatment k,

$\bar{Z}_{..} = (1/bt) \sum_i^b \sum_k^t Z_{ik},$ and

ε_{ik} = error for the experimental unit
          of block i and treatment k.

Assumption 5, in summation notation, is

5.   Restrict the model so that

$$\sum_{i=1}^{b} b_i = 0 \quad \text{and} \sum_{k=1}^{t} \tau_k = 0,$$

and constrain the estimators so that

$$\sum_{i=1}^{b} \hat{b}_i = 0 \quad \text{and} \sum_{k=1}^{t} \hat{\tau}_k = 0,$$

and assign the same number of experimental

units to each treatment in each block.

Subsequent numerical examples use one experimental unit

per treatment per block.

The standard matrix notation does not center the

covariate variable about its mean.   By redefining the

mean of the summation model equation as

$$\mu = \mu^* - \beta \bar{Z} \qquad (2.3.7)$$

the model equation of (2.3.6) becomes

$$Y_{ik} = \mu + b_i + \tau_k + \beta Z_{ik} + \epsilon_{ik} . \qquad (2.3.8)$$

It is the μ of (2.3.8), not the μ^* of (2.3.6), which

corresponds to the μ element of (2.3.5), (p. 19).   See

Searle [1987, p. 438-441] for further discussion of the

differences between the matrix and summation model

equations (2.3.5) and (2.3.6).

Let ij denote the $j^{\underline{th}}$ experimental unit in the $i^{\underline{th}}$

block.   The covariate values Z_{ij} are attached to the $ij^{\underline{th}}$

experimental unit, but, for one experimental unit per

treatment and after the assignment of treatments, the

covariate values may be identified by the treatment

applied.   Thus, the literature and (2.3.8) use Z_{ik}

instead of Z_{ij}.   For any number of experimental units per

treatment, the covariates are to be combined on the basis

of the treatments they receive, thus, the ik subscripting

is more convenient.   (See Section 2.5, (p. 33), and

(2.4.29), (p. 32).)

The summation model considers the ε_{ik} to be

independently distributed with identical variances.   For

this, Assumption 3^* becomes

Assumption 3.   The errors have variance

covariance matrix $\mathbb{I}\sigma^2$ with σ^2 unknown.

We adopt this assumption for the matrix model equation.

Use of the more general variance-covariance matrix Ψ is

an extension.

Note that Assumption 8, independent randomization

within each level of the blocking variable, is nowhere

used in the model equation, in the probability space, or

in other assumptions.   Also, the label identifying the

experimental unit, ij, has everywhere been replaced by

ik.   The identity of the individual experimental unit has

been lost; only its block and treatment membership is

retained.   These two aspects set off the normally

distributed random errors model from the model of Chapter

Three.

We next examine the matrix and summation formulas

for the estimators of the model equation parameters.

Section 2.4

Estimators

Section 2.4 presents the estimators typically of interest in the analysis of covariance. The first sub-section sketches their derivation in matrix notation. The second develops the "model Σ". The final sub-section displays the estimators in summation notation.

Of importance are the model Σ residuals (2.4.10), (p. 28), the product of (2.4.13b), (p. 29), the estimators of (2.4.19-2.4.22 and 2.4.24), (p. 30-31), and their summation versions in (2.4.27-2.4.29), (p. 32).

2.4.1 Sketch of Derivation in Matrix Notation

One task of the experiment is to estimate functions of the elements of the vectors τ and β and the statistical precision of these estimates, which requires estimating σ^2. Initiated by Legendre [1806], Adrian [1808] and Gauss [1809], the method of least squares provides the minimum variance unbiased estimators for τ and β under the normally distributed random errors model. These estimators follow from the expectation of (2.3.4), (p. 19),

$$E_N(Y) = \Sigma\tau + \mathbb{Z}\beta , \qquad (2.4.1)$$

by the usual calculus method of minimization. That is, one equates the first derivative of the sum of the squared errors, $\varepsilon'\varepsilon$, $\dfrac{\partial(\varepsilon'\varepsilon)}{\partial\begin{bmatrix} \tau \\ \beta \end{bmatrix}}$, to zero and shows the

second derivative (equal to the leftmost bracketed term

of (2.4.2)) to be positive definite. Assumptions 4 and 5

ensure that $[X \mid Z]$ is of full column rank, $p + q$.

Applying Corollary 1.4.2 (3) of Graybill, [1976, p. 22]

to $[X \mid Z]$ proves the second derivative to be positive

definite. When X does not have full column rank, it is

only the <u>estimable functions</u> of β^+ which have minimum

variance for any generalized inverse of X. See Searle,

1971, p. 182. The least squares equations corresponding

to (2.4.1) are

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z \end{bmatrix} \begin{bmatrix} \tau^0 \\ \beta^0 \end{bmatrix} = \begin{bmatrix} X'Y \\ Z'Y \end{bmatrix} \tag{2.4.2}$$

Multiplying by the inverse and solving directly for

τ and by substitution for β gives the solutions of the

least squares equation (2.4.2), which in the case of a

full rank X matrix, are also estimators. They are

$$\tau^0 = (X'X)^- X'Y - (X'X)^- X'Z \hat{\beta}^0 \quad \text{and,} \tag{2.4.3}$$

$$\hat{\beta}^0 = \{Z'[I - X(X'X)^- X']Z\}^{-1} \{Z'[I - X(X'X)^- X']Y\}$$

$$= \{Z'\mathbb{R}_X Z\}^{-1} \{Z'\mathbb{R}_X Y\} \tag{2.4.4}$$

$$= \{ [\mathbb{R}_X Z]'[\mathbb{R}_X Z] \}^{-1} [\mathbb{R}_X Z]'Y . \tag{2.4.5}$$

$$= \{ \mathbb{R}_Z' \mathbb{R}_Z \}^{-1} \mathbb{R}_Z' Y , \tag{2.4.6}$$

where $\mathbb{R}_X = I - X(X'X)^{-1}X'$ and $\mathbb{R}_Z = \mathbb{R}_X Z$. (2.4.7)

By the symmetry and idempotency of $\mathbb{R}_X$, $Z'\mathbb{R}_X Z =$

$(\mathbb{R}_X Z)'(\mathbb{R}_X Z)$, with $\text{rank}(Z'\mathbb{R}_X Z) = \text{rank}(\mathbb{R}_X Z)$ [Graybill,

1976, p. 6, Theorem 1.2.21]. Assumption 4, linear

independence of X and Z, (p. 18), ensures that $R_X Z$ has

full column rank q. Hence the rank of the q x q matrix

$Z'R_X Z$ is q and its inverse exists. Likewise, the full

column rank of Z ensures the existence of the inverse of

$Z'Z$. Thus $\hat{\beta}^0$, or for short $\hat{\beta}$, is the unique solution, so

denoted by a hat (^). Under Assumption 5, (p. 20),

$(X'X)^{-1}$ exists, making $\hat{\tau}^0$ the unique solution.

By (2.4.5), $\hat{\beta}^0$ is also the least squares estimator

for β in the model equation

$$Y = (R_X Z)\beta + \varepsilon = R_Z \beta + \varepsilon. \qquad (2.4.8)$$

2.4.2 The "Model Σ"

The "model Σ" provides an alternative to (2.4.2) and

is useful for the covariance model. See Searle [1971, p.

342-343], Cox and McCullagh [1982, p. 543-544] and

Zyskind, *et al.* [1964, p.139-141] for discussions of this

model. From each column of Z, Z_g, one obtains the

"residuals" one would obtain under the "model Σ"

$$Z_g = \Sigma \tau + \varepsilon , \quad g = 1, 2, \ldots , q.$$

For each of the q covariates, first obtain the least

squares estimate of Z_g, $\hat{Z}_g$,

$$\tau_g^0 = (\Sigma'\Sigma)^- \Sigma' Z_g , \qquad (2.4.9)$$

$$\hat{Z}_g = \Sigma \tau_g^0 = \Sigma[(\Sigma'\Sigma)^- \Sigma) Z_g] ,$$

and then the "residuals"

$$\mathbb{R}_{z,g} = \mathbb{Z}_g - \hat{\mathbb{Z}}_g = \mathbb{Z}_g - \mathbb{X}\tau_g^0 = \mathbb{Z}_g - [\mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}']\mathbb{Z}_g$$

$$= [\mathbb{I} - \mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}']\mathbb{Z}_g = \mathbb{R}_X\mathbb{Z}_g \ . \qquad (2.4.10)$$

Making an n x q matrix from the q vectors $\mathbb{R}_X\mathbb{Z}_g$ gives

$$\mathbb{R}_X\mathbb{Z} = [\mathbb{I} - \mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}']\mathbb{Z} = \mathbb{R}_Z \ . \qquad (2.4.11)$$

Because of Assumption 4, that the covariates of $\mathbb{Z}$ are causally unrelated to the treatment variables of $\mathbb{X}$, (p. 18), no one would actually compute these "residuals." Furthermore, the τ_g^0, differ with each of the g=1,2,...,q covariates, $\mathbb{Z}_g$. The line above equation (2.4.10) has the shape or *form* of the definition of residuals, as in equation (2.3.3), (p. 17), when the $\hat{\beta}^{+o}$ or β^{+o} of (2.3.3) is written as $(\mathbb{X}'\mathbb{X})^-\mathbb{X}'\mathbb{Y}$. These residuals, $\mathbb{R}_{z,g}$, are the *formal* residuals under the "model $\mathbb{X}$".

Next, to (2.4.1) add and subtract $\mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}'\mathbb{Z}\beta$, giving

$$E(\mathbb{Y}) = \mathbb{X}\tau + \mathbb{Z}\beta + \mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}'\mathbb{Z}\beta - \mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}'\mathbb{Z}\beta$$

$$= \mathbb{X}\tau + \mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}'\mathbb{Z}\beta + [\ \mathbb{I} - \mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}\ ]\mathbb{Z}\beta$$

$$= \mathbb{X}[\ \tau + (\mathbb{X}'\mathbb{X})^-\mathbb{X}'\mathbb{Z}\beta\ ] + \mathbb{R}_X\mathbb{Z}\beta$$

$$= \mathbb{X}\tau^* + \mathbb{R}_X\mathbb{Z}\beta$$

$$= \begin{bmatrix} \mathbb{X} & | & \mathbb{R}_X\mathbb{Z} \end{bmatrix} \begin{bmatrix} \tau^* \\ \beta \end{bmatrix} , \qquad (2.4.12)$$

where $\tau^* = \tau + (\mathbb{X}'\mathbb{X})^-\mathbb{X}'\mathbb{Z}\beta$. $\qquad (2.4.13a)$

The least squares equations for (2.4.12) are derived as were those of (2.4.1), (p. 25), and are

$$\begin{bmatrix} \Sigma'\Sigma & \Sigma'\mathbb{R}_X Z \\ Z'\mathbb{R}_X\Sigma & Z'\mathbb{R}_X Z \end{bmatrix} \begin{bmatrix} \tau^* \\ \beta \end{bmatrix} = \begin{bmatrix} \Sigma'\Upsilon \\ Z'\mathbb{R}_X\Upsilon \end{bmatrix} .$$

By the Corollary of p. 20 of Searle [1971],

$$\Sigma'\mathbb{R}_X = [\Sigma' - \Sigma'\Sigma(\Sigma'\Sigma)\!-\!\Sigma'] = [\Sigma' - \Sigma'] = 0,$$

as does $\Sigma'\mathbb{R}_Z$ and their transposes.   (2.4.13b)

Hence, the least squares equation for (2.4.12) becomes

$$\begin{bmatrix} \Sigma'\Sigma & 0 \\ 0 & Z'\mathbb{R}_X Z \end{bmatrix} \begin{bmatrix} \tau^* \\ \beta \end{bmatrix} = \begin{bmatrix} \Sigma'\Upsilon \\ Z'\mathbb{R}_X\Upsilon \end{bmatrix} .$$ (2.4.14)

The solution of (2.4.14) provides another estimate for τ. Typically, the analysis of covariance has multiple estimators (or solutions) for the model equation parameters τ and β depending upon what other parameters are in the model equation used to derive the estimator. To distinguish these various estimators, we subscript each estimator with the parameters used in the model equation which provided the estimation. The estimated parameter is always included in this equation, so the parameter being estimated is always repeated as a subscript.

Under Assumption 5, that $(\Sigma'\Sigma)^{-1}$ exists, (p. 20), the solution is unique and the estimators (in a shorthand notation) are

$$\hat{\tau}^*_\tau = \hat{\tau}_\tau = (\Sigma'\Sigma)^{-1}\Sigma'\Upsilon , \qquad \text{and} \qquad (2.4.15)$$

$$\hat{\beta}_{\tau,\beta} = \{Z'\mathbb{R}_X Z\}^{-1} \{Z'\mathbb{R}_X\Upsilon\} \qquad (2.4.16)$$

$$= \{\mathbb{R}_Z'\mathbb{R}_Z\}^{-1} \mathbb{R}_Z'\Upsilon \qquad (2.4.17)$$

$$= \hat{\beta}^0_{\tau,\beta} , \qquad \text{by (2.4.6).}$$

Substituting into (2.4.3) gives

$$\hat{\tau}^{o}_{\tau,\beta} = \hat{\tau}^{*}_{\tau} - (\Sigma'\Sigma)^{-1}\Sigma'Z\hat{\beta}^{o}_{\tau,\beta} \ , \quad \text{or in brief,}$$

$$\hat{\tau}_{\tau,\beta} = \hat{\tau}_{\tau} - (\Sigma'\Sigma)^{-1}\Sigma'Z\hat{\beta}_{\tau,\beta} \ . \tag{2.4.18}$$

The estimate for τ ignoring the covariates is, in the

more brief notation, $\hat{\tau}_{\tau}$, while $\hat{\tau}_{\tau,\beta}$ adjusts for the

covariates.

For the randomized block design using Assumption 5,

(p. 20), and the partitioning of (2.3.5), (p. 19), the

brief notation for the above estimators is

$$\hat{\tau}_{\mu,b,\tau} = (\Sigma'\Sigma)^{-1}\Sigma'Y \tag{2.4.19}$$

$$\hat{\beta}_{\mu,b,\tau,\beta} = \{Z'R_{X}Z\}^{-1}\{Z'R_{X}Y\} \tag{2.4.20}$$

$$= \{R_{Z}'R_{Z}\}^{-1}R_{Z}'Y \tag{2.4.21}$$

$$\hat{\tau}_{\mu,\beta,\tau,\beta} = \hat{\tau}_{\mu,b,\tau} - (\Sigma'\Sigma)^{-1}\Sigma'Z\hat{\beta}_{\mu,b,\tau,\beta} \tag{2.4.22}$$

The first, (2.4.19), estimates the treatment effects

after adjusting for the mean and the blocking variable(s)

only. The last, (2.4.22), estimates the treatment

effects after adjusting for the mean, the blocking

variable(s) and covariate(s).

A reduced model omitting the treatment parameters,

the elements of τ, estimates β adjusted for the mean and

blocks but not the treatment effects. This model

equation is

$$Y = [\Sigma_{\mu} \mid \Sigma_{b}]\begin{bmatrix} \mu \\ b \end{bmatrix} + Z\beta + \varepsilon \ . \tag{2.4.23}$$

As with (2.4.1), (p. 25), one may apply the "model X" to this model equation and compute a residual matrix corresponding to $\mathbb{R}_Z$, denoted $\mathbb{R}_{\mathfrak{Z}}$, which adjusts only for the mean and the blocks, not the treatments. The least squares estimator, $\hat{\mathbb{P}}_{\mu,b,\beta}$, is found as above and is

$$\hat{\mathbb{P}}_{\mu,b,\beta} = \{\mathbb{R}_{\mathfrak{Z}}'\mathbb{R}_{\mathfrak{Z}}\}^{-1}\mathbb{R}_{\mathfrak{Z}}'\Upsilon \ . \tag{2.4.24}$$

From (2.4.9), (p. 27), we see that $(\Sigma'\Sigma)^{-}\Sigma'\mathbb{Z}_q$ is the matrix with columns of τ_g^0, one per $\mathbb{Z}_q$, from the "model Σ". The $n \times q$ matrix $\mathbb{R}_Z$ is the residual from the multivariate model

$$\mathbb{Z} = \Sigma T + E \tag{2.4.25}$$

where $T = [\tau, \tau, \ldots, \tau]$,

(τ repeated q times), p x q, and

$E = $ the multivariate equivalent of ε.

Each column of T is estimated separately using one column of $\mathbb{Z}$, then recombined to form $\hat{T}$, giving

$$\mathbb{R}_Z = \mathbb{Z} - \Sigma\hat{T} \ . \tag{2.4.26}$$

2.4.3 Estimators in Summation Notation

The following equations compare the summation and matrix notations for the estimators of interest, equations (2.4.21), (p. 30), through (2.4.22), (p. 31). See Kempthorne [1952, p. 97-103] or Cochran [1957] for the derivation of these terms using summation notation.

The estimate of β adjusted for the estimated mean, blocks and treatment effects is

$$\hat{\beta}_{\mu,b,\tau,\beta} = SSR_{yz} \,/\, SSR_{zz} \,. \qquad (2.4.27)$$

Section 2.5 defines the righthand side of (2.4.27). Under Assumption 6, (constant β for all treatments), $\hat{\beta}$ is the <u>pooled</u> <u>within-class</u> <u>regression</u> <u>estimate</u>.

The estimate of τ_k without adjustment for the covariate is

$$\hat{\tau}_{k;\mu,b,\tau} = b^{-1}\sum_{i=1}^{b}Y_{ik} - (bt)^{-1}\sum_{i=1}^{b}\sum_{k=1}^{t}Y_{ik} \,. \qquad (2.4.28)$$

The estimate of τ_k adjusted for the estimated covariate is

$$\hat{\tau}_{k;\mu,b,\tau,\beta} = \hat{\tau}_{k;\mu,b,\tau} - \hat{\beta}_{\mu,b,\tau,\beta}(\bar{Z}_{\cdot k} - \bar{Z}_{\cdot\cdot}), \qquad (2.4.29)$$

$$\text{where } \bar{Z}_{\cdot k} = (1/b)\sum_{i=1}^{b}Z_{ik} \,.$$

The SSR_{yz} term of (2.4.27) corresponds to the matrix term $Z'R_zY$, and the SSR_{zz} term corresponds to the matrix term $Z'R_zZ$, both of (2.4.20). Equation (2.4.28) corresponds to the matrix notation of (2.4.19), (p. 30), substituting summation notation for the elements of $X'X$. The diagonal elements of $X'X$, once inverted, provide the divisors of (2.4.28). The product $X'Y$ provides the summations. Likewise, the product of $X'Z$ of (2.4.22) provides the summations for $\bar{Z}_{k\cdot}$ and $\bar{Z}_{\cdot\cdot}$ of (2.4.29), while the $(X'X)^{-1}$ provides the divisors.

Next we examine the sums of squares tables.

Section 2.5

Sums of Squares Tables

Section 2.5.3 presents analysis of variance tables, one in matrix notation, the other in summation notation. The first sub-sections discuss the notation used in the tables.

The analysis of covariance typically presents two analysis of variance tables, one for the treatments adjusted for the covariates, and one for the covariates adjusted for the treatments. Table 2.5.1 combines these two tables and shows in matrix notation the formulas used to compute the sums of squares column. Table 2.5.2 does the same for the summation notation. Both tables use the $R(.|.)$ notation discussed in Appendix B, (p.384). Table 2.5.1 follows Searle [1971, Tables 8.3a and 8.3b, p. 344-345]. Table 2.5.2 follows Searle [1971, Tables 8.4a and 8.4b, p. 351-352]; both tables use Ostle [1954, p. 394]. The two tables provide identical numerical values for the example of Section 2.7 and the simulations of Chapter 4.

2.5.1 Matrix Notation

The sum of squares due to treatments, plus the sum of squares due to covariates adjusted for the treatments (both adjusted for the mean and the blocks) is $R(t|\mu,b)$ + $R(\beta|\mu,b,t)$. This sum equals the sum of squares due to covariates, plus the sum of squares due to treatments

adjusted for the covariates (both adjusted for the mean

and blocks), $R(\beta|\mu,b) + R(t|\mu,b,\beta)$, so the two pairs have

the same sum. Typically, one finds the numerical value

of the term $R(t|\mu,b,\beta)$ by the subtraction

$R(t|\mu,b,\beta) = [R(t|\mu,b) + R(\beta|\mu,b,t)] - R(\beta|\mu,b)$.

The matrix notation, Table 2.5.1, shows two methods

of deriving the sums of squares for the adjusted

covariate terms. One uses the projector $R_Z(R_Z'R_Z)^-R_Z'$;

the other uses the $R(\cdot|\cdot)$ notation. Similarly, two

methods are shown for the unadjusted covariate term.

Examination of the model Σ derivation of R_Z, (p. 28),

(and R_3, (p. 31)), appears to show that the effects of

the discrete variables are removed before computing the

quadratic form. The second method, using $R(\cdot,\cdot,\cdot\cdot)$,

computes two quadratic forms and then uses the second

quadratic form to subtract the effect of the discrete

variables from the first quadratic form.

2.5.2 Summation Notation

The sums of squares and cross products in

summation notation are   (2.5.1)

total
$$SSTot_{..} = \sum_{i=1}^{b} \sum_{k=1}^{t} [\cdot ik][\cdot ik] \quad ,$$

mean
$$SSM_{..} = \frac{1}{bt} \left[\sum_{i=1}^{b} \sum_{k=1}^{t} [\cdot ik] \right] \left[\sum_{i=1}^{b} \sum_{k=1}^{t} [\cdot ik] \right] \quad ,$$

blocks
$$SSB_{..} = \frac{1}{t} \sum_{i=1}^{b} \left[\sum_{k=1}^{t} [\cdot ik] \right] \left[\sum_{k=1}^{t} [\cdot ik] \right] \quad ,$$

treatments
$$SST_{..} = \frac{1}{b} \sum_{k=1}^{t} \left[\sum_{i=1}^{b} [\cdot ik] \right] \left[\sum_{i=1}^{b} [\cdot ik] \right] \quad and$$

residual
$$SSR_{..} = (SSTot_{..} - SSM_{..}) - (SSB_{..} - SSM_{..})$$
$$- (SST_{..} - SSM_{..})$$
$$= SSTot_{..} - SSB_{..} - SST_{..} + SSM_{..} ,$$

where each dot ($\cdot$) is replaced by the appropriate

subscript, y or z, as shown in Table 2.5.2. Here, the

dot ($\cdot$) indicates replacement by y or z, not averaging

over a subscript.

Note that SSR_{yy} of the summation notation does not equal the SSR of the matrix notation. The later is

$$SSR = total - R(\mu) - R(b|\mu)$$

$$- \text{one of } \{ \ [\ R(t|\mu,b) \quad + R(\beta|\mu,b,t) \ ] \text{ or}$$

$$[\ R(t|\mu,b,\beta) + R(\beta|\mu,b) \ ] \ \}$$

$$= \Upsilon'\Upsilon - R(\mu,b,\tau,\beta)$$

$$= \Upsilon'\Upsilon - \{\Upsilon'[\mathbb{R}_Z(\mathbb{R}_Z'\mathbb{R}_Z)^-\mathbb{R}_Z]\Upsilon + \Upsilon'[\mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}']\Upsilon\}$$

$$= \Upsilon'\{ \ \mathbb{I} - \ \mathbb{R}_Z(\mathbb{R}_Z'\mathbb{R}_Z)^-\mathbb{R}_Z - \mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}' \ \}\Upsilon. \qquad (2.5.2)$$

The two terms within braces, {}, are equal. A similar subtraction provides the degrees of freedom for the residual sum of squares.

The term $R(\beta|\mu,b)$ is more easily understood as

$$R(\beta|\mu,b) = \frac{[(SSTot_{yz}-SSM_{yz})-(SSB_{yz}-SSM_{yz})]^2}{[(SSTot_{zz}-SSM_{zz})-(SSB_{zz}-SSM_{zz})]} . \qquad (2.5.3)$$

The numerator is a SSR_{yz} term and the denominator is a SSR_{zz} term for a model equation without the treatment effects. The form is the same as the sum of squares term for the covariate after the mean, blocks and treatments, $R(\beta|\mu,b,\tau)$. Both summation forms correspond to the respective matrix forms using $\mathbb{R}_Z$ and $\mathbb{R}_{\mathfrak{Z}}$.

In Table 2.5.1, the total uncorrected sum of squares is written in a form which emphasizes that it too is a quadratic form.

In the section following the tables, we examine, in some detail, the formulas used in hypothesis testing.

2.5.3 Tables of Sums of Squares

TABLE 2.5.1

COMBINED ANALYSIS OF VARIANCE TABLE -- MATRIX NOTATION

Source of Variation	d.f.	Sum of Squares
Total corrected for the mean (μ)	n-1	$\mathbf{Y}'\{\mathbb{I}\,(\mathbb{I}'\,\mathbb{I}\,)^{-1}\mathbb{I}'\}\mathbf{Y}$ $-\mathbf{Y}'\{\mathbf{X}_\mu(\mathbf{X}_\mu{}'\mathbf{X}_\mu)^{-1}\mathbf{X}_\mu{}'\}\mathbf{Y}$ $=\ \text{total}\ -\ R(\mu)$
Blocks after the mean (b$\mid\mu$)	b-1	$R(\mu,b) - R(\mu)$ $= R(b\mid\mu)$
Treatments and interactions after the mean and blocks (t$\mid\mu$,b)	t-1	$R(\mu,b,t) - R(\mu,b)$ $= R(t\mid\mu,b)$
Covariates after the mean, blocks, treatments and interactions ($\beta\mid\mu$,b,t)	q	$\mathbf{Y}'\{\mathbb{R}_Z(\mathbb{R}_Z{}'\mathbb{R}_Z)^{-1}\mathbb{R}_Z{}'\}\mathbf{Y}$ or $R(\mu,b,t,\beta)$ $-\ R(\mu,b,t)$ $= R(\beta\mid\mu,b,t)$
Covariates after the mean and blocks ($\beta\mid\mu$,b)	q	$\mathbf{Y}'\{\mathbb{R}_{\mathcal{Z}}(\mathbb{R}_{\mathcal{Z}}{}'\mathbb{R}_{\mathcal{Z}})^{-1}\mathbb{R}_{\mathcal{Z}}{}'\}\mathbf{Y}$ or $R(\mu,b,\beta)-R(\mu,b)$ $= R(\beta\mid\mu,b)$
Treatments and interactions after the mean, blocks and covariates (t$\mid\mu$,b,β)	t-1	$R(\mu,b,t,\beta)$ $-\ R(\mu,b,\beta)$ $= R(t\mid\mu,b,\beta)$
Residual	n-p-q	SSR (by subtraction)

TABLE 2.5.2

COMBINED ANALYSIS OF VARIANCE TABLE - SUMMATION NOTATION

Source of Variation	d.f.	Sum of Squares
Total corrected for the mean (μ)	n-1	$SSTot_{yy} - SSM_{yy}$ $= total - R(\mu)$
Blocks after the the mean $(b\mid\mu)$	b-1	$SSB_{yy} - SSM_{yy}$ $= R(b\mid\mu)$
Treatments after the mean and blocks $(\tau\mid\mu,b)$	t-1	$SST_{yy} - SSM_{yy}$ $= R(\tau\mid\mu)$
Covariate after the mean, blocks and treatments $(\beta\mid\mu,b,\tau)$	1	$\left[SSR_{yz}\right]^2 \left[SSR_{zz}\right]^{-1}$ $= R(\beta\mid\mu,b,\tau)$
Covariate after the mean and blocks $(\beta\mid\mu,b)$	1	$[SSTot_{yz}-SSB_{yz}]^2/$ $[SSTot_{zz}-SSB_{zz}]$ $= R(\beta\mid\mu,b)$
Treatments after the mean, blocks and covariate $(\tau\mid\mu,b,\beta)$	t-1	$SST_{yy} - SSM_{yy}$ $+R(\beta\mid\mu,b,\tau)$ $-R(\beta\mid\mu,b)$ $=R(\tau\mid\mu,b,\beta)$
Residual	n-b-t	$SSR_{yy} - R(\beta\mid\mu,b,\tau)$ $= SSR$ of the matrix notation

Section 2.6

The F-Ratio and Tests of Hypotheses

Section 2.6 develops the F-ratio and the noncentrality parameter for the hypotheses of interest in the analysis of covariance. Projectors for three partitionings of $[X \mid Z]$ are derived. They allow disentangling the matrix under the control of the experimenter, the design matrix, from the other matrices in some F-ratios and noncentrality parameters.

Important results are Theorems 2.6.15, (p. 55), and 2.6.19 - 2.6.22, (p. 59-69). Section 3.7 will further develop Theorems 2.6.19 - 2.6.22.

2.6.1 The Test of the Full Model

For the model with model equation (2.3.1), (p. 16), the test statistic for testing the null hypothesis

$$H_o: \quad \beta^+ = \mathbb{0} \qquad\qquad (2.6.1)$$

is the F-ratio

$$F \; = \; \frac{ Y'[\; \Sigma^+(\Sigma^{+\prime}\Sigma^+)^-\Sigma^{+\prime}\;]Y \; / \; [\; \sigma^2 \; n \;] }{ Y'[\; \mathbb{I} \; - \; \Sigma^+(\Sigma^{+\prime}\Sigma^+)^-\Sigma^{+\prime}\;]Y \; / \; [\; \sigma^2 \; d \;] } \; , \qquad (2.6.2)$$

with the non-centrality parameter

$$\lambda \; =$$

$$(1/2) \; (\Sigma^+\beta^+)'[\Sigma^+(\Sigma^{+\prime}\Sigma^+)^-\Sigma^{+\prime}](\Sigma^+\beta^+) \; / \; [\; \sigma^2 \;], \qquad (2.6.3)$$

$$= (1/2) \; \beta^{+\prime}[\; \Sigma^{+\prime}\Sigma^+]\beta^+ \; / \; [\; \sigma^2 \;], \qquad\qquad (2.6.4)$$

and $n = \text{rank}(\Sigma^+)$, $d = \text{rank}(\mathbb{I} - \Sigma^+(\Sigma^{+\prime}\Sigma^+)^-\Sigma^{+\prime})$.

Recall that $\Sigma^+ = [\Sigma_\mu \mid \Sigma_b \mid \Sigma_T \mid Z]$, as in (2.3.4), (p. 19).

Equation (2.6.4) follows from (2.6.3) by the definition of the generalized inverse (R1.1 of Section 2.8.7, p. 83). See Searle [1971, p. 54-64] for a development of the distribution of F from the assumptions of the model. When $\beta^+=0$, the parameters of the F-distribution (n, d, $\lambda=0$), are known, and thus the F-ratio is a statistic. When $\beta^+\neq0$, known or not, λ is not known, as by Assumption 3[*] and 3, (p. 16 and 24), σ^2 is unknown. In this case, the F-ratio is not a statistic, since one of its parameters, the σ^2 of its noncentrality parameter, is unknown. However, one may assume value(s) for the ratio β^+/σ, or assume a range of values for σ^2 for given value(s) of β^+. Morrison [1983, p. 127-131] does the former in determining the sample size needed to achieve a given power. Alternatively, one may compare ratios of λ as computed from different model matrices, X^+, and/or different β^+ vectors. Comparing ratios of two β's, such as $(\beta_1^+/\sigma^2)/(\beta_2^+/\sigma^2)$, or of two λ's, cancels the unknown σ^2's. Further investigation of handling an unknown σ^2 is an extension.

When the null hypothesis (2.6.1) is true, $\beta^+=0$, and by (2.6.4), $\lambda=0$. Under Assumptions 4 and 5, X^+ has full column rank, thus $X^{+'}X^+$ is positive definite [Searle, 1971, p. 36, Lemma 5]. By the definition of positive definitiveness, $\lambda>0$ except for $\beta^+=0$. Thus, the converse

holds: when $\lambda=0$, β^+ must equal $\mathbb{O}$. This gives the equivalence of

$$H_0: \beta^+= \mathbb{O} \quad \text{and}$$

$$H_0: \lambda = 0. \qquad\qquad (2.6.5)$$

The first is "a set of linear statements" about the parameters β^+. The second is a statement about a quadratic form of the parameters. See Searle [1987, p. 233-236] for this and additional forms of the hypotheses. The general linear hypothesis, $H_0: \mathbb{K}\beta^+=m$ is an extension.

The power of the test is one minus the probability of failing to detect a false null hypothesis, that is, (1-the probability of a Type II error)=$1-\beta_{II}$. For any given alternative $\beta^+ \neq \mathbb{O}$ (equivalently $\lambda \neq 0$), the power of the test increases with increasing λ. By (2.6.4) and the shape of the density function of F, increasing the power requires

 a. manipulating Σ^+ to maximize those λ corres-

  ponding to alternative, nonzero β^+'s, or

 b. increasing d, or

 c. decreasing n,   (2.6.6a-c)

while keeping the other two parameters constant [Seber, 1980, p.35]. Sections 3.7.5, (p. 216) and 4.8, (p. 335), examine option (a).

2.6.2 Testing Model Equation Parameters Adjusted

for Other Model Equation Parameters

The tests of hypotheses of interest in the analysis of covariance are that some elements of β^+, those of the hypothesis, are zero after adjusting for the remaining elements. Partition the model equation parameters into those of the hypotheses (h) and all of those remaining (r). The parameter vector and model matrix are partitioned conformably, as in

$$\beta^+ = \begin{bmatrix} \beta_h \\ -- \\ \beta_r \end{bmatrix} \quad \text{and}$$

$$\Sigma^+ = [\ \Sigma_h\ |\ \Sigma_r\ ]. \tag{2.6.7}$$

The columns of Σ^+ and rows of β^+ are reordered to permit this partitioning. Denote the parameters of the model equation providing the full model by subscripts on H_o, and denote the parameters of the model being tested by a subscript on β, as shown below,

$$H_{o;h,r}\ \beta_h = \mathbb{O}.$$

The partitioning yielding (2.3.4), (p. 19), is a special case of (2.6.7). The R-notation for the sums of squares associated with this hypothesis is R(h|r)=R(h,r)-R(r).

The denominator of (2.6.2) remains unchanged, while the R-notation gives as the numerator $R(h|r)=R(h,r)-R(r)=$

$$(1/2)\ \Upsilon'[\ \Sigma^+(\Sigma^{+'}\Sigma^+)^-\Sigma^{+'} - \Sigma_r(\Sigma_r'\Sigma_r)^-\Sigma_r'\ ]\ \Upsilon$$

$$/\ \{\ \sigma^2\ [\ rank(\Sigma^+) - rank(\Sigma_r)\ ]\}. \tag{2.6.8}$$

As always, the <u>non-centrality</u> <u>parameter</u> has the same center matrix, as in the brackets [] in (2.6.8), as the numerator of the associated F ratio, and replaces Υ with $\Sigma^+\beta^+$, where the value of β^+ is that of the alternative hypothesis. The non-centrality parameter is

$$\lambda = [1/(2\ \sigma^2)]\ x$$

$$(\Sigma_h\beta_h)'[\ \Sigma^+(\Sigma^{+'}\Sigma^+)^-\Sigma^{+'}- \Sigma_r(\Sigma_r'\Sigma_r)^-\Sigma_r']\ (\Sigma_h\beta_h) \tag{2.6.9}$$

$$= [1/(2\ \sigma^2)]\ x$$

$$\beta_h'\ \{\Sigma_h'\Sigma_h - \Sigma_h'[\Sigma_r(\Sigma_r'\Sigma_r)^-\Sigma_r']\Sigma_h\}\ \beta_h \tag{2.6.10}$$

$$= [1/(2\ \sigma^2)]\ x$$

$$\beta_h'\Sigma_h'\ \{I - [\Sigma_r(\Sigma_r'\Sigma_r)^-\Sigma_r']\}\ \Sigma_h\beta_h. \tag{2.6.11}$$

Equation (2.6.10) follows from (2.6.9) by Theorem 2.6.3, below, (p. 47). The first term within the braces {·} of (2.6.10) is positive definite if Σ_h has full column rank and is positive semidefinite if Σ_h has less than full column rank. The second term is positive definite if (a) Σ_h has full column rank, (b) this rank is less than or equal to the rank of Σ_r, and (c) the columns of the two matrices are linearly independent. Otherwise, the second term is positive semidefinite, or, as in (2.6.12), zero.

[Searle, 1971, p.36, Lemma 5 and p. 37, Lemma 7.] As the second term follows a negative sign, its contribution to the non-centrality parameter is to lower (the second term is greater than zero) or to leave unchanged (the second term equals zero) the value of λ and thus to lower or to leave unchanged the power of the hypothesis test, as discussed in (2.6.6a), (p. 41).

As Weisberg [1980, p. 89] points out, when X_h and X_r are orthogonal, (2.6.10) reduces to

$$\lambda = (1/2)\ \beta_h' \ [\ X_h'X_h]\ \beta_h\ /\ \sigma^2. \qquad (2.6.12)$$

In general, λ depends upon β_h, σ^2, $X_h'X_h$, $X_r(X_r'X_r)^-X_r'$, and the product $X_h'X_r$. Designs with an equal number of experimental units receiving each treatment and without a covariate may select design matrices X_h and X_r such that $X_h'X_r = 0$. In general, a covariate forces $X_h'X_r \neq 0$. When this product is altered by the randomization used to conduct the experiment, then the power of the test is conditioned on the randomization used. Section 3.7.5 (p. 216), further discusses this. On the other hand, the observed significance level is derived from the null hypothesis, which usually assumes $\beta_h = 0$. In this case, altering the randomization has no effect on the non-centrality parameter, as by assumption the non-centrality parameter is zero.

See Searle [1971, p. 343-344] for the non-centrality parameter (2.6.10) in $\mathbb{R}_Z$ notation. Morrison [1983, p.122-124] develops a special case of (2.6.9). In his Equation 8 [p. 123], jα twice erroneously appears as αj. Pope and Webster [1972, p. 330-331] state without proof the form of λ for stepwise regression with a missing parameter.

2.6.3 Products of Partitioned Matrices

This section expresses the estimate of β, (2.4.21), (p. 30), the F-ratio (2.6.2), (p. 39), and the noncentrality parameter (2.6.10), (p. 43), in an alternative form. The reexpressions are based upon the main results of this section, Theorems 2.6.16, (p. 56), and 2.6.17, (p. 57). Theorem 2.6.3 is referred to after Equation (2.6.11), (p. 43), above.

Theorem 2.6.1.

For $\Sigma = [\ \mathbb{A}\ |\ \mathbb{B}\ ]$, with Σ nx(p+q), $\mathbb{A}$ nxp and $\mathbb{B}$ nxq,

$$(\Sigma'\Sigma)^- =$$

$$\begin{bmatrix} (\mathbb{A}'\mathbb{A})^- + (\mathbb{A}'\mathbb{A})^-(\mathbb{A}'\mathbb{B})\mathbb{Q}^-(\mathbb{B}'\mathbb{A})(\mathbb{A}'\mathbb{A})^- & -(\mathbb{A}'\mathbb{A})^-(\mathbb{A}'\mathbb{B})\mathbb{Q}^- \\ -\mathbb{Q}^-(\mathbb{B}'\mathbb{A})(\mathbb{A}'\mathbb{A}) & \mathbb{Q}^- \end{bmatrix},$$

with $\mathbb{W} = [\mathbb{A}(\mathbb{A}'\mathbb{A})^-\mathbb{A}']$ and

$$\mathbb{Q} = (\mathbb{B}'\mathbb{B}) - (\mathbb{B}'\mathbb{A})(\mathbb{A}'\mathbb{A})^-(\mathbb{A}'\mathbb{B})$$

$$= \mathbb{B}'\mathbb{B} - \mathbb{B}'\mathbb{W}\mathbb{B} = \mathbb{B}'[\ \mathbb{I} - \mathbb{W}\ ]\mathbb{B}. \qquad (2.6.13)$$

Note that $\mathbb{Q}$ is the multivariate analogue for the residual sum of squares for the multivariate model equation

$$\mathbb{B}_{nxq} = \mathbb{A}_{nxp}\ \beta_{pxq} + E_{nxq}\ .$$

Proof:
$$X'X = \begin{bmatrix} A' \\ B' \end{bmatrix} \begin{bmatrix} A & | & B \end{bmatrix} = \begin{bmatrix} A'A & A'B \\ B'A & B'B \end{bmatrix}.$$

The generalized inverse of $(X'X)$ gives the desired result [Searle, 1971, p. 27, Equation (49)].

Theorem 2.6.2.

For $X = [\, A \mid B \,]$,

$$X(X'X)^- X' = W + W[BQ^-B']W'$$

$$- [BQ^-B']W - W[BQ^-B'] + [BQ^-B'] \,,$$

with $W = [A(A'A)^-A']$ and

$$Q = B'B - B'WB = B'[\, I - W \,]B. \qquad (2.6.14)$$

Proof: In partitioned form the projector $X(X'X)^- X'$

$$= \begin{bmatrix} A & | & B \end{bmatrix} \begin{bmatrix} A'A & A'B \\ B'A & B'B \end{bmatrix}^- \begin{bmatrix} A' \\ B' \end{bmatrix}.$$

Inserting (2.6.13) and multiplying gives

$$X(X'X)^- X'$$

$$= [A(A'A)^-A'] + [A(A'A)^-A'][BQ^-B'][A(A'A)^-A']$$

$$- [BQ^-B'][A(A'A)^-A'] - [A(A'A)^-A'][BQ^-B']$$

$$+ [BQ^-B']. \qquad (2.6.15)$$

Defining W and Q as in (2.6.14) gives the result.

Theorem 2.6.3.

For $\mathbb{X} = [\ \mathbb{A}\ |\ \mathbb{B}\ ]$,

$$\mathbb{A}'[\mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}']\mathbb{A} = \mathbb{A}'\mathbb{A}.$$

Proof: Substituting (2.6.15) into the bracketed term of

$\mathbb{A}'[\mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}']\mathbb{A}$ and regrouping terms gives

$\mathbb{A}'[\mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}']\mathbb{A}$

$$= +\ (\mathbb{A}'\mathbb{A})(\mathbb{A}'\mathbb{A})^-(\mathbb{A}'\mathbb{A})$$

$$+\ [(\mathbb{A}'\mathbb{A})(\mathbb{A}'\mathbb{A})^-\mathbb{A}'][\mathbb{B}\mathbb{Q}^-\mathbb{B}'][\mathbb{A}(\mathbb{A}'\mathbb{A})^-(\mathbb{A}'\mathbb{A})]$$

$$-\ \mathbb{A}'[\mathbb{B}\mathbb{Q}^-\mathbb{B}'][\mathbb{A}(\mathbb{A}'\mathbb{A})^-(\mathbb{A}'\mathbb{A})]$$

$$-\ [(\mathbb{A}'\mathbb{A})(\mathbb{A}'\mathbb{A})^-\mathbb{A}'][\mathbb{B}\mathbb{Q}^-\mathbb{B}']\mathbb{A}$$

$$+\ \mathbb{A}'[\mathbb{B}\mathbb{Q}^-\mathbb{B}']\mathbb{A}.$$

Repeated application of Equation 29 and the Corollary,

both on p. 20 of Searle [1971], gives

$\mathbb{A}'[\mathbb{X}(\mathbb{X}'\mathbb{X})^-\mathbb{X}']\mathbb{A}$

$$= +\ \mathbb{A}'\mathbb{A}$$

$$+\ \mathbb{A}'[\mathbb{B}\mathbb{Q}^-\mathbb{B}']\mathbb{A}$$

$$-\ \mathbb{A}'[\mathbb{B}\mathbb{Q}^-\mathbb{B}']\mathbb{A}$$

$$-\ \mathbb{A}'[\mathbb{B}\mathbb{Q}^-\mathbb{B}']\mathbb{A}$$

$$+\ \mathbb{A}'[\mathbb{B}\mathbb{Q}^-\mathbb{B}']\mathbb{A}$$

$$= \mathbb{A}'\mathbb{A}. \qquad\qquad (2.6.16)$$

We now develop the projection matrix for three

partitions of $\mathbb{X}$. For each, the initial step utilizes

Assumption 5, that the matrices $\mathbb{X}_D$ and $\mathbb{X}_T$ are the effects

version of the design matrix, (p. 74). Thus, each

partition has full rank and a unique inverse. By Theorem 2.8.3, (p. 76), the final step, the projector, is the same for any matrix which satisfies the conditions of the theorem. Thus, each projector holds for a class of matrices, not only those satisfying Assumption 5.

The matrices are defined for one experimental unit per treatment per block, thus p=t. The adjustment to multiple experimental units, say r, per treatment per block is to repeat each row r times in the initial partitioned form and add r as a multiplier or divisor in subsequent steps.

The first partition is $X = \left[\ X_\mu\ |\ X_b\ \right]$.

Theorem 2.6.4.

For $X = \left[\ X_\mu\ |\ X_b\ \right]$, with t experimental units in each of b levels of the blocking factor and where X_b is the effects version of the blocks partition of the design matrix, the matrix

$$X = \begin{bmatrix} \mathbb{1}_{t \times 1} & \mathbb{1}_{t \times 1} & \mathbb{0}_{t \times 1} & \cdots & \mathbb{0}_{t \times 1} \\ \mathbb{1}_{t \times 1} & \mathbb{0}_{t \times 1} & \mathbb{1}_{t \times 1} & & \mathbb{0}_{t \times 1} \\ \vdots & \vdots & & \ddots & \vdots \\ & & & & \mathbb{0}_{t \times 1} \\ \mathbb{1}_{t \times 1} & \mathbb{0}_{t \times 1} & \cdots & \mathbb{0}_{t \times 1} & \mathbb{1}_{t \times 1} \\ \mathbb{1}_{t \times 1} & & -\mathbb{1}_{t \times (b-1)} & & \end{bmatrix}_{(bt) \times b} .$$

(2.6.17)

Proof: Apply Assumption 5 and the partitioning of X.

Theorem 2.6.5.

For $\mathbf{X} = \left[\begin{array}{c|c} \mathbf{X}_\mu & \mathbf{X}_b \end{array} \right]$, with t experimental units in each of b levels of the blocking factor and where $\mathbf{X}_b$ is the effects version of the blocks partition of the design matrix, the matrix

$$\mathbf{X'X} = t \left[\begin{array}{c|c} b & \mathbb{O}_{1\times(b-1)} \\ \hline \mathbb{O}_{(b-1)\times 1} & \mathbb{I}_{(b-1)\times(b-1)} + \mathbb{J}_{(b-1)\times(b-1)} \end{array} \right]_{b\times b} .$$

(2.6.18)

Proof: Matrix multiplication of (2.6.17).

Theorem 2.6.6.

For $\mathbf{X} = \left[\begin{array}{c|c} \mathbf{X}_\mu & \mathbf{X}_b \end{array} \right]$, with t experimental units in each of b levels of the blocking factor and where $\mathbf{X}_b$ is the effects version of the blocks partition of the design matrix, the matrix

$$(\mathbf{X'X})^{-1} =$$

$$\frac{1}{bt} \left[\begin{array}{c|c} 1 & \mathbb{O}_{1\times(b-1)} \\ \hline \mathbb{O}_{(b-1)\times 1} & b\mathbb{I}_{(b-1)\times(b-1)} - \mathbb{J}_{(b-1)\times(b-1)} \end{array} \right]_{b\times b}.$$

(2.6.19)

Proof: Matrix multiplication of (2.6.18) by (2.6.19) yields the identity matrix. The inverse is unique since, by assumption, $\mathbf{X'X}$ has full rank.

Theorem 2.6.7.

For $X = \left[\ X_\mu\ |\ X_b\ \right]$, with t experimental units in each of b levels of the blocking factor and where X_b is the effects version of the blocks partition of the design matrix or one satisfying Theorem 2.8.3, (p. 87), the projector

$$X(X'X)^-X' = \frac{1}{t}\begin{bmatrix} J_{txt} & O_{txt} & O_{txt} & \cdots & O_{txt} \\ & J_{txt} & O_{txt} & \cdots & O_{txt} \\ & & \ddots & & \vdots \\ & \text{symmetric} & & \ddots & O_{txt} \\ & & & & J_{txt} \end{bmatrix}_{(bt)x(bt)}. \quad (2.6.20)$$

$$= (1/t)\{\text{diag}[J_{txt}]\}_{bxb}. \quad \text{(Defined as } W \text{ below.)}$$

Proof: Matrix multiplication of (2.6.19) by (2.6.17), the latter twice. The projector (2.6.20) may be obtained directly from the summation notation by a Kronecker product process developed by Craig Jefferson and the author. Appendix C, (p. 385), presents this process.

The second partition is $\mathbf{X} = \begin{bmatrix} \mathbf{X}_T \end{bmatrix}$.

Theorem 2.6.8.

For $\mathbf{X} = \begin{bmatrix} \mathbf{X}_T \end{bmatrix}$, with t experimental units in each of b levels of the blocking factor and where $\mathbf{X}_T$ is the effects version of the treatments partition of the design matrix, the matrix

$$\mathbf{X} = \begin{bmatrix} \mathbb{I}_{(t-1)x(t-1)} \\ -\mathbb{1}_{1x(t-1)} \\ \hline \mathbb{I}_{(t-1)x(t-1)} \\ -\mathbb{1}_{1x(t-1)} \\ \hline \vdots \\ \hline \mathbb{I}_{(t-1)x(t-1)} \\ -\mathbb{1}_{1x(t-1)} \end{bmatrix}_{(bt)x(t-1)} . \qquad (2.6.21)$$

Proof: Apply Assumption 5 and the partitioning of $\mathbf{X}$.

Theorem 2.6.9.

For $\mathbf{X} = \begin{bmatrix} \mathbf{X}_T \end{bmatrix}$, with t experimental units in each of b levels of the blocking factor and where $\mathbf{X}_T$ is the effects version of the treatments partition of the design matrix, the matrix

$$\mathbf{X'X} = b \begin{bmatrix} \mathbb{I}_{(t-1)x(t-1)} + \mathbb{J}_{(t-1)x(t-1)} \end{bmatrix}_{(t-1)x(t-1)}.$$
$$(2.6.22)$$

Proof: Matrix multiplication of (2.6.21).

Theorem 2.6.10.

For $\Sigma = \begin{bmatrix} \Sigma_T \end{bmatrix}$, with t experimental units in each of b levels of the blocking factor and where Σ_T is the effects version of the treatments partition of the design matrix, the matrix

$$(\Sigma'\Sigma)^{-1} = \frac{1}{bt}\begin{bmatrix} t\mathbb{I}_{(t-1)\times(t-1)} & - & \mathbb{J}_{(t-1)\times(t-1)} \end{bmatrix}_{(t-1)\times(t-1)} . \quad (2.6.23)$$

Proof: Matrix multiplication of (2.6.23) by (2.6.22) yields the identity matrix. The inverse is unique since, by assumption, $\Sigma'\Sigma$ has full rank.

Theorem 2.6.11.

For $\Sigma = \begin{bmatrix} \Sigma_T \end{bmatrix}$, with t experimental units in each of b levels of the blocking factor and where Σ_T is the effects version of the treatments partition of the design matrix, or one satisfying Theorem 2.8.3, (p. 87), the projector $\Sigma(\Sigma'\Sigma)^-\Sigma'$ has b x b identical blocks, each t x t. Thus,

$$\Sigma(\Sigma'\Sigma)^-\Sigma' = \frac{1}{bt}\begin{bmatrix} t\mathbb{I} - \mathbb{J} & t\mathbb{I} - \mathbb{J} & & t\mathbb{I} - \mathbb{J} \\ & t\mathbb{I} - \mathbb{J} & & t\mathbb{I} - \mathbb{J} \\ & & \ddots & t\mathbb{I} - \mathbb{J} \\ \text{symmetric} & & & t\mathbb{I} - \mathbb{J} \end{bmatrix}_{(bt)\times(bt)} . \quad (2.6.24)$$

Proof: Matrix multiplication of (2.6.23) by (2.6.21), the latter twice.

Note that the coding version of $\mathbf{X}_T$ does not satisfy

Theorem 2.8.3 (p. 87) and thus does not yield the

projector matrix of (2.6.24); see Section 2.8.4 below,

(p. 79). The coding version yields a projector similar

to (2.6.24), but the diagonal elements of the diagonal

blocks are not all equal; one is zero. The projector

(2.6.24) may be obtained directly from the summation

notation by the Kronecker product process of Appendix C.

The third partition is $\mathbf{X} = \begin{bmatrix} \mathbf{X}_\mu & | & \mathbf{X}_b & | & \mathbf{X}_T \end{bmatrix}$.

Theorem 2.6.12.

For $\mathbf{X} = \begin{bmatrix} \mathbf{X}_\mu & | & \mathbf{X}_b & | & \mathbf{X}_T \end{bmatrix}$, with t experimental units

in each of b levels of the blocking factor and where $\mathbf{X}_b$

and $\mathbf{X}_T$ are the effects version of the blocks and

treatments partitions of the design matrix, the matrix

$$
\mathbf{X} = \left[
\begin{array}{c|cccc|c}
\mathbf{1}_{t\times 1} & \mathbf{1}_{t\times 1} & \mathbf{0}_{t\times 1} & \cdots & \mathbf{0}_{t\times 1} & \mathbf{I}_{(t-1)\times(t-1)} \\
 & & & & & -\mathbf{1}_{1\times(t-1)} \\
\hline
\mathbf{1}_{t\times 1} & \mathbf{0}_{t\times 1} & \mathbf{1}_{t\times 1} & & \mathbf{0}_{t\times 1} & \text{as above} \\
\vdots & \vdots & & \ddots & \vdots & \vdots \\
 & & & & \mathbf{0}_{t\times 1} & \\
\hline
\mathbf{1}_{t\times 1} & \mathbf{0}_{t\times 1} & \cdots & \mathbf{0}_{t\times 1} & \mathbf{1}_{t\times 1} & \text{as above} \\
\mathbf{1}_{t\times 1} & & -\mathbf{I}_{t\times(b-1)} & & & \text{as above}
\end{array}
\right]
$$

$$[bt]\times[1+(b-1)+(t-1)]. \qquad (2.6.25)$$

Proof: Apply Assumption 5 to the indicated partition.

Theorem 2.6.13.

For $\mathbf{X} = \left[\ \mathbf{X}_\mu\ |\ \mathbf{X}_b\ |\ \mathbf{X}_\tau\ \right]$, with t experimental units

in each of b levels of the blocking factor and where $\mathbf{X}_b$

and $\mathbf{X}_\tau$ are the effects version of the blocks and

treatments partitions of the design matrix, the matrix

$$\mathbf{X'X} = \left[\begin{array}{c|c|c}
bt & \mathbb{O}_{1\times(b-1)} & \mathbb{O}_{1\times(t-1)} \\
\hline
\mathbb{O}_{(b-1)\times 1} & \begin{array}{c} t\mathbb{J}_{(b-1)\times(b-1)} \\ +\ t\mathbb{I}_{(b-1)\times(b-1)} \end{array} & \mathbb{O}_{(b-1)\times(t-1)} \\
\hline
\mathbb{O}_{(t-1)\times 1} & \mathbb{O}_{(t-1)\times(b-1)} & \begin{array}{c} b\mathbb{J}_{(t-1)\times(t-1)} \\ +\ b\mathbb{I}_{(t-1)\times(t-1)} \end{array}
\end{array}\right]$$

$$(b+t+1)\times(b+t+1). \qquad (2.6.26)$$

Proof: Apply matrix multiplication to (2.6.25).

Theorem 2.6.14.

For $\mathbf{X} = \left[\ \mathbf{X}_\mu\ |\ \mathbf{X}_b\ |\ \mathbf{X}_\tau\ \right]$, with t experimental units

in each of b levels of the blocking factor and where $\mathbf{X}_b$

and $\mathbf{X}_\tau$ are the effects version of the blocks and

treatments partitions of the design matrix, the matrix

$(\mathbf{X'X})^{-1} =$

$$\frac{1}{bt}\left[\begin{array}{c|c|c}
1 & \mathbb{O}_{1\times(b-1)} & \mathbb{O}_{1\times(t-1)} \\
\hline
\mathbb{O}_{(b-1)\times 1} & \begin{array}{c} b\mathbb{I}_{(b-1)\times(b-1)} \\ -\ \mathbb{J}_{(b-1)\times(b-1)} \end{array} & \mathbb{O}_{(b-1)\times(t-1)} \\
\hline
\mathbb{O}_{(t-1)\times 1} & \mathbb{O}_{(t-1)\times(b-1)} & \begin{array}{c} t\mathbb{I}_{(t-1)\times(t-1)} \\ -\ \mathbb{J}_{(t-1)\times(t-1)} \end{array}
\end{array}\right]$$

$$(b+t+1)\times(b+t+1) \qquad (2.6.27)$$

Proof: Matrix multiplication of (2.6.27) by (2.6.26)

yields the identity matrix. The inverse is unique since,

by Assumption 5, (p. 20), $\mathbf{X'X}$ has full rank.

Theorem 2.6.15.

For $\mathbf{X} = \begin{bmatrix} \mathbf{X}_\mu & | & \mathbf{X}_b & | & \mathbf{X}_\tau \end{bmatrix}$, with t experimental units

in each of b levels of the blocking factor and where $\mathbf{X}_b$

and $\mathbf{X}_\tau$ are the effects version of the blocks and

treatments partitions of the design matrix or ones

satisfying Theorem 2.8.3, (p. 87), the projector

$\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}' =$

$$\frac{1}{bt}\begin{bmatrix}
t\mathbf{I}+(b-1)\mathbf{J} & t\mathbf{I}-\mathbf{J} & & t\mathbf{I}-\mathbf{J} \\
& t\mathbf{I}+(b-1)\mathbf{J} & & t\mathbf{I}-\mathbf{J} \\
\text{symmetric} & & \ddots & t\mathbf{I}-\mathbf{J} \\
& & & t\mathbf{I}+(b-1)\mathbf{J}
\end{bmatrix}$$
$$\text{(bxt) x (bxt)} \qquad (2.6.28)$$

Proof: Matrix multiplication of (2.6.27) by (2.6.25),

the latter twice. The projector (2.6.28) may be obtained

directly from the summation notation by the Kronecker

product process of Appendix C, (p. 385).

Note that there are but two types of submatrices in

the projector matrices of (2.6.20), (p. 50), (2.6.24),

(p. 52) and (2.6.28). The submatrices on the diagonal

are one type. The submatrices off of the diagonal are

another. In each of the diagonal submatrices, all diago-

al elements have the same value, as do all off-diagonal

elements. The two sets of diagonal elements may be equal

as in (2.6.20), or unequal as in (2.6.24) and (2.6.28).

In each of the off-diagonal submatrices, again all diagonal elements have the same value and all off-diagonal elements have the same value. The two sets of elements may be equal as in (2.6.20), or not equal as in (2.6.24) and (2.6.28).

Finally, the diagonal blocks may equal the off-diagonal blocks, as in (2.6.24) or they may be unequal as in (2.6.20) and (2.6.28). These structural characteristics have important effects upon the randomization model as discussed in Section 3.7.2, (p. 187).

Theorems 2.6.16 and 2.6.17 decompose projectors used in the usual analysis of covariance hypotheses. The decomposition separates the treatments part of the appropriate model matrix from the other parts. The first theorem is a special case of Theorem 2.6.2, (p. 47).

Theorem 2.6.16.

For versions of Σ_T with all columns summing to zero within each block, such as the effects version, and $\Sigma_r = [\ \Sigma_\mu |\ \Sigma_b |\ \Sigma_T]$, the projector

$$[\Sigma_r(\Sigma_r'\Sigma_r)^-\Sigma_r'] = W + \Sigma_T(\Sigma_T'\Sigma_T)^-\Sigma_T'\ ,\ \text{where}$$

$$W = (1/t)\{diag[J_{t \times t}]\}_{b \times b} \text{ as in (2.6.20)}. \quad (2.6.29)$$

Note that Σ of (2.4.7), (p. 26), is, here Σ_r.

Proof: Partition Σ_r as $\Sigma_r = [A\ |\ B]$, where $A = [\Sigma_\mu\ |\ \Sigma_b]$

and $B = \Sigma_T$. Theorem 2.6.2, (p. 47), gives $A(A'A)^-A'$ as W; Theorem 2.6.7, (p. 50), gives the above W.

Theorem 2.6.2 splits $X_T(X_T{'}X_T)^-X_T{'}$ into five terms. The middle three contain WB and/or $B'W'$. This product sums one block's worth of columns of X_T (=B). By Assumption 5, (p. 20), each such sum is zero. The fifth term has the form $X_T Q^-X_T{'}$, where $Q = X_T{'}[\ I\ -\ W\ ]X_T$, as in (2.6.13), (p. 54). Again $X_T{'}W = 0$, leaving $Q=X_T{'}IX_T$; thus, $X_T Q^-X_T = X_T(X_T{'}X_T)^-X_T{'}$. The first term from Theorem 2.6.2 remains unchanged, providing the desired result of separating X_T from the other partitions of this model matrix.

Theorem 2.6.17.

For a treatment matrix X_T orthogonal to both the blocks matrix X_b and the mean vector X_μ, as when X_T and X_b are the effects version, and for $X_r =[\ X_\mu|\ X_b|\ Z\ ]$,

$X_T{'}\ [X_r(X_r{'}X_r)^-X_r{'}]\ X_T$

$= X_T{'}\ \{Z\ [Z'(I-W)Z]^-\ Z'\}\ X_T$, where

$W = (1/t)\{diag[J_{txt}]\}_{bxb}$ as in (2.6.20).   (2.6.30)

Proof: Partition X_r as $X_r = [A\ |\ B]$, where $A = [X_\mu\ |\ X_b]$ and $B =Z$. Theorem 2.6.2, (p. 47), gives $A(A'A)^-A'$ for W; Theorem 2.6.7, (p. 50), gives $W = (1/t)\{diag[J_{txt}]\}_{bxb}$.

Theorem 2.6.2 splits $\mathbf{X}_T(\mathbf{X}_T'\mathbf{X}_T)^-\mathbf{X}_T'$ into five factors, all except the last beginning and/or ending in $\mathbb{W}$. Pre-multiplication by $\mathbf{X}_T'$ or post-multiplication by $\mathbf{X}_T$ gives

$$\mathbf{X}_T'\mathbb{W} = \mathbf{X}_T' \ \{ \ [\mathbf{X}_\mu | \mathbf{X}_b](\mathbb{A}'\mathbb{A})^-\mathbb{A}' \ \} \quad , \ or$$

$$\mathbb{W}\mathbf{X}_T = \ \{ \ \mathbb{A}(\mathbb{A}'\mathbb{A})^-[\mathbf{X}_\mu | \mathbf{X}_b]' \ \} \ \mathbf{X}_T.$$

Each of these four factors is zero by the assumed mutual orthogonality of $[\mathbf{X}_\mu | \mathbf{X}_b]$ and $\mathbf{X}_T$, as discussed below (2.3.5), (p. 127). The last factor of Theorem 2.6.2 is $\mathbb{B}$ $\mathbb{Q}^-\mathbb{B}' = \mathbb{B} [\mathbb{B}'(\mathbb{I}-\mathbb{W})\mathbb{B}]^-\mathbb{B}'$. Upon substituting $\mathbb{Z}$ for $\mathbb{B}$, this provides the desired result of separating $\mathbf{X}_T$ from the other partitions of this model matrix.

2.6.4 Alternative Forms of $\hat{\mathbf{\beta}}_{\mu,b,T,\beta}$, the F-Ratio and the Non-Centrality Parameter

The following theorems disentangle the matrix under the control of the experimenter, $\mathbf{X}_T$, from the other matrices, $\mathbf{X}_\mu$, $\mathbf{X}_b$ and $\mathbb{Z}$. In each theorem $\mathbf{X}_T$ remains as part of a projector matrix having the form of Equation (2.6.25) of Theorem 2.6.12.

Theorem 2.6.18.

For versions of Σ_T with all columns summing to zero

for each block, such as the effects version,

$$\mathbb{R}_Z = \mathbb{R}_X \mathbb{Z} = [\mathbb{I} - \mathbb{W} - \Sigma_T (\Sigma_T ' \Sigma_T)^- \Sigma_T '] \mathbb{Z}$$

where $\mathbb{W} = (1/t) \{ \text{diag}[\mathbb{J}_{t \times t}] \}_{b \times b}$.

as in (2.6.20).   (2.6.31)

Proof: Apply Theorem 2.6.16 to Equation 2.4.7, (p. 26).

Theorem 2.6.19.

For versions of Σ_T with all columns summing to zero

for each block, such as the effects version,

$$\hat{\beta}_{\mu, b, \tau, \beta} = \{ \mathbb{Z}' [\mathbb{I} - \mathbb{W} - \Sigma_T (\Sigma_T ' \Sigma_T)^- \Sigma_T '] \mathbb{Z} \}^{-1}$$

$$\times \{ \mathbb{Z}' [\mathbb{I} - \mathbb{W} - \Sigma_T (\Sigma_T ' \Sigma_T)^- \Sigma_T '] \} \mathbb{Y}$$

where $\mathbb{W} = (1/t) \{ \text{diag}[\mathbb{J}_{t \times t}] \}_{b \times b}$

as in (2.6.20).   (2.6.32)

Proof: Apply Theorem 2.6.18 to Equation 2.4.4, (p. 26).

This gives the desired result of disentangling the

treatments matrix from the other submatrices of the model

matrix.

<u>Theorem 2.6.20.</u>

For versions of $\mathbb{X}_T$ with all columns summing to zero

for each block, such as the effects version,

$$F(\beta|\mu,b,\tau) \; = \; \frac{\Upsilon'\{\mathbb{R}_Z(\mathbb{R}_Z'\mathbb{R}_Z)^{-1}\mathbb{R}_Z'\}\Upsilon \; /q}{SSR/(n-p-q)}$$

$$= \; \Big\{ \; \Upsilon'\{ \; [\mathbb{I} - \mathbb{W} - \mathbb{X}_T(\mathbb{X}_T'\mathbb{X}_T)^-\mathbb{X}_T']\mathbb{Z}$$

$$\times \; \{\mathbb{Z}'[\mathbb{I} - \mathbb{W} - \mathbb{X}_T(\mathbb{X}_T'\mathbb{X}_T)^-\mathbb{X}_T']\mathbb{Z}\}^{-1}$$

$$\times \; \mathbb{Z}'[\mathbb{I} - \mathbb{W} - \mathbb{X}_T(\mathbb{X}_T'\mathbb{X}_T)^-\mathbb{X}_T'] \; \}\Upsilon \; / \; q \; \Big\} \; \times$$

$$\Big\{\{\Upsilon'[\mathbb{I} - \mathbb{X}_T(\mathbb{X}_T'\mathbb{X}_T)^-\mathbb{X}_T' - \mathbb{R}_Z(\mathbb{R}_Z'\mathbb{R}_Z)^-\mathbb{R}_Z']\Upsilon\} \; / \; \{n-p-q\} \Big\}^{-1}$$

where

$$\mathbb{W} = (1/t)\{diag[\mathbb{J}_{t \times t}]\}_{b \times b} \text{ as in } (2.6.20). \qquad (2.6.33)$$

Here, p is the rank of $\mathbb{X}$, q is the rank of $\mathbb{Z}$.

Proof: Apply Theorem 2.6.18 to Table 2.5.1, (p. 37),

using the idempotency of $\mathbb{R}_Z$. Note that in the numerator

the treatments matrix is disentangled from the other

submatrices of the model matrix. The projector of $\mathbb{R}_Z$ in

the denominator may be similarly reexpressed with $\mathbb{X}_T$

separated from the other submatrices.

<u>Theorem 2.6.21.</u>

For versions of $\mathbf{X}_T$ with all columns summing to zero for each block, such as the effects version,

$\lambda_{\beta | \mu, b, \tau, \beta}$

$= 1/(2\sigma^2)$ x

$\{\beta'\mathbf{Z}'\ [\mathbf{I}-\mathbf{W}]\ \mathbf{Z}\beta\ -\ \beta'\mathbf{Z}'\ [\mathbf{X}_T(\mathbf{X}_T'\mathbf{X}_T)^-\mathbf{X}_T']\ \mathbf{Z}\beta\}$, (2.6.34)

$= 1/(2\sigma^2)$ x $\{\beta'\mathbf{Z}'\ [\mathbf{I}-\mathbf{W}-\mathbf{X}_T(\mathbf{X}_T'\mathbf{X}_T)^-\mathbf{X}_T']\ \mathbf{Z}\beta\}$,   (2.6.35)

$= 1/(2\sigma^2)$ x

$\beta'\{\mathbf{Z}'[\mathbf{I}-\mathbf{W}-\mathbf{X}_T(\mathbf{X}_T'\mathbf{X}_T)^-\mathbf{X}_T']\}\{[\mathbf{I}-\mathbf{W}-\mathbf{X}_T(\mathbf{X}_T'\mathbf{X}_T)^-\mathbf{X}_T']\mathbf{Z}\}\beta$,

where $\mathbf{W} = (1/t)\{diag[\mathbf{J}_{txt}]\}_{bxb}$

as in (2.6.20).   (2.6.36)

Proof: From Equation (2.6.10), (p. 43), with $\mathbf{X}_h=\mathbf{Z}$ and $\mathbf{X}_r=[\mathbf{X}_\mu | \mathbf{X}_b | \mathbf{X}_T]$,

$\lambda_{\beta | \mu, b, \tau, \beta} =$

$(1/2)\ \beta'\{\mathbf{Z}'\mathbf{Z} - \mathbf{Z}'[\mathbf{X}_r(\mathbf{X}_r'\mathbf{X}_r)^-\mathbf{X}_r']\mathbf{Z}\}\beta\ /\sigma^2$ (2.6.37)

By Theorem 2.6.16, (p. 56), the projector of this $\mathbf{X}_r$ is $\mathbf{W} + \mathbf{X}_T(\mathbf{X}_T'\mathbf{X}_T)^-\mathbf{X}_T'$. Insertion into (2.6.37) and rearranging gives (2.6.35) and (2.6.34). The idempotency of $\mathbb{R}_Z$ gives (2.6.36). In each of (2.6.34) to (2.6.36), the treatments matrix $\mathbf{X}_T$ is disentangled from the other submatrices of the model matrix.

Theorem 2.6.22.

For treatments matrices Σ_T orthogonal to both the blocks matrix Σ_b and the mean vector Σ_μ, as when Σ_T and Σ_b are the effects versions of these matrices,

$$\lambda_{T|\mu,b,\tau,\beta} = 1/(2\sigma^2) \text{ x} \qquad (2.6.38)$$

$$\{\tau'\Sigma_T'\Sigma_T\tau \quad - \quad \tau' \Sigma_T' [Z(Z'(I - W)Z)^{-}Z'] \Sigma_T \tau\},$$

where $W = (1/t)\{diag[J_{txt}]\}_{bxb}$, as in (2.6.20).

Proof: From Equation (2.6.10), (p. 43), with $\Sigma_h = \Sigma_T$ and $\Sigma_r = [\Sigma_\mu | \Sigma_b | Z]$,

$$\lambda_{T|\mu,b,\tau,\beta} =$$

$$(1/2) \tau'\{\Sigma_T'\Sigma_T - \Sigma_T'[\Sigma_r(\Sigma_r'\Sigma_r)^{-}\Sigma_r']\Sigma_T\}\tau /\sigma^2. \qquad (2.6.39)$$

Application of Theorem 2.6.17, (p. 57), to the second term of (2.6.39) and rearranging the τ gives the desired result. Note that Σ_T is disentangled from the other submatrices of the model matrix.

2.6.5 Testing Hypotheses of Interest in the Randomized Block Design with Covariates.

The following list of hypotheses and F-tests partitions Σ^+ as $[\Sigma | Z]$ and expresses the F-test of (2.6.2), (p. 39), in R notation. The numerator of the F-test is a form of (2.6.8), (p. 43). The denominator is that of (2.6.2). Under the assumption of a true null hypothesis, the F-statistic yields a p-value, herein

termed the <u>observed significance level</u>, OSL_F, in this case, as computed from the F distribution.

The hypothesis that the covariates, adjusted for the mean, blocks and treatment effects have no effect on the response is

$$H_{o; \mu, b, \tau} \; \beta = 0 \; . \qquad\qquad (2.6.40)$$

The F-statistic used to test this hypothesis is

$$F(\beta | \mu, b, \tau) = \frac{R(\beta | \mu, b, \tau)/q}{SSR/(n-p-q)} \; . \qquad\qquad (2.6.41)$$

The associated noncentrality parameter is (2.6.34), (p. 61),

$$\lambda_{\beta | \mu, b, \tau, \beta} = 1/(2\sigma^2) \; x$$

$$\{\beta' Z' \; [I - W] \; Z\beta \; - \; \beta' Z' \; [X_\tau (X_\tau' X_\tau)^{-} X_\tau'] \; Z\beta \},$$

where $W = (1/t)\{diag[J_{t \times t}]\}_{b \times b}$ as in (2.6.20), (p. 50).

The hypothesis that the treatments, adjusted for the mean and blocks, but without adjustment for the covariates, are all equal is

$$H_{o; \mu, b} \; \tau = 0 \; . \qquad\qquad (2.6.42)$$

Under Assumption 5, "all treatments equal", (p. 20), implies that each element of τ is zero. The F-statistic used to test this hypothesis is

$$F(\tau | \mu, b) = \frac{R(\tau | \mu, b)/(t-1)}{SSR/(n-p-q)} \; . \qquad\qquad (2.6.43)$$

Normally one replaces the above hypothesis with the following one which uses the treatment sum of squares adjusted also for the covariates. The hypothesis that the treatments, adjusted for the mean, blocks and covariates, are all equal is

$$H_{o;\mu,b,\beta} \ \tau = 0 \ . \qquad\qquad (2.6.44)$$

The F-statistic used to test this hypothesis is

$$F(\tau|\mu,b,\beta) = \frac{R(\tau|\mu,b,\beta)/(t-1)}{SSR/(n-p-q)} \ . \qquad (2.6.45)$$

The associated noncentrality parameter is (2.6.38),

$$\lambda_{\tau|\mu,b,\tau,\beta} = 1/(2\sigma^2) \ x$$

$$\{\tau'\Sigma_\tau'\Sigma_\tau\tau \ - \ \tau' \ \Sigma_\tau' \ [Z(\ Z'(I - W)Z)^-Z'] \ \Sigma_\tau \ \tau\},$$

where $W = (1/t)\{diag[J_{txt}]\}_{bxb}$

as in (2.6.20), (p. 50).

The hypothesis that the covariates, adjusted for the mean and blocks, but without adjustment for the treatments, have no effect upon the response is

$$H_{o;\mu,b} \ \beta = 0 \ . \qquad\qquad (2.6.46)$$

The F-statistic used to test this hypothesis is

$$F(\beta|\mu,b) = \frac{R(\beta|\mu,b)/q}{SSR/(n-p-q)} \ . \qquad (2.6.47)$$

Normally one replaces this last hypothesis with the one first above, $H_{o;\mu,b,\tau,\beta} \ \beta = 0$, which tests for the effect of the covariates after adjusting for treatment effects as well as those of the mean and blocks.

Next, a numerical example illustrates Sections 2.3-2.6.

Section 2.7

Numerical Example

Section 2.7 uses matrix and summation methods to calculate the combined analysis of variance tables, Tables 2.5.1 and 2.5.2, (p. 37-38), the estimates, $\hat{\beta}_{\mu,b,\beta}$, $\hat{\beta}_{\mu,b,\tau,\beta}$, $\hat{\tau}_{\mu,b,\tau}$, and $\hat{\tau}_{\mu,b,\tau,\beta}$, (p. 32), and associated F-statistics and noncentrality parameters of Equations (2.6.41) - (2.6.47). The artificial data represent one random assignment of treatments to experimental units.

For this example, the mean is $\mu = 10.0$, block effects are $b_1 = -b_2 = 1.5$, treatment effects are $\tau_k = 0$ for all k, and the covariate is $\beta = 2.0$. Table 2.7.1 shows the assigned treatment, k, the response, Y_{ik}, and the covariate, Z_{ik}, for each block-plot experimental unit. The plot ($j = \cdot$) and treatment ($k = \cdot$) identification numbers happen to be the same for this example.

TABLE 2.7.1
TREATMENTS AND RESPONSES FOR EXAMPLE ONE

	plot j=1	j=2	j=3	j=4	Block Totals
block i = 1	k=1 17.3956 Z_{11}= 5	k=2 29.8077 9	k=3 34.1529 17	k=4 52.9714 19	134.3276 50
block i = 2	k=1 38.4230 14	k=2 35.6973 11	k=3 7.4435 2	k=4 20.4594 1	102.0232 28
Totals Treat.: Covariate:	55.8186 19	65.5050 20	41.5964 19	73.4308 20	Grand Totals 236.3508 78

Examples of calculations, using the matrix notation

under Assumptions 1 - 8 with $\mathbb{X}=[\mathbb{X}_\mu|\mathbb{X}_b|\mathbb{X}_\tau]$, are

$$
\mathbb{X} = \begin{bmatrix} 1 & | & 1 & | & 1 & 0 & 0 \\ 1 & | & 1 & | & 0 & 1 & 0 \\ 1 & | & 1 & | & 0 & 0 & 1 \\ 1 & | & 1 & | & -1 & -1 & -1 \\ 1 & | & -1 & | & 1 & 0 & 0 \\ 1 & | & -1 & | & 0 & 1 & 0 \\ 1 & | & -1 & | & 0 & 0 & 1 \\ 1 & | & -1 & | & -1 & -1 & -1 \end{bmatrix}
\qquad
\mathbb{Y} = \begin{bmatrix} 17.3956 \\ 29.8077 \\ 34.1529 \\ 52.9714 \\ 38.4230 \\ 35.6973 \\ 7.4435 \\ 20.4594 \end{bmatrix}
\qquad
\mathbb{Z} = \begin{bmatrix} 5 \\ 9 \\ 17 \\ 19 \\ 14 \\ 11 \\ 2 \\ 1 \end{bmatrix}
$$

$$
(\mathbb{X}'\mathbb{X})^{-1} = \begin{bmatrix}
0.125 & 0 & 0 & 0 & 0 \\
0 & .125 & 0 & 0 & 0 \\
0 & 0 & 0.375 & -0.125 & -0.125 \\
0 & 0 & -0.125 & 0.375 & -0.125 \\
0 & 0 & -0.125 & -0.125 & 0.375
\end{bmatrix}
$$

$$
\mathbb{R}_\mathbb{Z} = \begin{bmatrix} -7.25 \\ -3.75 \\ 4.75 \\ 6.25 \\ 7.25 \\ 2.75 \\ -4.75 \\ -6.25 \end{bmatrix}
\qquad
\mathbb{R}_3 = \begin{bmatrix} -7.50 \\ -3.50 \\ 4.50 \\ 6.50 \\ 7.00 \\ 4.00 \\ -5.00 \\ -6.00 \end{bmatrix}
\qquad
\begin{aligned}
\hat{\beta}_{\mu,b,\beta} &= 2.0038 \\[1em]
\hat{\beta}_{\mu,b,\tau,\beta} &= 1.9672
\end{aligned}
$$

The vectors $\tau_{(.)}' = [\ \hat{\mu},\ \hat{b}_1,\ \hat{\tau}_1,\ \hat{\tau}_2,\ \hat{\tau}_3\ ]'$ are

$$
\tau_g^0 = \begin{bmatrix} 9.75 \\ 2.75 \\ -0.25 \\ 0.25 \\ -0.25 \end{bmatrix},
\qquad
\hat{\tau}_{\mu,b,\tau} = \begin{bmatrix} 29.54 \\ 4.04 \\ -1.63 \\ 3.21 \\ -8.75 \end{bmatrix},
\qquad
\hat{\tau}_{\mu,b,\tau,\beta} = \begin{bmatrix} 10.35 \\ -1.37 \\ -1.14 \\ 2.72 \\ -8.25 \end{bmatrix}.
$$
(Model $\mathbb{X}$)

By Assumption 5, $\hat{\tau}_{4;\mu,b,\tau} = 7.17$ and $\hat{\tau}_{4;,\mu,b,\tau,\beta} = 6.68$.

The summation notation yields the same estimates as the matrix notation

$\hat{\beta}$ = 504.6043/256.5000 = 1.9673 , and estimates of τ

not corrected for the covariate $(\hat{\tau}_{\mu,b,\tau})$

$\hat{\tau}_1$= (1/2)(55.8186) - 29.5438 = -1.6345 ,

$\hat{\tau}_2$= (1/2)(65.5050) - 29.5438 = 3.2086 ,

$\hat{\tau}_3$= (1/2)(41.5964) - 29.5438 = -8.7456 ,

$\hat{\tau}_4$= (1/2)(73.4380) - 29.5438 = 7.1716 , and

estimates of τ corrected for the covariate $(\hat{\tau}_{\mu,b,\tau,b})$

$\hat{\tau}_1$= -1.6345 - 1.9673(9.5 - 9.75) = -1.1427 ,

$\hat{\tau}_2$= 3.2086 - 1.9673(10   - 9.75) = 2.7168 ,

$\hat{\tau}_3$= -8.7456 - 1.9673(9.5 - 9.75) = -8.2538 ,

$\hat{\tau}_4$= 7.1715 - 1.9673(10   - 9.75) = 6.6798 .

The first set of summation formulas of Section 2.5.2 replaces the $\cdot\cdot$ of the summation formulas with yy,

$SSTot_{yy}$ = (17.3956)(17.3956)+...+(20.4594)(20.4594)
          = 8,388.1126 ,

SSM_{yy} = 1/(2·4) [(17.3956 +...+ 20.4594)
          (17.3956 +...+ 20.4594)] = 6,982.7126 ,

SSB_{yy} = (1/4)[(134.3276)^2 + (102.0232)^2]= 7,113.1594,

SST_{yy} = (1/2) [(55.8186)^2 + (65.5050)^2 + (41.5964)^2
          + (73.4308)^2] = 7,264.4820 ,

SSR_{yy} = 8,388.1126 - 7,113.1594 - 7,264.4820 +
          6,982.7126 = 993.1838 .

Recall that SSR is the sum of squares for the residual.

The sums of squares for terms in Z with subscripts zz are

$$SSTot_{zz} = (5)(5) + \ldots + (1)(1) \quad = \quad 1,078 \ ,$$

$$SSM_{zz} = 1/(2 \cdot 4)[\ (5 + \ldots + 1) \ (5 + \ldots + 1) \ ]= 760.5,$$

$$SSB_{zz} = (1/4) \ [\ (50)^2 + (28)^2 \ ] \quad = \quad 821 \ ,$$

$$SST_{zz} = (1/2) \ [\ (19)^2 + (20)^2 + (19)^2 + (20)^2 \ ] = 761,$$

$$SSR_{zz} = 1,078 - 821 - 761 + 760.5 \ = 256.5.$$

The sums of YZ crossproducts with subscripts yz are

$$SSTot_{yz}= (17.3956)(5) + \ldots + (20.4594)(1)=2,908.2419,$$

$$SSM_{yz} = 1/(2 \cdot 4) \ [(17.3956 +\ldots+ 20.4594) \ (5 +\ldots+ 1)]$$

$$= 2,304.4303 \ ,$$

$$SSB_{yz} = (1/4) \ [\ (134.3276)(50) + (102.0232)(28) \ ]$$

$$= 2,393.2574 \ ,$$

$$SST_{yz} = (1/2) \ [\ (55.8186)(19) + \ldots + (73.4308)(20) \ ]$$

$$= 2,314.8005 \ ,$$

$$SSR_{yz} = 2,908.2419 - 2,393.2574 - 2,314.8005 +$$

$$2,304.4203 \ = \ 504.6043 \ .$$

The values in Table 2.7.2 correspond to Table 2.5.1, as derived by the R(·|·) notation and "the Model Σ".

TABLE 2.7.2

COMBINED ANALYSIS OF VARIANCE TABLE - MATRIX COMPUTATION

Source of Variation	d.f.	Sum of Squares
Total corrected for the mean (μ)	n-1	8,388.1126 −6,982.7126 = 1,405.4000
Blocks after the mean (b|μ)	b-1	7,113.1594 −6,982.7126 = 130.4468
Treatments and interactions after the mean and blocks (t|μ,b)·	t-1	7,394.9288 −7,113.1594 = 281.7694
Covariates after the mean, blocks, treatments and interactions (β|μ,b,t)	q	992.6920 or 8,387.6208 −7,394.9288 = 992.6920
Covariates after the mean and blocks (β|μ,b)	q	1,031.9418 or 8,145.1011 −7,113.1594 = 1,031.9418
Treatments and inter- actions after the mean, blocks, and covariates (t|μ,b,β)	t-1	8,387.6208 −8,145.1011 = 242.5197
Residual error	n-p-q	0.4918 (by subtraction)

The numerical values of Table 2.7.3 correspond to Table 2.5.2 as derived by the summation notation.

TABLE 2.7.3

COMBINED ANALYSIS OF VARIANCE TABLE-SUMMATION COMPUTATION

Source of Variation	d.f.	Sum of Squares
Total corrected for the mean (μ)	8-1	8,388.1126 -6,982.7126 =1,405.4000
Blocks after the the mean (b$\mid\mu$)	2-1	7,113.1594 -6,982.7126 = 130.4468
Treatments after the mean and blocks ($\tau\mid\mu$,b)	4-1	7,264.4820 -6,982.7126 = 281.7694
Covariate after the mean, blocks, and treatments ($\beta\mid\mu$,b,τ)	1	$\dfrac{(504.6043)^2}{256.5000}$ = 992.6920
Covariate after the mean and blocks ($\beta\mid\mu$,b)	1	$\dfrac{(514.9845)^2}{257.0000}$ = 1,031.9418
Treatments after the mean, blocks, and covariate ($\tau\mid\mu$,b,β)	4-1	281.7694 + 992.6920 - 1,031.9418 = 242.5196
Residual error	8-2-4 = 2	993.1838 - 992.6920 = .4918

Tables 2.7.2 and 2.7.3 yield the same following F-statistics.   In each case, the OSL_F is effectively 1.0.

For the covariate adjusted for the mean, blocks and treatments, the F-statistic for testing the null hypothesis

$$H_{o;\mu,b,\tau,\beta} \quad \beta = 0$$

is $F(\beta|\mu,b,\tau) = [992.6920/1] \: / \: [.4918/2]$

$$= 4,036.9744.$$

For $\beta=2.0$ the noncentrality parameter is

$$\lambda_{\beta|\mu,b,\tau,\beta} = [1/(2\sigma^2)][1028-2]$$

$$= [1/(2\sigma^2)][1026] = [1/\sigma^2][513].$$

This test properly rejects the null hypothesis.

For the treatments adjusted for the mean and blocks, but not the covariate, the F-statistic for testing the null hypothesis

$$H_{o;\mu,b,\tau} \quad \tau = 0$$

is $F(\tau|\mu,b) = [281.7694/3] \: / \: [.4918/2]$

$$= 381.9566.$$

Since $\tau = 0$, the noncentrality parameter is zero. This test erroneously rejects the null hypothesis.

For the covariate adjusted for the mean and blocks, but not the treatments, the F-statistic for testing the null hypothesis $H_{o;\mu,b,\beta}$ $\beta = 0$

is $F(\beta|\mu,b) = [1,031.9418/1] / [.4918/2]$

$$= 4,196.5913.$$

For $\beta=2.0$, the noncentrality parameter is

$$\lambda_{\beta|\mu,b,\beta} = [1/(2\sigma^2)][1024] = [1/\sigma^2][512].$$

Both tests of β, adjusted for the treatments or not adjusted, properly reject the null hypothesis.

The F-statistic for testing the null hypothesis $H_{o:\mu,b,\tau,\beta}$ $\tau = 0$ (adjusted for the covariate) is

$$F(\tau|\mu,b,\beta)=[242.5196/3]/[.4918/2]=328.7510.$$

Since $\tau = 0$, the noncentrality parameter, λ, is zero. Both tests of τ, adjusted for the covariate or not adjusted, erroneously reject the null hypothesis.

We next prove a characteristic of projection matrices and digress a bit to discuss the implications of the proof.

Section 2.8

Matrices Producing Identical Projectors

Two different matrices may have identical projectors, $X(X'X)^-X'$, if, and only if, they are related by the conditions of Theorem 2.8.3, (p. 87). The first section introduces effects and coding matrices and the use their projectors play in the analysis of variance. The next section demonstrates the theorem. Sections 2.8.3 - 2.8.5 apply the theorem to, respectively, estimable functions, reductions in the sums of squares, and the estimator $\hat{\beta}$. The theorem is proved in Section 2.8.6. The final section is a literature review and a suggested reading list. Theorem 2.8.3 gives the conditions for reparamaterizing a model equation.

2.8.1 The Task of The Theorem

Consider an experiment testing one factor with four levels, using two experimental units per level. The model equation in matrix notation is

$$Y = X\beta + \varepsilon \qquad (2.8.1)$$

with Y an 8 by 1 vector of responses,

X an 8 by 4 design matrix,

β a 4 by 1 vector of parameters, and

ε an 8 by 1 vector of errors with mean

zero and variance-covariance matrix $I\sigma^2$.

One possible design matrix is the "effects" matrix, X_e below, which incorporates the usual constraints of Assumption 5, (p. 20). Using X_e, the resulting elements of $\hat{\beta}_e$, the usual least squares estimator of β_e, equal those obtained from the summation notation model, $\hat{\mu}$, $\hat{\tau}_1$, $\hat{\tau}_2$ and $\hat{\tau}_3$. The estimate for τ_4 is the negative of the sum of the estimates of τ_1, τ_2 and τ_3.

Other design matrices are possible, such as the dummy or coding matrix, X_C below. The first element of $\hat{\beta}_C$ will not equal $\hat{\mu}$, nor will the last three elements equal the $\hat{\tau}_k$, k=1,2,3. The estimates $\hat{\beta}_C$ are familiar to users of the SOLUTIONS option in the MODEL statement of PROC GLM in the SAS statistical computing package.

The two design matrices are

$$
X_e = \begin{bmatrix} 1 & | & 1 & 0 & 0 \\ 1 & | & 0 & 1 & 0 \\ 1 & | & 0 & 0 & 1 \\ 1 & | & -1 & -1 & -1 \\ 1 & | & 1 & 0 & 0 \\ 1 & | & 0 & 1 & 0 \\ 1 & | & 0 & 0 & 1 \\ 1 & | & -1 & -1 & -1 \end{bmatrix} \quad \text{and} \quad X_C = \begin{bmatrix} 1 & | & 1 & 0 & 0 \\ 1 & | & 0 & 1 & 0 \\ 1 & | & 0 & 0 & 1 \\ 1 & | & 0 & 0 & 0 \\ 1 & | & 1 & 0 & 0 \\ 1 & | & 0 & 1 & 0 \\ 1 & | & 0 & 0 & 1 \\ 1 & | & 0 & 0 & 0 \end{bmatrix} \qquad (2.8.2)
$$

The rank of X_e and of X_C equals the number of their columns, giving them full column rank and $X'X$ a true inverse. The matrices may be partitioned as

$$X_e = [\ X_\mu\ |\ X_{\tau e}\ ] \quad \text{and} \quad X_C = [\ X_\mu\ |\ X_{\tau C}\ ] . \qquad (2.8.3)$$

The full model sum of squares, due to fitting the

model equation with both mean and treatment effects, is

$$Y' \; [X(X'X)^{-1}X'] \; Y. \tag{2.8.4}$$

The product within the brackets is the projection matrix

or <u>projector</u>. Interestingly enough, either X_e or X_c or

many other forms of the model matrix produce the same

projector. When X does not have full column rank (2.8.4)

becomes

$$Y' \; [X(X'X)^{-}X'] \; Y, \tag{2.8.5}$$

where the minus sign indicates any member of the set of

generalized inverses defined in Section 2.8.6. The

numerical value (2.8.5) is the same for any of the

(nonunique) generalized inverses $(X'X)^{-}$ of $X'X$. (See for

example, Graybill [1983], Theorem 6.6.9, p. 134, or

Searle [1987], Theorem 7.1(14), p. 218.)

One would not expect the projector matrices to be

equal for every pair of matrices. One may ask what

relationship between two matrices, X_e and X_c, not neces-

sarily design matrices, results in the equality of their

projectors, and, conversely, what such equality tells us

about the relationship between arbitrary pairs of

matrices X_e and X_c. While the relationship is well

known, we find no proof of it in more commonly used

linear models textbooks. It appears to follow from the

Corollary of p. 129 of Zyskind, *et al*. [1964], who state

it, without proof, on their page 102. Pringle and Rayner

[1971] present the relationship under "Reparametrization" p. 88-90. Graybill [1976, p. 493-498, Theorem 13.2.7] proves equivalent relationships are required for reparameterization to hold. Section 2.8.2 states what the relationship must be; the proof is deferred until Section 2.8.6.

2.8.2 Theorem 2.8.3 and Two Numerical Examples

Theorem 2.8.3

For any matrices Σ_a (nxq), Σ_b (nxp), and M (pxq)

$$\Sigma_a(\Sigma_a'\Sigma_a)^-\Sigma_a' = \Sigma_b(\Sigma_b'\Sigma_b)^-\Sigma_b' \qquad (2.8.6)$$

$$<=> \begin{cases} \Sigma_a = \Sigma_b M \text{ for some } M, \text{ and} & (2.8.7) \\ \text{rank}(\Sigma_a) = \text{rank }(\Sigma_b) . & (2.8.8) \end{cases}$$

As an example, straightforward multiplication will show that the projection matrices are equal for Σ_e and Σ_c of (2.8.2). Their ranks are both equal to four and $\Sigma_c = \Sigma_e M$ for

$$M = 1/4 \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix} . \qquad (2.8.9)$$

The matrices need not be design matrices with elements of 0 and ± 1. Multiplication (computer assisted, one hopes) will confirm the equality of the projector matrices for the pair

$$
\Sigma_a = \begin{bmatrix}
234 & 250 & 311 & 284 & 144 & 234 \\
290 & 347 & 350 & 333 & 152 & 290 \\
170 & 175 & 265 & 189 & 101 & 170 \\
264 & 297 & 286 & 331 & 170 & 264 \\
185 & 229 & 365 & 287 & 161 & 185 \\
208 & 236 & 337 & 256 & 135 & 208 \\
218 & 230 & 227 & 216 & 96 & 218 \\
161 & 165 & 174 & 153 & 69 & 161
\end{bmatrix}
\quad \text{and}
$$

(8 x 6)
rank=4

$$
\Sigma_b = \begin{bmatrix}
7 & 5 & 8 & 3 & 7 & 5 & 8 \\
2 & 14 & 9 & 1 & 2 & 14 & 9 \\
3 & 2 & 7 & 9 & 3 & 2 & 7 \\
6 & 6 & 11 & 5 & 6 & 6 & 11 \\
5 & 1 & 13 & 5 & 5 & 1 & 13 \\
3 & 4 & 10 & 8 & 3 & 4 & 10 \\
3 & 9 & 4 & 3 & 3 & 9 & 4 \\
2 & 6 & 3 & 4 & 2 & 6 & 3
\end{bmatrix}
$$

(2.8.10)

(8 x 7)
rank=4.

The two matrices have the same number of rows and the same rank. They do not have the same number of columns.

One matrix connecting Σ_a and Σ_b of (2.8.10) by

$\Sigma_a = \Sigma_b M$ is

$$
M = \begin{bmatrix}
11 & 1 & 9 & 4 & 4 & 11 \\
8 & 7 & 5 & 9 & 3 & 8 \\
2 & 6 & 13 & 6 & 5 & 2 \\
7 & 4 & 8 & 2 & 1 & 7 \\
3 & 9 & 2 & 10 & 3 & 3 \\
7 & 9 & 5 & 3 & 1 & 7 \\
3 & 5 & 7 & 9 & 4 & 3
\end{bmatrix}
$$

(7 x 6).
rank=5

Two others are

$$
M = \begin{bmatrix}
7 & 5 & 5.5 & 7 & 3.5 & 7 \\
7.5 & 8 & 5 & 6 & 2 & 7.5 \\
2.5 & 5.5 & 10 & 7.5 & 4.5 & 2.5 \\
7 & 4 & 8 & 2 & 1 & 7 \\
7 & 5 & 5.5 & 7 & 3.5 & 7 \\
7.5 & 8 & 5 & 6 & 2 & 7.5 \\
2.5 & 5.5 & 10 & 7.5 & 4.5 & 2.5
\end{bmatrix}
= \Sigma_b^+ \Sigma_a, \text{ and}
$$

(7 x 6)
rank=4

$$\mathbb{M} = \begin{bmatrix} 14 & 10 & 11 & 14 & 7 & 14 \\ 15 & 16 & 10 & 12 & 4 & 15 \\ 5 & 11 & 20 & 15 & 9 & 5 \\ 7 & 4 & 8 & 2 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \Sigma_b^- \Sigma_a \,,$$

(7 x 6)
rank=4

where Σ_b^+ is the Moore-Penrose generalized inverse and Σ_b^- is a generalized inverse constructed by the method of Searle [1982, p.217]. Result R1 of Section 2.8.6, (p. 83), discusses such inverses.

Note that for each of the above five matrices, the rank equals neither the number of rows nor the number of columns. The rank of $\mathbb{M}$ is five; the rank is four for the other four matrices.

2.8.3 Estimable Functions

When $\mathbb{X}$ does not have full rank, the generalized inverse of $\mathbb{X}'\mathbb{X}$ is not unique, and consequently neither is the estimate of β, $\hat{\beta}$. However, certain linear functions of the elements of $\hat{\beta}$ do yield identical (invariant) values for all generalized inverses. Such invariant functions are called estimable functions, estimable because the estimate is not altered by (is invariant to) the choice of the generalized inverse. It is known that the linear function $\mathbb{A}\beta$ is estimable if and only if a matrix $\mathbb{B}$ exists such that $\mathbb{A} = \mathbb{B}\mathbb{X}$; also, that the "best" (minimum variance, unbiased) linear estimator of $\mathbb{A}\beta$ is $\mathbb{A}\hat{\beta}$. [See Graybill 1983, p. 134 and the Note ending p. 135, or Searle 1987, p. 221 and 287-288.]

Such an estimate is also invariant over all pairs of
matrices satisfying Theorem 2.8.3. Consider the
estimable function $\Lambda_a \beta_a$. It is estimated by $\Lambda_a \hat{\beta}_a$, where
X_a is used in the model equation to estimate β. Thus,

$$\Lambda_a \hat{\beta}_a = B X_a \hat{\beta}_a = B [\ X_a \ (X_a'X_a)^- \ X_a \ ] \ Y. \qquad (2.8.11)$$

Provided X_b satisfies Theorem 2.8.3, one may substitute
it into the above brackets and obtain

$$\Lambda_a \hat{\beta}_a = B [\ X_b \ (X_b'X_b)^- \ X_b \ ] \ Y = B X_b \hat{\beta}_b = \Lambda_b \hat{\beta}_b. \qquad (2.8.12a)$$

Because Λ_b equals $B X_b$, the function $\Lambda_b \beta_b$ is estimable and
its estimate equals the estimate of the function $\Lambda_a \beta_a$, as
shown in (2.8.12a).

2.8.4 Reductions in Sums of Squares

The model equation of (2.8.1) fits to the data both
the mean and treatment effects. The sum of squares due
to fitting this model is $R(\mu, \tau)$, and equals (2.8.4) (and
(2.8.5), (p. 76)). The sum of squares due to the mean
alone is $R(\mu)$. This sum of squares has the same form as
(2.8.4), but the X matrix is a column of 1's, as in X_μ of
(2.8.3), (p. 74). The reduction in the sum of squares
due to adding treatments, once the mean is already in the
model, is $R(\tau | \mu) = R(\mu, \tau) - R(\mu)$.

Provided the conditions of Theorem 2.8.3 are met,
as in (2.8.2), (p. 74), the reduction in the sum of
squares using the dummy or coding matrix is equal to the
reduction using the usual-constraints model matrix.

Each of the columns of Σ_{Te} of (2.8.3), (p. 74), is

orthogonal to the column of ones of Σ_{μ}. Consequently,

$R_e(\mu, \tau) = R_e(\mu) + R_e(\tau)$, and $R_e(\tau|\mu) = R_e(\mu, \tau) - R_e(\mu) =$

$R(\mu) + R_e(\tau) - R(\mu) = R_e(\tau)$. But this orthogonality is

not the case for Σ_{TC}, the dummy coding matrix, and for it

$R_C(\tau|\mu) \neq R_C(\tau)$. Theorem 2.8.3 tells us why the sums of

squares due only to treatments differ in the two model

matrices, that is, why $R_e(\tau) \neq R_C(\tau)$. (Removing a row of Σ_e

would correspond to one treatment's having fewer

observations than other treatments. In such an unbalanced

case, the columns of Σ_{Te} are not orthogonal to the column

of Σ_{μ} and $R_e(\mu, \tau) \neq R_e(\tau)$.)

Examination of the two design matrices shows that no

matrix M exists such that $\Sigma_{TC} = \Sigma_{Te}M$, as seen in

$$\Sigma_{Te} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix} \quad \text{and} \quad \Sigma_{TC} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} . \qquad (2.8.12b)$$

Any matrix M capable of transforming the top three rows

of Σ_{Te} into the top three rows of Σ_{TC} is incapable of

transforming the fourth row of Σ_{Te}, (-1 -1 -1), into

the fourth row of Σ_{TC}, (0 0 0). The columns of Σ_{TC} are

not linear combinations of the columns of Σ_{Te}. Their

column spaces differ, thus Theorem 2.8.3 cannot be

satisfied and the projector matrices cannot be equal. If
the column of 1's is included, then a $\mathbb{M}$ matrix does exist
and Theorem 2.8.3 applies as demonstrated in (2.8.9),
(p. 76).

2.8.5 Converting the Estimator $\hat{\beta}_c$ to $\hat{\beta}_e$

Theorem 2.8.3 provides an easy proof for Theorem
2.8.1 which is due to Ott [1977, first ed., p. 524], who
implies the sufficient conditions stated below. It
allows one to obtain the estimates ($\hat{\beta}_e$) from one design
matrix ($\mathbb{X}_e$) directly from the estimates ($\hat{\beta}_c$) obtained
from another design matrix ($\mathbb{X}_c$) without computing $\mathbb{X}_e'\mathbb{Y}$.
This theorem of Ott led to Section 3.8. It is a full-
rank reparameterization of β_c.

Theorem 2.8.1

For matrices $\mathbb{X}_c$ nxq, $\mathbb{X}_e$ nxp, and $\mathbb{M}$ pxq with

$$\mathbb{X}_c = \mathbb{X}_e\mathbb{M} \text{ and rank}(\mathbb{X}_e) = \text{rank}(\mathbb{X}_c) = p,$$

then $\hat{\mathbb{Y}}_e = \hat{\mathbb{Y}}_c$, and   (2.8.13)

$$\hat{\beta}_e = \{ (\mathbb{X}_e'\mathbb{X}_e)^{-1}\mathbb{X}_e'\mathbb{X}_c \} \hat{\beta}_c , \qquad (2.8.14)$$

where $\hat{\mathbb{Y}}_{(\cdot)}$ and $\hat{\beta}_{(\cdot)}$ are the usual least squares
estimators.

Proof.

Since the conditions for Theorem 2.8.3 are satisfied, one may replace the projector matrix for $\mathbf{X}_e$ with that of $\mathbf{X}_c$ and obtain

$$\hat{\mathbf{Y}}_e \equiv \mathbf{X}_e \hat{\boldsymbol{\beta}}_e = \mathbf{X}_e (\mathbf{X}_e' \mathbf{X}_e)^{-1} \mathbf{X}_e' \mathbf{Y}$$

$$= \mathbf{X}_c (\mathbf{X}_c' \mathbf{X}_c)^{-} \mathbf{X}_c' \mathbf{Y} = \mathbf{X}_c \hat{\boldsymbol{\beta}}_c \equiv \hat{\mathbf{Y}}_c \quad . \qquad (2.8.15)$$

Then one equates $\mathbf{X}_e \hat{\boldsymbol{\beta}}_e$ and $\mathbf{X}_c \hat{\boldsymbol{\beta}}_c$ and solves for $\hat{\boldsymbol{\beta}}_e$. The assumption that rank$(\mathbf{X}_e) = p$, its number of columns, gives $\mathbf{X}_e' \mathbf{X}_e$ a true inverse and thus the unique solution for $\hat{\boldsymbol{\beta}}_e$.

A use of (2.8.14) is to convert the SAS PROC GLM solution vector to the usual-constraints solution. Using PROC IML, one constructs $\mathbf{X}_e$ and $\mathbf{X}_c$ as in (2.8.2) and obtains the matrix in the brackets of (2.8.14). Then, dropping the elements of $\hat{\boldsymbol{\beta}}_c$ which SAS has set to zero, one obtains $\hat{\boldsymbol{\beta}}_e$ via (2.8.14). Finally, one equates the estimate for the level with the largest subscript, here $\hat{\tau}_4$, to the negative of the sum of the estimates for the other levels, here $-[\ \hat{\tau}_1 + \hat{\tau}_2 + \hat{\tau}_3\ ]$.

2.8.6 Proof of Theorem 2.8.3

Results R1 through R8 are used to prove Theorem 2.8.3. All matrices have real elements.

R1. For any matrix A there is a generalized inverse matrix, G, satisfying any one or more of the following conditions

R1.1 $AGA = A$,

R1.2 $GAG = G$ reflexive,

R1.3 $(AG)' = AG$ symmetric, and

R1.4 $(GA)' = GA$ symmetric.

[Ben-Israel and Greville 1980, p. 7-22.]

No notation is standard. We use $A^{(1,2,3)}$ to indicate membership in the set of generalized inverses G satisfying conditions R1.1, R1.2, and R1.3. Others use G_{1r}^-, the letter "l" to indicate the set of least squares generalized inverses $\{A^{(1,3)}\}$, and the "r" to indicate the set of reflexive generalized inverses $\{A^{(1,2)}\}$. We drop the set indication $\{\cdot\}$; it is implied throughout. For $A^{(1,2,4)}$ another notation is G_{mr}^-, the set of minimum-norm reflexive generalized inverses. Only the Moore-Penrose generalized inverse, $A^{(1,2,3,4)} \equiv A^+$ is unique. A generalized inverse satisfying at least condition R1.1 is denoted A^-. The particular matrix selected is immaterial provided it is a member of the class indicated by the superscript(s).

R2. rank$[\Sigma]$ = rank$[\Sigma(\Sigma'\Sigma)^-\Sigma']$. [Graybill 1983, Theorem 6.6.8, p. 134.]

R3. $\Sigma(\Sigma'\Sigma)^-\Sigma' = \Sigma\Sigma^+$. For Σ nxp, all generalized inverses, including Σ^+, are pxn; thus the multiplication of the right hand side is permitted. When Σ has full column rank $\Sigma(\Sigma'\Sigma)^{-1}\Sigma'$ also equals $\Sigma\Sigma^+$. [Graybill 1983, Corollary 6.6.9.3, p. 135, or Searle 1982, p. 221-222.]

R4. For any matrices A and B of suitable dimensions, the Moore-Penrose generalized inverse of their product, $(AB)^+$, can always be expressed as

$$(AB)^+ = B^{(1,2,4)}A^{(1,2,3)} \qquad (= B^-_{mr}A^-_{lr}).$$

[Shinozake and Sibuya 1974, Theorem 3.2, p. 34.]

R5. For a $A^{(1,2,3)}$ such that $B^{(1,2,4)}$ is a solution to $(AB)^+$ of R4, the form of $B^{(1,2,4)}$ may be written as

$$B^{(1,2,4)} = B^+ + \{ (AB)^+A - [I - (AB)^+AB]B^+A^{(1,2,3)}A \}$$

$$x \{ I - BB^+ \}$$

$$+ K\{ I - [(I - BB^+)A^{(1,2,3)}] [(I - BB^+A^{(1,2,3)})] \}$$

$$x \{ I - BB^+ \},$$

where K is an arbitrary qxp matrix, A is nxp, and B is pxq. [Wibker, Howe, and Gilbert 1979, Theorem 4, Equation 17, p. 112.] That such $A^{(1,2,3)}$ and $B^{(1,2,4)}$ exist follows from R4.

R6. <u>Theorem 2.8.2.</u>

For $A^{(1,2,3)}$ and $B^{(1,2,4)}$ satisfying R5,

$$ABB^{(1,2,4)}A^{(1,2,3)} = (AB)(AB)^+AA^{(1,2,3)} \ .$$

The proof proceeds by postmultiplying the equation

of R5 by $A^{(1,2,3)}$. This leaves the third term equal to

zero as shown by straightforward multiplication and by

Wibker, Howe, and Gilbert [1979] on the bottom of their

page 112. Thus, the third term is dropped.

Removing the internal grouping in the second term of

R5 gives

$$B^{(1,2,4)} = B^+ + \{ (AB)^+A - [\ I - (AB)^+(AB)] \ B^+A^{(1,2,3)}A \}$$

$$\times \{ \ I - BB^+ \}$$

$$= B^+ + \{(AB)^+A - B^+A^{(1,2,3)}A + (AB)^+(AB)B^+A^{(1,2,3)}A\}$$

$$\times \{ \ I - BB^+ \}$$

$$= B^+ + (AB)^+A \ - \ B^+A^{(1,2,3)}A$$

$$+ \ (AB)^+(AB)B^+A^{(1,2,3)}A$$

$$- \ (AB)^+(AB)B^+$$

$$+ \ B^+A^{(1,2,3)}(AB)B^+$$

$$- \ (AB)^+(AB)B^+A^{(1,2,3)}(AB)B^+ \ .$$

Premultiplication by AB and postmultiplication by $A^{(1,2,3)}$ gives

$$AB B^{(1,2,4)} A^{(1,2,3)}$$

$$= (AB) B^+ A^{(1,2,3)} + (AB)(AB)^+ AA^{(1,2,3)}$$

$$- (AB) B^+ [A^{(1,2,3)} AA^{(1,2,3)}]$$

$$+ [(AB)(AB)^+(AB)] B^+ [A^{(1,2,3)} AA^{(1,2,3)}]$$

$$- [(AB)(AB)^+(AB)] B^+ A^{(1,2,3)}$$

$$+ (AB) B^+ A^{(1,2,3)} (AB) B^+ A^{(1,2,3)}$$

$$- [(AB)(AB)^+(AB)] B^+ A^{(1,2,3)} (AB) B^+ A^{(1,2,3)} .$$

Applying R1.1, or R1.2 as appropriate, to the terms within the brackets gives

$$AB B^{(1,2,4)} A^{(1,2,3)}$$

$$= (AB) B^+ A^{(1,2,3)} + (AB)(AB)^+ AA^{(1,2,3)}$$

$$- (AB) B^+ A^{(1,2,3)}$$

$$+ (AB) B^+ A^{(1,2,3)}$$

$$- (AB) B^+ A^{(1,2,3)}$$

$$+ (AB) B^+ A^{(1,2,3)} (AB) B^+ A^{(1,2,3)}$$

$$- (AB) B^+ A^{(1,2,3)} (AB) B^+ A^{(1,2,3)} .$$

Cancellation of terms gives the result.

R7.   If $\text{rank}(AB) = \text{rank}(A) \Rightarrow [B(AB)^-] = A^-$.
[Searle 1982, Exercise 39(a), p. 226.]

R8.   For any matrix A, $AA^{(1,2,3)} = AA^+$.
[Wibker, Howe, and Gilbert 1979, p. 109.]

We now repeat the statement of and prove Theorem 2.8.3.

<u>Theorem 2.8.3.</u>

For any matrices $\mathbb{X}_s$ (nxq), $\mathbb{X}_a$ (nxp), and $\mathbb{M}$ (pxq)

$$\mathbb{X}_a(\mathbb{X}_a'\mathbb{X}_a)^-\mathbb{X}_a' \;=\; \mathbb{X}_b(\mathbb{X}_b'\mathbb{X}_b)^-\mathbb{X}_b' \qquad (2.8.16)$$

$$<=> \left\{ \begin{array}{ll} \mathbb{X}_a = \mathbb{X}_b\mathbb{M} \text{ for some } \mathbb{M}, \text{ and} & (2.8.17) \\ \text{rank}(\mathbb{X}_a) = \text{rank}(\mathbb{X}_b) \ . & (2.8.18) \end{array} \right.$$

Proof.

$(2.8.16) \Rightarrow (2.8.18)$ by R2.

By R3 one may replace (2.8.16) with

$$\mathbb{X}_a\mathbb{X}_a^+ = \mathbb{X}_b\mathbb{X}_b^+ \ . \qquad (2.8.16*)$$

$(2.8.16*) \Rightarrow (2.8.17)$

Multiplying (2.8.16*) on the left by $\mathbb{X}_a$ and using R1.1 gives

$$\mathbb{X}_a = \mathbb{X}_a\mathbb{X}_a^+\mathbb{X}_a \;=\; \mathbb{X}_b(\mathbb{X}_b^+\mathbb{X}_a) \text{ , or}$$

$$\mathbb{X}_a = \mathbb{X}_b\mathbb{M}^* \text{ , with } \mathbb{M}^* \equiv \mathbb{X}_b^+\mathbb{X}_a.$$

Note that $\mathbb{M}^*$ need not equal $\mathbb{M}$; $\mathbb{M}$ need not be unique.

$(2.8.17)$ and $(2.8.18) \Rightarrow (2.8.16)$

By assumption (2.8.17), $\mathbb{X}_a\mathbb{X}_a^+ = \mathbb{X}_b\mathbb{M}(\mathbb{X}_b\mathbb{M})^+$.

Applying R4 to $(\mathbb{X}_b\mathbb{M})$ gives

$$\mathbb{X}_a\mathbb{X}_a^+ = \mathbb{X}_b\mathbb{M}\mathbb{M}^{(1,2,4)}\mathbb{X}_b^{(1,2,3)} \ .$$

Applying R6 (Theorem 2.8.2) to the entire right hand side of the above gives

$$X_a X_a^+ = X_b \ [M(X_b M)^+] \ X_b X_b^{(1,2,3)} \ .$$

By assumption (2.8.17) $X_b M = X_a$; assumption (2.8.18) adds rank$(X_b M)$ = rank(X_b), the condition for R7, which, applied to the term within brackets [] above, gives

$$X_a X_a^+ = X_b [X_b^-] X_b X_b^{(1,2,3)} \ .$$

Applying R1.1 to the first three matrices of the right hand side gives $\ X_a X_a^+ = X_b X_b^{(1,2,3)} \ .$

Applying R8 gives the result $\ X_a X_a^+ = X_b X_b^+ \ .$

Finally, R3 permits returning to the desired form

$$X_a (X_a' X_a)^- X_a' \ = \ X_b (X_b' X_b)^- X_b' \ .$$

Note that Assumption (2.8.18) requires rank(X_a) ≤ rank(M) (and rank(X_b) ≤ rank(M)), but none of the matrices need be of full row or full column rank. The matrix M need not be square, in which case $p \neq q$.

Proof of Theorem 2.8.3 is straightforward under the more restricted cases of

a) X_a and X_b have full column rank ($p=q$) and M has a true inverse M^{-1}, and

b) $(X_a M)^+ = M^+ X_a^+$ and M has a true inverse.

The difficulty in (b) is to set the minimum restrictions

on Σ_a and M such that $(\Sigma_a M)^+ = M^+ \Sigma_a^+$. These restrictions

must link both the column and null spaces of the two

matrices. See Greville [1966] for conditions under which

this equality holds. Shinozake and Sibuya's Theorem 3.2

[1974] permits the use of the less restrictive inverses

$M^{(1,2,4)}$ and $\Sigma_a^{(1,2,3)}$, thus avoiding the difficulty.

The projection matrix for Σ_a is a projector on the

range or column space or manifold (equivalent terms) of

Σ_a along the null space of Σ_a, and likewise for the

projection matrix for Σ_b. The equality of the two

projectors is the equality of the column spaces of Σ_a and

Σ_b and the equality of the null spaces of Σ_a and Σ_b.

Postmultiplication of Σ_b by M generates a new member in

the set of all linear combinations of the column vectors

of Σ_b. Thus, Σ_a has the same column space as Σ_b.

Conditions (2.8.17) and (2.8.18) of Theorem 2.8.3 imply

the equality of the column spaces of Σ_a, Σ_b, and hence of

their respective projectors. Thus, their orthogonal

basis may be chosen to be identical, giving identical

orthogonal complements and null spaces. (See Graybill

[1983], Theorem 5.4.3, p. 87 and Corollary 5.4.4, p. 89.)

2.8.7   Readings in Generalized Inverses

The novice to the topic of generalized inverses with courses in linear algebra and linear models could begin with Chapter 8 of Searle [1982]. Several numerical examples introduce the topic. Chapter 6 of Graybill [1983] covers the Moore-Penrose and generalized inverses. Care should be taken with his notation, especially if simultaneously reading Searle. Cline [1979] has geometric illustrations as well as numerical examples and exercises. He focuses on the Moore-Penrose inverse; later chapters touch on other generalized inverses and introduce the Drazin inverse.

The first three chapters of Ben-Israel and Greville [1980] contain a good introduction to the multi-condition generalized inverses, such as $A^{(1,2,3)}$, used in Theorem 2.8.3 and in statistical applications. Subsequent chapters cover applications in areas of mathematics outside of statistics. Boullion and Odell [1971] present partitioned matrices and statistical applications including sequential least squares, quadratic forms, and stochastic matrices. Their Notes and Comments provide useful insights and direct the reader to the major and more readable early papers. Pringle and Rayner [1971] bring together many results on generalized inverses. They discuss the multi-condition generalized inverses and

how to compute them, the over-parameterized linear model
(X not of full column rank), and the linear model with a
singular variance-covariance matrix.

Rao and Mitra [1971] thoroughly cover the statis-
tistical applications of generalized inverses, including
the case of a singular variance-covariance matrix. They
discuss many additional multi-condition matrices created
by attaching additional characteristics to those of R1.
The novice will find the going there easier after
grasping the suggested parts of Searle, Graybill,
Ben-Israel and Greville, and Cline. The specialist
should consult Nashed [1976] for an annotated bibliog-
raphy nearly complete up to 1976. Zielke [1978] and
Campbell [1982] provide more recent bibliographies.

We return to the normally distributed random error
model with a discussion of early understanding of this
type of error. As Chapter Three will make clear, there
are types of errors which do not have the characteris-
tics discussed in the following section.

Section 2.9

Early Understanding of Random Error

Section 2.9 illustrates the early understanding of random error using citations from Galileo and Gauss. The citations are grouped into fourteen categories. The chapter ends by introducing two different types of errors and contrasting their characteristics with those of the errors discussed by Galileo and Gauss. Chapter Three investigates these two different types of error.

The linkage of normally distributed random errors with least squares estimation grew out of the expansion of global navigation and nationwide surveying, and the beginning of modern astronomy in the late 1500's in Europe. Workers in these areas made careful and repeated measurements of an object assumed to be in a fixed place, for example a star. As few, if any, of the measurements precisely equaled one another, workers came to realize that all measurements erred from the true value. While improving their methods and instruments, they also realized that errors inevitably remained and described these errors.

The following excerpts illustrate the thinking of the times and the understanding of the errors which today are addressed by the normal distribution model and least squares estimation technique. Providing the excerpts are Galileo Galilei [GG] (1564-1642) from his work in stellar distances and Karl Friedrich Gauss [KG] (1777-1885) from

his work in the determination of orbits.   Similar

perceptions are found in the writings of others

throughout the two centuries.   Maistrov [1974, pages

30-35 and 148-157] presents a more extensive review of

the historical development of the theory of errors.   See

also Sheynin [1966].

1.   A unique true value exists; errors are deviations

from it.

> Doubtless one must say that [the supernova
>
> of 1572] was located in a single place at a
>
> unique and determinate distance from the
>
> earth.   [GG 281]

> Whenever the calculations made from the
>
> observations . . . do not agree in putting it
>
> in the same place, there must be errors in the
>
> observations [GG 289]

> The quantities ε, ε', ε'', ε''', etc.
>
> will furnish as many different determinations
>
> of the same quantity. [KG 250]

2.   Errors are measured in units of the observations,

not in units of values calculated from them.

> The size of the instrumental errors, so
>
> to speak, must not be reckoned from the
>
> outcome of the calculation, but according to
>
> the number of degrees and minutes actually
>
> counted on the [astronomical] instrument.
>
> [GG 293]

If the astronomical observations and

other quantities, on which the computation of

orbits is based, were absolutely correct... .

[KG 249]

3.   Types of errors form a continuum from systematic to
random.

[Systematic errors] may be predicted and

evaluated since these are either constant or

vary in a regular manner. . . . Such a division

of error into two kinds is relative, and in

many cases depends upon the problem at hand.

[KG in Maistrov, page 155]

4.   Both types of error may be reduced by experimental
technique and measuring devices.

Tycho Brahe [the astronomer] measured the same

object under varied conditions for the purpose

of elimination of errors. [Maistrov, p. 152]

What faith can we have in calculations

founded upon observations which . . . are made

with less convenient and more unreliable

instruments? [GG 310, see also GG 315-316]

5.   Physically impossible values should be discarded.

Whenever the calculations imply that the

two angles [of a triangle] exceed two right

angles, the observations are to be taken as

unquestionably mistaken. [GG 289, see also GG

290-291]

6.   Random errors remain after systematic errors have
been removed.

> In the differences between the observed
> and computed places there will remain no trace
> of any law, which it would be possible to
> remove or sensibly diminish by a correction of
> the elements. [KG 252]

7.   Random errors are inevitable.

> We know too well that in the taking of only
> one altitude of the pole [Polaris] with the
> same instrument, in the same place, by the same
> observer who has repeated the observation a
> thousand times, there will still be a variance
> of one, or sometimes of many minutes. [GG 290]

> All our measurements and observations are
> nothing more than approximations to the truth.
> [KG 249]

8.   Small random errors are more likely than large
ones.

> One must believe that they would be more
> likely to err little than much. [GG 290]

> Generally speaking, small errors oftener
> occur than large ones. [KG 253]

9.   Positive random errors are as likely as negative
random errors.

> Or must the errors be always of one kind, so
> that when they err they are always mistaken by
> an excess, or always be a defect and never by
> an excess?

> I do not doubt that they are equally prone to
> err in one direction [as] the other. [GG 291]

> [As from the data no reason exists otherwise]
> we must assume, [based] upon the principles of
> probability, that greater or less errors are
> equally possible. [KG 253]

10.   The distribution of random errors is symmetric.

> The probability to be assigned to each
> error $\triangle$ [delta]. . . should be . . . equal,
> generally, for equal opposite values of $\triangle$. [KG 254]

11.   The maximum probability of the true value is that
which gives a zero value for the random error as
measured from the true value.

> And among the possible places, the actual place
> must be believed to be that in which concur the
> greatest number of distances, calculated on
> the most exact observations. [GG 293]

> The probability to be assigned to each
> error $\triangle$ . . . should be a maximum for $\triangle = 0$.
> [KG 254]

12. For theoretical purposes, positive probability is assigned to large random errors.

> The probability to be assigned to each
>
> error $\triangle$ . . . should vanish, if, for $\triangle$ is taken
>
> the greatest error, or value greater than the
>
> greatest error . . . [and the probability must]
>
> converge to zero on both sides, asymptotically
>
> as it were from $\triangle$ = 0. [KG 254]

13. Random errors are mutually independent.

> Wherefore, since we are authorized to regard
>
> all observations as events independent of each
>
> other. . . . [KG 255] [The translation, at least,
>
> is obscure as to the source of the authorization.]

14. Multiple observations of the same object are to be combined by averaging.

> If the degree of accuracy is to be presumed
>
> in all the observations . . . [it] must
>
> . . . be considered an axiom with the same
>
> propriety as the [axiom that the] arithmetical
>
> mean of several observed values of the same
>
> quantity is adopted as the most probable value.
>
> [KG 260]

> But if . . . the [astronomical] observations
>
> are distant from each other by too great an
>
> interval of time, . . . so that it would not be
>
> admissible to regard their deviations as
>
> constant for all observations . . . [then] the

mean deviation . . . cannot be regarded as

common to all the observations. [KG 260]

Many experimental studies differ in an important

aspect from the studies in navigation and astronomy which

lead to the normally distributed random errors theory.

In these experimental studies the observations, which are

combined, are no longer made upon the same object. When

such differing objects or experimental units are organic

in nature, they are seldom identical in every aspect,

even when the "same" experimental unit is measured at a

later date.

When the experimental units differ, different

responses, as measured without error, may arise from

units receiving the same treatment. Such errors are not

the result of errors in multiple observations of the same

or physically identical objects. They arise indepen-

dently of the error due to inadvertent changes in

experimental conditions and/or subsequent measurements,

that is, measurement errors. There is no a priori reason

to expect the distribution of these errors to fit the

above descriptions, especially numbers 8, 9, 10, and 12;

number 11 does not apply. Physically adjacent experi-

mental units in agricultural studies, such as fields or

plots, may not be independent, violating number 13.

Some organic experimental units may have a higher or lower response regardless of the treatment applied. Such units are more robust or less so than others. Such deviations from the mean response are experimental unit errors.

An experimental unit may perform like others receiving treatment A, but quite differently from those receiving treatment B when it also receives treatment B. This is experimental unit x treatment interaction error.

Either type of error causes an instability in the combined data not present in the studies of Galileo and Gauss or, typically, in studies using inorganic experimental units. For example, the outcome may be altered if most of the robust experimental units receive treatment A while most of the nonrobust units receive treatment B, as opposed to the reverse assignment of treatments. Or, should the penicillin-allergic experimental units receive the placebo, all will go well; not so if the penicillin goes to the allergic experimental units. In extreme cases, the outcome of the experiment depends upon the randomization used to link treatments with experimental units. There appears to be no provision within the normally distributed random errors theory for this type of instability of results.

If either experimental unit error and/or experimental unit x treatment interaction error is an amalgamation of elementary errors, one could employ a central limit

theorem to support treating its distribution as approximately normal.   For example, compare the three assumptions of Hagen's Hypothesis [C.R. Rao, 1973, p. 161] with, respectively, excerpts 6, 13, and 9 above.   But, these errors need not be an amalgamation; they may arise from a single cause, different perhaps for each experimental unit.   The measurement error is the more likely to be an amalgamation of elementary errors satisfying some central limit theorem, and having a distribution approximately normal.

One could argue that such errors arise from extreme experimental units which form a different population and, hence, should be removed from the experiment.   An example is inclusion in the study population of only those whose skin test shows them not to be allergic to penicillin. However, such a test may not be known and, if known, would need be done in both the experimental material and the future applications of the results of the experiment. Furthermore, one may need to test for a plethora of conditions.

Chapter Three examines such such errors and the implications they have upon the classical analysis of covariance.

CHAPTER III

THE RANDOMIZATION MODEL

Section 3.1

Introduction

Suppose every experimental unit of the study is assigned, at the same time, to treatment k. Suppose further that the responses from all experimental units are identical. Yet further, suppose that this holds when the same experimental units are assigned en masse to each of the other treatments. In such a situation, the normally distributed random errors theory of Chapter 2 would suffice. All errors would be failures to properly measure the response, or failures to apply the treatments identically, and/or the effects of factors not measured. But, in some investigations, experimental units and treatments are such that experimental units receiving the same treatment consistently fail to respond identically. Some experimental units may respond with higher yields to all treatment, others may respond with yields higher to some treatments and lower to others. Chapter Three examines the analysis of covariance as applied to such experimental units.

The foundation is the randomization model. This model considers the conceptual population of values actually obtained if the experiment were conducted anew using each of the possible assignments of treatments to experimental units. Reassigning all treatments to experimental units, each time repeating the experiment, is seldom, if ever, possible in practice. At the very least the experimental units are older and have been exposed to the treatments of previous experiments. Thus, the population of outcomes is conceptual, existing only in artificial simulations.

Subsequent chapter subheadings and topics are

Section 3.2

Randomization

Section 3.2.1 introduces the experimental unit assignment probability. Sections 3.2.2 - 3.2.4 discuss methods of linking experimental units and treatments. Section 3.2.5 links assignment and selection probabilities. Sections 3.2.6 and 3.2.7 discusses restricted randomization and unequal selection probabilities. The final section, 3.2.8, discusses why one randomizes. Major results are the $N(\mathcal{E})$ of (3.2.4), (p. 108), the selection process of (3.2.7), (p. 109), and Theorem 3.2.1, (p. 112), linking the two types of probabilities.

3.2.1 Definition of Randomization

Consider one fixed treatment, k. By <u>randomization</u> we mean that the experimental unit(s) assigned to treatment k are not fixed <i>a priori</i>. Instead, the experimental unit(s) are assigned by a process which makes the treatment to which each unit is assigned a random event. The process must uphold the assumptions of the model, here being Assumption 8, (p. 22), that randomization occurs independently within each block, and Assumption 5, (p. 20), which is satisfied by an equal number, r, of experimental units per treatment per block.

The notation for experimental unit ijl receiving treatment k is U_{ijlk}. Within the $i^{\underline{th}}$ block, r times p

experimental units are available, of which r are assigned
to each treatment. The subscripts l (the letter el) and
j have values of, respectively, l=1,2,...,r, and
j=1,2,...,p. As before, p=t, the number of treatments.
When there is but one experimental unit per treatment per
block, r=1, and the subscript l is dropped.

One way to randomly link experimental units to
treatments is to select the experimental units at random
from a larger, perhaps theoretically infinite, popula-
tion. For some authors, randomization implies this.
Here, the experimental units are the finite number in
hand and awaiting assignment to treatments. The random
linkage is of these experimental units to the treatments.
Section 3.5.5, (p. 163), considers these experimental
units as part of a larger population.

"Randomization" here has nothing to do with the survey
sampling technique also called randomization or randomized
response used in obtaining probabilistic answers to
sensitive polling questions, as in Warner [1965].

Many randomization processes are possible. All
begin by uniquely labeling experimental units and
treatments. Directly or indirectly, all assign to each
$jl^{\underline{th}}$ experimental unit of block i, an experimental unit
assignment probability for each treatment, k. This is
denoted as $p(ijl:k)$, the probability that experimental
unit U_{ijl} is assigned to treatment k. Typically, the
assignment is made "at random," meaning with equal

probability. Such a process makes all $p(ijl:k)$ equal.

3.2.2 A Sequential Selection Process

The random process may be (a) a sequential selection process, (b) a mass draw process, or (c) a multi-stage process combining aspects of (a) and (b).

One sequential selection process selects, "at random", p of the experimental units and assigns them one by one "at random" to each of the t treatments, then repeats the process r times. As the two processes of (a) selection into group r and (b) assignment to treatment k operate independently,

$$p(ijl:k \mid r) = p(ijl:k, \ r)/p(r)$$
$$= p(ijl:k)p(r)/p(r)$$
$$= p(ijl:k) \ , \hspace{2cm} (3.2.1)$$

where $p(r)$ is the probability that U_{ijl} is in group r.

3.2.3 A Two-Stage Selection Process

A process in-between the sequential and mass draw processes proceeds block-wise in two stages per block. The first stage lists the set $\mathfrak{G}$ of unique, distinguishable groupings of the pr=tr experimental units of a block into t groups of r experimental units each. This set has $N(\mathfrak{G})$ elements. The example below lists four such groupings. There are (rt)! ways of ordering, or listing, the rt experimental units. However, not all are distinguishable for purposes of the test statistics and/or estimators. For each list, such as (a) below, one

divides the rt experimental units into t subgroups of r

each.   The first r comprise subgroup 1, the second r

comprise subgroup two and so on for each list.   The

parentheses of the example indicate this grouping.   Sub-

groups with the same experimental units, regardless of

the order of the experimental units, are indistinguish-

able from one another, such as (a) and (b) of the

example.   Each subgroup of the same r experimental units

may be ordered in r! ways.   This holds for each of the t

subgroups.   Thus, the initial (rt)! lists, each of rt

experimental units, reduce to $(rt)!/[(r!)^t]$ lists, each

of t subgroups of r experimental units each.   The

groupings of the reduced list are distinguishable.

The t subgroups can be arranged in t! ways which,

with an assumption, are indistinguishable, as in (a),

(b) and (c) of the example.   The assumption is that the

function of the data used subsequently, as in a

hypothesis test, does not depend upon the order of

arranging the subgroups, nor on the order of the

experimental units within each subgroup.   In the example,

this is the ordering of (a) the subgroups bound by

parentheses and (b) the experimental units within each

subgroup.   This is the case for the usual estimators,

(p. 30), and F-ratio test statistics, (p. 62).   This is

not the case when t=2 and the test statistic is the one-

tailed Student's t-test.   In such a case, there are 2N(ₛ)

unique, distinguishable groupings of the experimental

units, half of which differ only in the sign of their
test statistic from the other half.   The example below
illustrates this by considering only the first two
columns (t=2).   Note that sets (a) and (b) become
distinguishable from (c).   For fully-worked examples
using the t-test, see Kempthorne [1952, p. 232] or
Kempthorne and Doerfler [1969, p. 232, Table 1].   For the
functions used herein, the initial (rt)! lists reduce to
N($\mathfrak{G}$), the number of unique, distinguishable groupings of
experimental units per block,

$$N(\mathfrak{G}) = \left\{ \frac{(rt)!}{(r!)^{t} \, t!} \right\} . \qquad\qquad (3.2.2)$$

The first stage ends by selecting, "at random," one of
the N($\mathfrak{G}$) unique, distinguishable groupings.   This stage
is a mass draw process.

The example below lists four of the (rt)! orderings
of the experimental units.   The numbers represent the
experimental units.   There are t=3 subgroups of r=2
experimental units each; $N(\mathfrak{G}) = (6!)/[(2^{3})(3!)] =$
720/[(8)(6)] = 15.   Note that (a), (b) and (c) are
indistinguishable from each other, while each is
distinguishable from (d).   Thus, only two elements of the
set $\mathfrak{G}$ are listed below,

$$
\begin{array}{llll}
\text{(a)} & (1, \ 2) & (3, \ 4) & (5, \ 6) \\
\text{(b)} & (2, \ 1) & (4, \ 3) & (5, \ 6) \\
\text{(c)} & (3, \ 4) & (1, \ 2) & (5, \ 6) \\
\text{(d)} & (4, \ 1) & (2, \ 5) & (3, \ 6) \quad .
\end{array}
$$

The second stage begins with the single grouping selected at the end of stage one, such as the grouping which includes (a), (b) and (c) above. Then one selects "at random" one of the t sets (of r experimental units each) to assign to treatment k=1. Repeat this for each of the k=2,3,...,t treatments. There are t! ways of assigning the t sets of experimental units to the t treatments. As described, the second stage is a sequential assignment process, but it can be performed as a mass draw process.

The set of all <u>block-wise</u> <u>randomizations</u> is $\mathcal{B}$. Combining stages one and two gives as the total number of unique, block-wise randomizations,

$$N(\mathcal{B}) = \{N(\mathcal{G})\}\{t!\}$$

$$= \{(rt)!/[(r!)^t t!]\}\{t!\}$$

$$= (rt)!/[(r!)^t] \qquad (3.2.3)$$

This is the multinomial coefficient for distributing rt objects into t groups of r objects each. When t=2, (3.2.3) reduces to $\binom{rt}{r}$, as in the above cited examples of the t-test. When t=3 and r=2, as in the example, $N(\mathcal{B})=(3\cdot2)!/(2!)^3 =720/8 =90$.

The set of all <u>experiment-wise</u> <u>randomizations</u> is $\mathcal{E}$. For a model which randomizes independently in each of b blocks (Assumption 8), there are

$$N(\mathcal{E}) = \{N(\mathcal{B})\}^b = \{N(\mathcal{G})t!\}^b = \{(rt)!/[(r!)^t]\}^b \qquad (3.2.4)$$

unique, experiment-wise randomizations. When there is

but one experimental unit per treatment per block, r=1,

N($\mathfrak{G}$)=1 and the numbers of unique block- and

experiment-wise randomizations are, respectively,

$$N(\mathfrak{B}) = (t!) \quad \text{and} \qquad (3.2.5)$$
$$N(\mathfrak{E}) = (t!)^b. \qquad (3.2.6)$$

The values of N($\mathfrak{G}$), N($\mathfrak{B}$) and N($\mathfrak{E}$) in (3.2.2) − (3.2.6)

are the largest consistent with the model assumptions of

Chapter Two; smaller values may also be consistent, as in

(3.2.12) and (3.2.13), found below (p. 115).

The two-stage process permits unequal probabilities

at each stage. Restricted randomization, discussed in

Section 3.2.6, (p. 114), is an example of this.

3.2.4 A Mass Draw Selection Process

The mass draw process conceptually performs all of

the possible experiment-wise randomizations. The experi-

mental units are not physically assigned all possible

ways. Rather, the N($\mathfrak{E}$) experiment-wise randomization

labels are ordered in an arbitrary, but fixed manner,

say, s= 1,2,3,...,N($\mathfrak{E}$). To each label, an experiment-wise

randomization _selection_ _probability_ is attached, p(s).

These probabilities guide the process of selecting the

one experiment-wise randomization used for the actual

experiment. The process is

(a) select with uniform probability R, 0 < R < 1;
(b) find the value of 1 ≤ S ≤ N($\mathfrak{E}$) satisfying

$$\sum_{s=0}^{S-1} p(s) \leq R < \sum_{s=0}^{S} p(s), \text{ with } p(0) = 0; \text{ then}$$

(c) select the plan with label s=S.   (3.2.7)

[Kempthorne and Doerfler, 1969, p. 232, as corrected.]

Because R is uniform (0,1), the probability of R falling into any interval is the length of the interval, p(s). Experimental randomizations with larger values of p(s) have greater probabilities of capturing R. The selected label, s, then directs the actual assignment of experimental units to treatments. The mass draw process attaches one experiment-wise randomization selection probability, p(s), to each of the N($\mathcal{E}$) experiment-wise randomizations. In contrast, the sequential selection process attaches t treatment assignment probabilities, p(ij1:k), to each of the bpr experimental units.

3.2.5 Assignment and Selection Probabilities

Typically, the literature uses "at random" to mean equal probability for each of the p(ij:k) assignment and p(s) selection probabilities, that is,

$$p(s) = 1/N(\mathcal{E}) \quad \text{and} \tag{3.2.8}$$

$$p(ij:k) = 1/t . \tag{3.2.9}$$

Because equal probability is assigned to p(s), p(s)=1/N with N = N($\mathcal{E}$) or t as indicated in (3.2.8) and (3.2.9).

We now prove that equality of all p(s) implies the equality of all p(ij1:k), that is (3.2.8) => (3.2.9). As examples, we use (3.2.2), (p. 107), (3.2.4), (p. 108), and the below (3.2.12), (p. 115), (3.2.13), (p. 116). Note that (3.2.6), (p. 109), is the special case of r=1 of (3.2.4) and for r=1, (3.2.2), (p. 102), has N($\mathcal{G}$)=1.

The proof uses <u>inclusion probabilities</u> [Cassel,

Särndal, Wretman, 1977, p. 11-12 and 68]. For an

arbitrary block, define $\mathfrak{C}_{j1:k}$ as the set of all block-

wise randomizations which assign experimental unit j1 to

treatment k. The number of such randomizations is easily

found with the two-stage process of Section 3.2.3. At

the first stage, the number of groupings is $N(\mathfrak{G})$, as in

(3.2.2) or the below (3.2.12). At the second stage, the

set containing experimental unit j1 is assigned to

treatment k. The remaining (t-1) sets may be assigned in

(t-1)! ways to the remaining t-1 treatments. This

process is repeated $N(\mathfrak{G})$ times, once per element of $\mathfrak{G}$,

giving as the number of elements in $\mathfrak{C}_{j1:k}$,

$$N(\mathfrak{C}_{j1:k}) = \{N(\mathfrak{G})\}\{(t-1)!\}. \tag{3.2.10}$$

For example, using $N(\mathfrak{G})$ of (3.2.2) and (3.2.12),

$$N(\mathfrak{C}_{j1:k}) = \begin{cases} \{(rt)!/[(r!)^t t!]\}\{(t-1)!\} & \text{for } (3.2.2) \\ \{(t!)^r / t!\} \{(t-1)!\} & \text{for } (3.2.12) \end{cases}$$

$$= \begin{cases} (rt)! / [(r!)^t t] & \text{for } (3.2.2) \\ (t!)^r / t & \text{for } (3.2.12) \end{cases}$$

There are $N(\mathfrak{B})$ block-wise randomizations for each of the

remaining b-1 blocks, giving

$$N(\mathfrak{C}_{ij1:k}) = N(\mathfrak{C}_{j1:k})[N(\mathfrak{B})]^{b-1}. \tag{3.2.11}$$

The first order inclusion probability is the probability

of assigning U_{ij1} to treatment k, which equals $p(ij1:k)$.

By definition, this is the total of all selection

probabilities summed over only those elements of $\mathcal{E}$ which assign experimental unit ijl to treatment k.

Theorem 3.2.1

a) All $p(s)$ are equal $<=>$ $p(s) = 1/N(\mathcal{E})$.

b) All $p(ijl:k)$ equal $<=>$ $p(ijl:k) = 1/t$.

c) All $p(s)$ are equal $=>$ all $p(ijl:k)$ are equal.

Proof:

Parts (a) and (b) follow directly from the definition of a probability set function. For (c),

$$p(ijl:k) \equiv \sum_{\mathcal{E}_{ijl:k}} p(s) = \sum_{\mathcal{E}_{ijl:k}} [N(\mathcal{E})]^{-1} = \sum_{\mathcal{E}_{ijl:k}} [N(\mathcal{B})]^{-b}$$

$$= \{N(\mathcal{E}_{ijl:k})\} \times \{[N(\mathcal{G})^{-1}(t!)^{-1}[N(\mathcal{B})]^{-(b-1)}\}$$

$$= \{N(\mathcal{E}_{jl:k})[N(\mathcal{B})]^{b-1}\}$$

$$\times \{[N(\mathcal{G})^{-1}(t!)^{-1}[N(\mathcal{B})]^{-(b-1)}\}$$

$$= \{N(\mathcal{G})[(t-1)!][N(\mathcal{B})]^{b-1}\}$$

$$\times \{[N(\mathcal{G})^{-1}(t!)^{-1}[N(\mathcal{B})]^{-(b-1)}\}$$

$$= [(t-1)!]/[t!]$$

$$= 1/t .$$

For the $N(\mathcal{E})$ of (3.2.4) and (3.2.13) the proof is

$$p(ijl:k) = \sum_{\mathcal{E}_{ijl:k}} p(s) = \sum_{\mathcal{E}_{ijl:k}} [N(\mathcal{E})]^{-1} = \sum_{\mathcal{E}_{ijl:k}} [N(\mathcal{B})]^{-b}$$

$$
= \begin{cases} \sum_{\mathbb{C}_{ij1:k}} \dfrac{(r!)^t}{(rt)!} \, [1/N(\mathbb{B})]^{b-1} & \text{for} \quad (3.2.4) \\[2ex] \sum_{\mathbb{C}_{ij1:k}} \dfrac{1}{(t!)^r} \, [1/N(\mathbb{B})]^{b-1} & \text{for} \quad (3.2.13) \end{cases}
$$

$$
= \begin{cases} \{N(\mathbb{C}_{ij1:k})\} \left\{ \dfrac{(r!)^t}{(rt)!} \, [1/N(\mathbb{B})]^{b-1} \right\} & \text{for} \quad (3.2.4) \\[2ex] \{N(\mathbb{C}_{ij1:k})\} \left\{ \dfrac{1}{(t!)^r} \, [1/N(\mathbb{B})]^{b-1} \right\} & \text{for} \quad (3.2.13) \end{cases}
$$

$$
= \begin{cases} \left[\dfrac{(rt)!}{[(r!)^t t]} [N(\mathbb{B})]^{b-1} \right] \left[\dfrac{(r!)^t}{(rt)!} \left\{ \dfrac{1}{N(\mathbb{B})} \right\}^{b-1} \right] & \text{for} \;(3.2.4) \\[2ex] \left[\dfrac{(t!)^r}{t} [N(\mathbb{B})]^{b-1} \right] \left[\dfrac{1}{(t!)^r} \left\{ \dfrac{1}{N(\mathbb{B})} \right\}^{b-1} \right] & \text{for} \;(3.2.13) \end{cases}
$$

$$
= 1/t \ .
$$

The second order inclusion probability, within a block, is defined as the probability that experimental units U_{j1} and $U_{j'1'}$, are both assigned to treatment k. The inclusion probability is the summation of $p(s)$ over the set $\mathbb{C}_{j1,j'1':k}$ defined similarly to $\mathbb{C}_{j1:k}$. This probability is prior to the creation of the sets $\mathbb{G}$. Afterward, the second order probability is 0 if 1 and 1' are in different sets j and j'. Two such experimental units will not be assigned to the same treatment k. If 1 and 1' are in the same j set, the second order inclusion probability is the probability of this set being assigned to treatment k. Higher order inclusion probabilities are defined likewise up to and including order r, after which all are zero. No more than r experimental units may be assigned to treatment k in any one block. A conjecture

is that the implication reverse to Theorem 3.2.1(c)
requires equality for all inclusion probabilities for
orders from 1 through r.

Bailey [1987, p. 712], discussing an additive design
(see Section 3.3.7 below) without a covariate, defines a
randomization process as

a) <u>unbiased</u> if $E_R[\hat{\tau}_k - \hat{\tau}_{k'}] = \tau_k - \tau_{k'}$, $k \neq k$ and as

b) <u>valid</u> if E_R[Treatment Mean Square]

$= E_R$[Error Mean Square] when $H_{o:\mu,b,\tau}\tau=0$.

A process is unbiased if and only if all $p(ijl:k)$ are
equal. Thus, by Theorem 3.2.1, biased additive designs
lacking a covariate must have unequal selection
probabilities, $p(s)$. "For most common designs validity
can be characterized in terms of [second order inclusion
probabilities]" [ibid. p. 712]. See Bailey [1987] for
references to earlier discussions of these definitions.
Often "unbiased" is used for (b).

Two processes, a and b, whose sets of $p(s)$ are
equal, that is $p_a(s) = p_b(s)$ for each s, are <u>equivalent
selection</u> <u>processes</u>. The two-stage selection process of
Section 3.2.3, (p. 105), with equal probability of
selection at each stage is equivalent to the mass draw
selection process with $p(s)=1/N(\mathcal{E})$.

3.2.6 Restricted Randomization

Randomization may be constrained or restricted by
adding a condition to the model, or equivalently,

assigning a zero value to some selection probabilities.
For example, in each block one could group the rt
experimental units to balance the range of the covariate
"initial height" by assigning one experimental unit of
the tallest t to each treatment, then one experimental
unit of the next tallest t to each treatment and so on,
finally assigning one experimental unit of the shortest t
to each treatment. This adds a condition to the model by
restricting the membership of each of the t subgroups of
$\mathfrak{G}$. Alternatively, one could assign zero values of $p(s)$
to selected groupings of $\mathfrak{G}$. For example, either tactic
could eliminate (d) from the list in Section 3.2.3.

We derive $N(\mathfrak{G})$ for such a condition on the model.
Taking the tallest t experimental units, assign them one
at a time to the t subgroups. There are t ways of
assigning the first experimental unit, t-1 ways of
assigning the second and so on. This gives a total of
t! ways of assigning the tallest t experimental units.
Repeat this process with the next tallest t experimental
units, again obtaining t! ways of assigning them.
Continuing for r such groups gives $(t!)^r$ initial lists.
The t subgroups can be arranged in t! ways, making the
number of distinguishable groups,

$$N(\mathfrak{G}) = (t!)^r/t! \ . \tag{3.2.12}$$

The number of possible experiment-wise randomizations is

$$N(\mathbb{E}) = \{N(\mathbb{B})\}^b = \{[N(\mathbb{G})][t!]\}^b$$

$$= \{[(t!)^r/t!]\ [t!]\}^b$$

$$= \{(t!)^r\}^b\ . \qquad\qquad (3.2.13)$$

With t=3 and r=2, this restriction gives $N(\mathbb{G})=(3!)^2/(3!)$ =6 and $N(\mathbb{B})=(3!)^2=36$. Without the restriction, $N(\mathbb{G})=90$, as shown below (3.2.3), (p. 108). Such a restriction helps to equalize the covariate values for each treatment within each block, improving the balance of the covariate in the design. Verrill [1993] terms this <u>predictor</u> <u>sort</u> <u>sampling</u> and discusses the characteristics of experiments so conducted. For an example, see Reisch and Webster [1969, p. 705] where $N(\mathbb{G})$ is that of (3.2.13), (p. 116).

The serpentine method of balancing ranks the experimental units by their single covariate, say from high to low, then assigns treatments 1,2,...,t in the following order (1,2,...,t), (t,t-1,t-2,...,1), (1,2,...,t), (etc.). This method so constrains $\mathbb{G}$ that

$$N(\mathbb{G})=1 \text{ for all r.} \qquad\qquad (3.2.14)$$

See Finney [1957 p. 374-375, his "objective rule of allocation"], who does not recommend it due to the unpredictable bias in estimation of experimental error it may produce. Various such balancing methods have been proposed [Finney, 1946; Lucas, 1950; Greenberg, 1953; Finney, 1957; Cox, 1957, Methods III, IV and V; Reisch and Webster, 1969, p. 705; Wu, 1981A].

3.2.7 Unequal Selection Probabilities

Instead of restricting randomization with zero-valued selection probabilities, one may assign unequal or weighted selection probabilities. For example, the values of $p(s)$ may be assigned via the concomitant variables or functions thereof. A few such weighting methods have been proposed and are further discussed in Section 4.7, (p. 320).

When the experimental unit assignment probabilities, $p(ijl:k)$, are unequal and/or dependent, it is difficult to compute the experiment-wise randomization selection probabilities, $p(s)$, from the $p(ijl:k)$. The mass draw process sidesteps this difficulty by assigning the $p(s)$ directly. The experimental unit assignment probabilities are assigned, if at all, via the first order inclusion probabilities. The remainder of Chapter 3 considers equal probabilities as in Theorem 3.2.1, (p. 112), and randomization restrained only by the model assumptions. $N(\mathcal{E})$ designates all possible randomizations; zero valued $p(s)$ are used to avoid those randomizations eliminated by, say, balancing on the covariate. Chapter 4.7, (p. 320), examines unequal, nonzero-one, assignment probabilities.

3.2.8 Why Randomize

Many reasons have been advanced for including a random component in the linking of experimental unit to treatment in designed experiments. Cox [1958, p. 74-85], Bailey [1982] and Folks [1984] provide reviews. Four of the more frequently cited benefits follow. The paper by Basu [1980] and accompanying discussion touches on each of these points. A Bayesian view is in Rubin [1978].

1. Randomization (with equal $p(s)$) provides an equal probability that any two first, second, etc. adjacent experimental units will be assigned to the same treatment or treatment pair, or triplicate etc. This is useful when errors in adjacent (physically, temporally, etc.) experimental units are of equal magnitude and correlated, as happens in field plots in agricultural experiments. As the number of experimental units increases, the "adjacent to" probabilities, and hence correlations, uniformly shrink. Under Assumption 2, that the errors are normally distributed, zero correlations implies independence of the errors, satisfying Assumption 3, that the variance-covariance matrix equals $I\sigma^2$. Thus, "the expected value of the total error for any one treatment is independent of that for any other treatment" [Anderson and Bancroft, 1952, p. 221].

The goal of the experimenter is a randomization which appears free of concentrations of adjacent

treatments. Of course, the randomization actually
selected may or may not so appear.

2.   Randomization protects one from favoring a particu-
lar treatment and from charges of having done so [Lane,
1980, p. 589]. Failure to randomize has invalidated
experiments, for example, the Lanarkshire milk trials of
1930 [Cox, 1958, p. 77-81]. Randomization is now
standard practice in medicine and other fields. Recall
the work by Thornett discussed in Chapter 2.3.1, (p. 21),
linking randomization to the independence of
(a) the assignment of treatments to experimental units
and (b) the responses of the experimental units.

     Goals One and Two conflict in that concentrations of
adjacent or nearby treatments can be detected in all but
a few randomizations. Should the selected randomization
have clear concentrations of treatments, the door is open
to charges of having favored a treatment, even if without
intention. Yet, failure to utilize the initial randomi-
zation opens the door to charges of manipulation. Yates
[1975, p. 586] claims that Fisher ducked the question of
what to do when the selected randomization has a
"systematic" arrangement of treatments. Cox [1958, p.
85-89] suggests three options, (a) add a condition to the
model statement, as in Section 3.2.6, (p. 114),
(b) ignore the suspect randomization and select another,
or (c) employ restricted randomization. For designs
without a covariate, Bailey [1985] presents unbiased and

valid, (p. 114), restricted randomization processes for designs with two or four treatments on eight experimental units. A second paper [Bailey, 1987] adds more designs and details the permutation group methods involved. Both papers specify all suspect randomizations *a priori*, and assign them zero values of p(s). As few unbiased and valid restricted randomization are available, in practice one follows Cox's suggestion (b) and discards suspect randomizations. The probability of a randomization actually being used does not, in practice, equal $[N(\xi)]^{-1}$. [Kotz and Johnson, Vol. 7, p. 524-530].

3.   Randomization provides robustness against model inadequacies [Hooke, 1958; Royal and Herson, 1973 (for survey sampling); Wu, 1981B (for linear models)]. As Godambe [1966] points out, under equal-probability sampling, the likelihood function is flat (uniform), thus uninformative. Hence, his argument for abandoning probabilistic sampling. His suggestion and manner of promoting it [Hartley and Rao, 1971; Godambe, 1975], have aroused much discussion [Godambe and Thompson, 1971]. Basu's 1980 paper continues the argument. If the model equation and assumptions are correct, the optimal design approach leads to nonrandom, purposeful, assignment of treatments to experimental units [Kiefer, 1959; Fedorov, 1969]. However, should the model be in error, purposeful assignment can lead to biased estimates.

4.   In several experimental designs, all lacking a covariate, randomization provides a basis for statistical inference without the necessity for assuming a distribution for the residuals [Kempthorne, 1975, p. 326, Section 1.15].   Kempthorne [1955, p. 949] points to the difficulty of testing the assumptions for the normally distributed random errors model of Chapter Two.   When the errors are additive (Section 3.3.7, p. 134) the randomization test (Section 3.6, p. 166,) avoids such assumptions.   Cochran and Cox [1957, p. 6-9] discuss these issues.

Assumption 3, that the variance is $I\sigma^2$, (p. 24), removes randomization Goal (1) from consideration.   Goal (3) is beyond the scope of this paper, but of importance. Section 3.7.5, (p. 216), shows that the presence of a single covariate gives some randomizations greater statistical power than others, at least for normally distributed random error.   Section 3.8.2 discusses why a covariate makes impossible the closed-form expectations helpful in Goal (4).   We feel Goal (2) is essential for confidence in and adoption of experimental findings when the errors discussed in the next section may be present.

Section 3.3

Errors and Their Implications

The algebraic identity of Section 3.3.1 introduces
two types of errors which are detailed in Sections 3.3.3
and 3.3.5. Section 3.3.2 reviews the "technical error".
An interaction error is presented in Section 3.3.4.
Section 3.3.6 connects various definitions of the error
and interaction terms. Section 3.3.7 discusses additiv-
ity. A covariate requires redefining the errors, as
shown in Section 3.3.8. Section 3.3.9 compares the
hypotheses with and without additivity. Of major
interest is (3.3.8), (p. 136), and the discussion of
(3.3.1) vs (3.3.3), (p. 138).

3.3.1 Two New Error Terms

The error term of the model equations of Chapter
Two, (2.3.1), (2.3.4) and (2.3.6), (p. 16-22), is
assumed to be due to inaccuracy of experimental
technique. It is postulated as part of the model, as
opposed to being defined explicitly. The randomization
model replaces it with one or more new error terms.
These new errors require the identifiability of the
experimental unit and, for one, the treatment assigned to
it. The label ijk now replaces the ik used in Chapter
Two. Throughout, we use r=1 and drop the subscript 1. A
superscript distinguishes between similar error terms;
the notation $(\cdot)$ indicates that any of the similar terms

applies. When the model lacks a covariate, the terms are

defined explicitly by a model equation which is an

algebraic identity, such as (3.3.1) and (3.3.6) below.

The averaging version of the dot notation simplifies such

identities. An example of this notation is

$$Y_{\cdot\cdot k} = \frac{1}{bp} \sum_{i=1}^{b} \sum_{j=1}^{p} Y_{ijk},$$

where Y_{ijk} is the unobservable true response when treat-

ment k is applied to experimental unit ij. Conceptually,

treatment k can be applied to all experimental units, and

experimental unit ij may receive in turn each treatment

k, as in $Y_{ij\cdot\cdot}$. One algebraic

identity is

$$
\begin{aligned}
Y_{ijk} &= Y_{\cdots} + (Y_{i\cdot\cdot} - Y_{\cdots}) + (Y_{\cdot\cdot k} - Y_{\cdots}) \\
&\quad + [(Y_{i\cdot k} - Y_{i\cdot\cdot}) - (Y_{\cdot\cdot k} - Y_{\cdots})] \\
&\quad + (Y_{ij\cdot} - Y_{i\cdot\cdot}) \\
&\quad + [(Y_{ijk} - Y_{ij\cdot}) - (Y_{i\cdot k} - Y_{i\cdot\cdot})], \\
&= \mu + b_i + \tau_k + (bt)_{ik} + e_{ij} + n_{ijk} . \qquad (3.3.1)
\end{aligned}
$$

The second equality supplies symbolic names for the

respective terms of the first equality. Variations of

this equation and its last three terms are the subject of

the following sections. Until Section 3.3.8, the model

equation has no covariate. Note that $\sum_{i=1}^{b} b_i = \sum_{k=1}^{t} \tau_k = 0$,

embodying Assumption 5, (p. 20).

3.3.2   Technical Error

"Technical error" is the term used by Neyman [1935, p. 110, 114 and 145], Wilk [1955, p. 70; 1957, p. 224], and Wilk and Kempthorne [1955, p. 1148] for the normally distributed random error term of Chapter Two. The last two authors also use the term "measurement errors" [p. 1148]. Neyman describes the technical error as that "due solely to the inaccuracy of experimental technique, the vegetative conditions in all our hypothetical [agricultural] experiments being exactly the same" [p. 110]. In Neyman's description, the vector of technical errors has dimension bpt x 1 instead of bt x 1 as in Chapter Two. This implies that his technical error may differ with each experimental unit-treatment combination, that is, with each reassigning of experimental units to treatments.

3.3.3   Experimental Unit Error

Kempthorne [1952, p. 145-146] defines the experimental unit error as

$$e_{ij}^K = Y_{ij.} - Y_{i..} \, , \qquad (3.3.2)$$

a term of the algebraic identity of (3.3.1). He describes it as "the deviation of the jth plot from the block mean averaging over all treatments" [p. 146]. It is "the difference between (a) the (conceptual) mean of the yields of all treatments on unit j of block i and (b) the mean over the whole of block i" [Wilk, 1955, p. 71].

Kempthorne uses the term "plot error"; Wilk uses "unit error." Roux [1982] uses this experimental unit error in a model equation containing a covariate, as discussed below in Section 3.3.8, (p. 136).

Bartlett [1935], Welch [1937], Pitman [1938] and Plackett's review of their results [1960, p.152-157], use an error term, e^P_{ij}, which is similar to e^K_{ij} in that it is not determined by the particular treatment applied to experimental unit ij. Instead of defining e^P_{ij} as a term in an algebraic identity, they postulate it as an error term, say, as "arising from the variability among the [experimental units] of a [block], errors in measurement, and other accidents affecting particular individuals" [Pitman, p. 322-323]. Pitman includes technical errors in this term; other authors [Kempthorne, Wilk, *op. cit.*] assume the technical errors to be small enough to ignore. The e^P_{ij} experimental unit error is not a conceptual deviation of the $ij^{\underline{th}}$ experimental unit as averaged over the treatments of the particular experiment; rather, it exists independently of any treatment. The e^K_{ij} of (3.3.2) may be large because the $ij^{\underline{th}}$ experimental unit has a large response (Y_{ijk}) for but one treatment. The e^P_{ij} would be large (or small) for all treatments; it is the robustness (or lack of) of experimental unit ij to all treatments.

Kempthorne [1952, p. 137] uses the e_{ij}^K of (3.3.2),

but with the model equation

$$Y_{ijk} = Y_{...} + (Y_{i..} - Y_{...}) + (Y_{ijk} - Y_{ij.})$$
$$+ (Y_{ij.} - Y_{i..})$$
$$= \mu + b_i + \tau_k^K + e_{ij}^K . \qquad (3.3.3)$$

In this model, the treatment effect, τ_k^K, is assumed to be

identical for each experimental unit. In (3.3.1),

(p. 123), τ_k is the treatment effect as averaged over all

experimental units. When the e_{ij}^K are considered as

random variables, forcing them to sum to zero within each

block imparts a correlation among them within each block.

The variance-covariance matrix is not $\mathbb{I}\sigma^2$ as in

Assumption 3, (p. 24), but, rather is $\Psi\sigma^2$ as in

Assumption 3^*, (p. 16).

3.3.4 Block-Treatment Interaction

Kempthorne [1952, p. 145-146] defines

$$(bt)_{ik} = (Y_{i \cdot k} - Y_{i..}) - (Y_{..k} - Y_{...}) \qquad (3.3.4)$$

as "the deviation of the $k^{\underline{th}}$ treatment from the overall

mean, averaging over all plots." A term of the identity

(3.3.1), (p. 123), it is "the difference between (a) the

effect of treatment k on block i and (b) the treatment

effect." Thus, "it is a measure of the extent to which

treatment k and block i interact" [Wilk, 1955, p.71]. It

is not a random variable unless the blocking factors

and/or their levels are a random sample from a larger

population.   Herein, as in Chapter Two, we consider them
fixed.   By Assumption 7, (p. 21), this term is either
zero or included in the normally distributed random
error.   Roux [1984] applies the multiplicative approach
of Mandel [1961] to block-unit interaction.

3.3.5 Experimental Unit-Treatment Interaction

The classical example of experimental unit-treatment
interaction is penicillin allergy [Kempthorne, 1975, p.
324].   Should the penicillin treatment be assigned to
an allergic experimental unit, the response is quite
different than if the penicillin had been assigned to a
nonallergic experimental unit.   The conclusion of an
experiment can be determined by a sufficient number of
such unfortunate assignments.

Discussing the response for treatment k in block i,
Neyman [1935, p. 111 and 145] defines the "soil error" as

$$\tau_{ijk}^{N1} = Y_{ijk} - Y_{i \cdot k} \cdot \qquad (3.3.5)$$

It is "the correction for fertility variation within the
block" [*ibid.* p. 111].   Neyman's definition is in terms
of the algebraic identity

$$Y_{ijk} = Y_{\cdot \cdot k} + (Y_{i \cdot k} - Y_{\cdot \cdot k}) + (Y_{ijk} - Y_{i \cdot k}) \cdot$$

The addition of terms for the mean and block requires
redefining this interaction term as follows

$$Y_{ijk} = Y_{\cdot \cdot \cdot} + (Y_{i \cdot \cdot} - Y_{\cdot \cdot \cdot}) + (Y_{\cdot \cdot k} - Y_{\cdot \cdot \cdot})$$

$$+ [Y_{ijk} + (Y_{\cdot \cdot \cdot} - Y_{i \cdot \cdot} - Y_{\cdot \cdot k})]$$

$$= \mu + b_i + \tau_k + \tau_{ijk}^{N2} \cdot \qquad (3.3.6)$$

This experimental unit-treatment interaction term
contrasts (a) the response of experimental unit ij to
treatment k with (b) the overall mean as adjusted for
(i) the response of block i averaged over all treatments
and experimental units and for (ii) the response of
treatment k averaged over all experimental units. Its
functional form is similar to the usual estimate for
factor A by factor B interaction.

Kempthorne [1952, p. 145-146] defines a different
experimental unit-treatment interaction,

$$\eta^K_{ijk} = (Y_{ijk} - Y_{ij\cdot}) - (Y_{i\cdot k} - Y_{i\cdot\cdot}). \qquad (3.3.7)$$

It is a term of the algebraic identity of (3.3.1) and is
"the deviation of the effect of the $k^{\underline{th}}$ treatment on the
$j^{\underline{th}}$ plot of the $i^{\underline{th}}$ block from the average effect in the
$i^{\underline{th}}$ block" [p. 146]. Wilk [1955, p. 72] describes it as
"the difference between (a) the effect of treatment k on
unit j of block i and (b) the effect of treatment k over
all of block i. It is a measure of the extent to which
treatment k and unit j of block i interact."

Roux [1984] defines an experimental unit-treatment
interaction for the completely randomized design. It is
identical to (3.3.7), but lacks the subscript i.

As with e^P_{ij}, one may postulate an experimental unit-
treatment interaction, η^P_{ijk}. The following graph
illustrates this interaction error term.
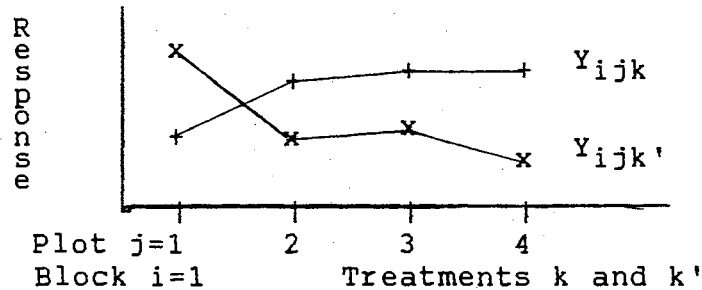
Graph 3.3.1

Experimental Unit - Treatment Interaction π_{ijk}

The graph is that of a factor A by factor B interaction
as in the normally distributed random error model, except
here the second "factor" is the experimental unit. Of
the experimental unit-treatment interaction terms, π_{ijk}^{N2},
most closely resembles the estimate of the A by B
interaction. We use it to indicate the meaning of the
postulated version of the experimental unit-treatment
interaction term, π_{ijk}^{P}.

Suppose several experimental units receive in turn
treatments k and k' as in Graph 3.3.2.

τ_k □   □   ▽*

τ_k .□*....................   ....τ_k

            □*   ▽*
                           □   □*

$\tau_{k'}$.▽..........▽*........   ...▽...□.............$\tau_{k'}$

      ▽*   □*   ▽

block i   1   1   2   2 1   1   2   2

plot j   1   2   1   2 3   4   3   4

        □ = $\tau_k + \eta_{ijk}$ ▽ = $\tau_{k'} + \eta_{ijk'}$ $k \neq k'$
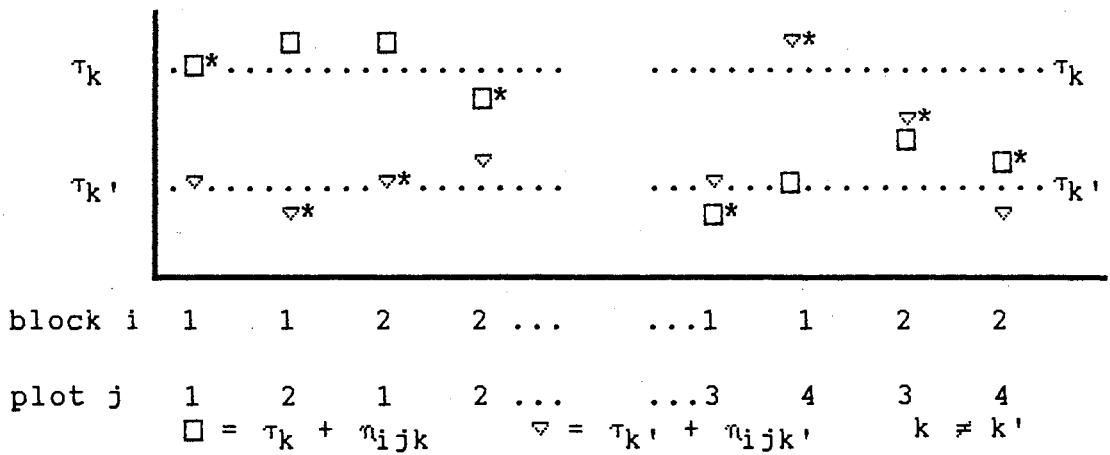
Graph 3.3.2

Error as Experimental Unit-Treatment Interaction

The experiment proceeds under the assumption that the experimental unit-treatment errors are small relative to the treatment effects, as shown in the left-hand side of Graph 3.2 for plots one and two. The experimental unit i=1, j=1 has zero interaction under either treatment k or k'. Other experimental units respond to treatment k (and/or k') but over or under shoot the "true" effect of treatment k (and/or k'). For these experimental units, any assignment of treatments to experimental units, such as the points with an asterisk (*), will lead to the correct conclusion that $\tau_k > \tau_{k'}$. In experimental units j = 3 or 4 to the right-hand side of Graph 3.2, the plot errors are large relative to the treatment effects. The conclusion is determined by the assignment made. The assignment shown leads to the incorrect conclusion

$\tau_k < \tau_{k'}$. The difference between points assigned to the same experimental unit is:

$$Y_{ijk} - Y_{ijk'} = (\tau_k - \tau_{k'}) + (\eta_{ijk} - \eta_{ijk'}),$$

and is a fixed shift due to treatment differences which is constant for all experimental units, plus a shift due to interaction which may differ with each experimental unit.

Each of the experimental unit-treatment interaction error terms is a random variable, dependent upon the plot j to which treatment k is applied. Each is independent of the technical error ε and of the $\eta_{ijk}^{(\cdot)}$ of a different block [Neyman, 1935, p. 146]. Two $\eta_{ijk}^{(\cdot)}$ in the same block are not statistically independent, as the following joint probability of their indices is not factorable,

$$p(IJK=ijk, \ IJ'K'=ij'k') = \begin{cases} 0 & \text{for } j=j' \text{ or } k=k' \text{ or both} \\ t^{-1}[t-1]^{-1} & \text{otherwise.} \end{cases}$$

When the entire set of η_{ijk} are considered as random variables, the entire set of bpt η_{ijk} are correlated, but the bp η_{ijk} appearing in any one randomization appear not to be. They may have a variance-covariance matrix of $\mathbf{I}\sigma^2$, as in Assumption 3, (p. 24).

3.3.6 Relationships Among Error Terms

The various error terms are related as follows

$$e_{ij}^K + \eta_{ijk}^K = \eta_{ijk}^{N1} \quad \text{and}$$

$$(bt)_{ik}^K + e_{ij}^K + \eta_{ijk}^K = \eta_{ijk}^{N2} \ . \tag{3.3.8}$$

When algebraic identities define the error terms, the model must either (a) use one of the Neyman definitions, $\eta_{ijk}^{N(\cdot)}$, without an experimental unit error term or (b) Kempthorne's definition η_{ijk}^K and include the experimental unit error term, e_{ij}^K. When $(bt)_{ik}^K$ is zero, the definitions of η_{ijk}^{N1} and η_{ijk}^{N2}, (3.3.5), (p. 127), and (3.3.6), (p. 132), are identical, as seen in (3.3.8).

The error term(s) may be postulated instead of defined in terms of an algebraic identity. In this case the model equation will include e_{ij}^P and/or η_{ijk}^P, instead of those of (3.3.8). The interaction term, η_{ijk}^P, indicates the response by experimental unit ij over and above (a) that due to its robustness relative to other experimental units, e_{ij}^P and (b) that due to the treatment effect, τ_k, considered as averaged over all experimental units.

When for several experimental units the effect of e_{ij}^P and/or η_{ijk}^P is large relative to the effect of τ_k, the effects of τ_k will be masked. Comparing subsets of the experimental units, say with the bootstrap procedure, may permit detection of treatment effects by removing

experimental units with large e_{ij}^P and/or η_{ijk}^P. The model would no longer have the same number of experimental units per treatment per block. Such methods are not considered here.

Wilk [1955, p.72] notes that $\tau_k = 0$ for all k does not imply $(bt)_{ik} = 0$ for all k, nor does $(bt)_{ik} = 0$ imply $\eta_{ijk}^K = 0$ for all i and k.

Wilk suggests that if blocking has successfully grouped experimental units that are homogeneous as to their responses to the treatments, then a zero block-treatment interaction effect would imply a zero experimental unit-treatment interaction. The converse does not hold; small experimental unit-treatment interaction effects can coexist with large block-treatment interaction effects. An example is animal experimental units blocked by parents. Litter mates may differ little in their response to any treatment k, making $\eta_{ijk}^{(\cdot)}$ small, but litter-to-litter differences may be substantial, making $(bt)_{ik}$ large [Kempthorne, 1955, p. 952; Roux, 1984, p. 144]. The consistency of transmission of relevant genes to offspring, say by inbreeding, determines the success of such blocking. Should one treatment have an unusual and equal effect on one experimental unit of each block, then $\eta_{ijk}^{(\cdot)}$ would be large for that treatment, but $(bt)_{ik}$ would be small.

3.3.7 Additivity

Model equations lacking interaction are said to be additive or to have mutual additivity [Wilk and Kempthorne, 1955, p. 1149; Cox, 1958, p. 14-17]. For such model equations,

$$Y_{ijk} = \mu + b_i + \tau_k + e_{ij} \text{ , or} \qquad (3.3.9)$$

$$Y_{ijk} = \mu + b_i + \tau_k + e_{ij} + \varepsilon \text{ .} \qquad (3.3.10)$$

Those model equations including interaction terms are non-additive. The interactions may be those of Section 3.3.4 [$(bt)_{ik}$] and/or Section 3.3.5 [η^p_{ijk}] and/or interaction between two factors. Additivity in the strict sense excludes all interactions, but permits an experimental unit error, $e^{(\cdot)}_{ij}$, in the model, as in (3.3.9). Additivity in the broad sense also excludes all interaction, but permits an experimental unit error, $e^{(\cdot)}_{ij}$, and a normally distributed random error, ε, as in (3.3.10) [Kempthorne, 1955, p. 952; W. J. Welch, 1990, p. 697]. Under broad additivity, it is the expected value of the difference between the responses of the same experimental unit subjected to two different treatments, k and k', which equals $\tau_k - \tau_{k'}$. The expectation is with respect to the normal distribution,

$$E_N(Y_{ijk} - Y_{ijk'}) = \tau_k - \tau_{k'} .$$

Equation (3.3.11) summarizes these types of additivity,

$$
Y_{ijk} - Y_{ijk'} = \begin{cases} \tau_k - \tau_{k'} & \text{for each } ij & \begin{array}{l}\text{additive}\\\text{strict sense}\end{array} \\[2ex] \tau_k - \tau_{k'} & \text{expected value } (E_n) & \begin{array}{l}\text{additive}\\\text{broad sense}\end{array} \\[2ex] \tau_k - \tau_{k'} + [(bt)_{ik} - (bt)_{ik'}] & & \text{non-additive.} \\ \quad + [\eta_{ijk} - \eta_{ijk'}] & & \end{cases}
$$

(3.3.11)

Treatments may be additive within a block, implying that all $\eta_{ijk}^{(\cdot)}=0$, but non-additive from block to block, implying that $(bt)_{ik} \neq 0$ [Wilk, 1955, p. 72]. Model equations with the former, all $\eta_{ijk}^{(\cdot)}=0$, have unit treatment additivity [Wilk and Kempthorne, 1955, p. 1150], also termed treatment-unit additivity [Wilk, 1955, p. 72].

A nonrandom covariate with constant coefficient β does not affect the additivity, or lack thereof, of the model equation, as the two treatments, k and k', are applied to the same experimental unit, whose covariate value is unchanged. If the coefficient changes with the treatment, β_k, then the model equation is non-additive.

Under additivity ($bt_{ik} = \eta_{ijk} = 0$) and without a covariate, Kempthorne points out [1955, page 956], that

> Randomization affects only the [experimen-
> tal unit] error, this statement indicates the
> fields of application where randomization is
> important. One can envision experimental
> situations where [experimental unit] errors
> are trivial and the additional errors large in

comparison to them. In such a situation the lack

of randomization will not seriously invalidate

the experimental conclusions. This seems to be

to be the main reason why the randomization

experiment has not been essential to progress

in some of the physical and chemical sciences.

The inanimate "subjects" of the physical scientists are

homogeneous experimental units which react identically to

identical changes in treatments. They face no interact-

tion between the experimental unit i,j and the treatment

k; the interaction η_{ijk} equals zero for all randomiza-

tions. Researchers whose subjects are plants, animals,

or humans face experimental units closer to the two

examples of Graph 3.3.2, (p. 130), hopefully closer to

those of the left-hand side. Ideally, the η_{ijk} and e_{ij}

are small enough so that only a small percentage of the

randomizations will provide incorrect conclusions.

3.3.8   Errors in the Covariate Model

The presence of a covariate prevents defining the

terms of the model equation by an algebraic identity such

as (3.3.1), (p. 123), or (3.3.3), (p. 124). One way

around this is to define

$$W_{ijk} = Y_{ijk} - \beta Z_{ij} \ . \qquad\qquad (3.3.12)$$

One then uses the W_{ijk} in place of the Y_{ijk} in (3.3.1) or

(3.3.3). This method postulates a known covariate slope

coefficient, β.

A covariate with a constant slope coefficient β is like experimental unit error in that it is attached to the experimental unit and is unaltered by the randomization. While the bp experimental unit errors are entirely unknown, the bp covariates are known, up to the unknown multiplicative constant β. When the slope coefficient changes with each treatment, β_k, the product $\beta_k Z_{ij}$ has some aspects of the experimental unit-treatment interaction error. Like this error, the product changes with the treatment as β_k changes to $\beta_{k'}$. Unlike this error, this product is known up to the multiplicative constants β_k, $\beta_{k'}$, etc. Also, the change from β_k to $\beta_{k'}$ is independent of the experimental unit, thus the ratio $(\beta_k Z_{ij})/(\beta_{k'} Z_{ij})$ is constant for each covariate Z_{ij}.

Roux [1982, p. 4] defines β as the term minimizing the variance of $\hat{\tau}_k$, when $\hat{\tau}_k$ is defined as in (2.4.29), (p. 32). His model equation is (3.3.3), (p. 124), with the covariate entering at the estimation stage. The definition of β is equal to the usual least squares estimator from the model equation

$$e_{ij} = \beta(Z_{ij} - Z_{i.}) + d_{ij} \ , \qquad (3.3.13)$$

where the d_{ij} represents deviation from regression. The resulting minimum variance of the estimated value of τ is

$$\mathrm{var}(\hat{\tau}_k) = \sum_{i=1}^{b} \sum_{j=1}^{p} (d_{ij})^2/[b^2 t] \ . \qquad (3.3.14)$$

More common is to postulate the model equation, as in
Kempthorne [1952, p. 159]. A price is that "we are
entirely dependent on the accuracy of the model . . . for
the test of the hypotheses that there are no treatment
effects." [*ibid.*, p. 159].

3.3.9 Two Hypotheses: In-Particular and On-Average

The hypotheses of interest in Chapter Two, (2.6.40),
(p. 63), and (2.6.44), (p. 64), test at least one
treatment being nonzero for the response of *every*
experimental unit. This is the in-particular version of
the hypotheses. In the randomization model, when there
is experimental unit error, with or without normally
distributed random error, again the hypotheses test a
nonzero treatment effect for *every* experimental unit.

When there is experimental unit-treatment interac-
tion, η_{ijk}, some experimental units may be positively
affected, others negatively affected, and yet others
unaffected, by one or more of the t treatments.
Nevertheless, the treatment effect, averaged over all
experimental units may be nonzero. In the randomization
model with experimental unit-treatment interaction, the
two hypotheses test at least one treatment being nonzero
for the responses as *averaged* over all experimental units
receiving that treatment. This is the on-average version
of the hypotheses. The truth of the in-particular hypoth-
esis implies the truth of the on-average hypothesis, but
the converse does not hold.

The randomization model equation of (3.3.3) lacks experimental unit treatment and block treatment interactions, but includes experimental unit error. This model equation tests the in-particular hypothesis, as the τ_k^K of (3.3.3), (p. 124), suggests. The of τ_k of (3.3.1), (p. 123), shows it to be averaged over all experimental units. It tests the on-average hypothesis.

The in-particular hypothesis interests the physician and his/her patient. They need to know if the treatment will work in their particular case. The on-average hypotheses interests the public health authorities and the society supporting them. They need to know if, on average, the society will benefit from the treatment, even if some are injured by it. The in-particular version interests the chemist verifying a mechanistic explanation for a physical phenomenon, such as generating oxygen and hydrogen by electrolysis. The on-average version interests the entomologist facing a diverse and evolving population, such as in mosquito control. When there is experimental unit-treatment interaction and/or block-treatment interaction, the standard analysis of variance of Chapter Two (with or without a covariate) gives the appearance of testing the in-particular hypothesis, when in fact, it only tests the on-average hypothesis. See Neyman [1935, p. 111, below (3) and 172-177], Welch [1937, p. 22-23] and Kempthorne [1952, p. 132-134; 1975, p. 324-325, Section 1.12] for discussions of these two hypotheses.

<u>Section 3.4</u>

The Randomization Model

Sections 3.4.1 and 3.4.2 present the randomization

model equation, assumptions and probability space.

Section 3.4.3 describes the parameter space and shows

that interest centers in statistics of these parameters.

Section 3.4.4 contrasts assumptions underlying the

randomization and normally distributed random errors

models. Connections between survey sampling and

randomization model concepts are discussed in sections

3.4.3 and 3.4.4. The major result is the probability

space of Section 3.4.2, (p. 142).

3.4.1 The Model Equation and Assumptions

The Randomization model equation is identical to

(2.3.4), (p. 19), or (2.3.6), (p. 22), except the error

term is e_{ij} and/or n_{ijk} in their postulated versions, as

is necessary with a covariate. For convenience we drop

the superscript P. As r=1, we drop the subscript 1.

Assumptions 1 and 3 - 8 of the normally distributed

random errors model hold. We add <u>Assumption 9</u>,

$$\sum_{j=1}^{p} e_{ij} \ = \ \sum_{j=1}^{p} n_{ijk} \ = \ \sum_{k=1}^{t} n_{ijk} = 0, \qquad (3.4.1)$$

for all blocks i=1,2,...,b.

In the unusual cases where these terms may be estimated,

discussed below in Section 3.9, (p. 232), their

estimators are similarly constrained.

3.4.2 The Probability Space for the Randomization Model

We follow Cassel, Särndal and Wretman [1977,p. 1-31]
in defining the probability space for the randomization
model.   They use ($\mathcal{U}$, $\mathcal{A}$, $p(s)$) to denote the probability
space as developed for survey sampling, instead of the
more common notation (Ω, $\mathcal{F}$, $\mathbb{P}$).   We adapt their notation
for the randomization model.

The randomization model has a finite population of
bpt block-plot-treatment labels, ijk, which identify
the responses Y_{ijk}.   Designate one set of all bpt of the
labels as $\mathcal{U}_u$ = {111, 112, ..., 11t, 121, ..., 112t,

..., ijk, ..., bpt},

where u = 1, 2, ..., bp.

The $\mathcal{U}_u$ are identical for each value of u. Any experiment-
wise randomization draws one label from $\mathcal{U}_u$ for each of
its bp experimental units.   That is, there are bp draws
of $\mathcal{U}_u$.

Assume that each draw is independent from all others. This gives as $\mathfrak{U}$ the Cartesian product

$$\mathfrak{U} = \mathfrak{U}_1 \times \mathfrak{U}_2 \times \cdots \times \mathfrak{U}_{bp}.$$

The set $\mathfrak{U}$ has bpt^{bp} elements. For two blocks and one experimental unit for each of four treatments, $bp=(4)(2)$ $=8$, and typical elements, ψ, of $\mathfrak{U}$ are

(111, 122, 133, 144, 211, 222, 233, 244), or

(111, 121, 131, 141, 211, 211, 211, 211), or

(111, 112, 113, 114, 111, 112, 113, 114).

Each row, not each triplet, is one element of $\mathfrak{U}$. Elements (entire rows), such as the last two, are useless or impossible experiment-wise randomizations. In the second row only one treatment is used; in the third row the same experimental unit is reused. Such randomizations will be assigned zero probability. Only elements such as the first are possible experiment-wise randomizations. This Cartesian product representation is due to Dr. Sahadeb Sarkar with help from Dr. Ignacy Kotlarski and the author.

Define $\mathcal{A}$ as the set of all sets of $\mathfrak{U}$, that is, the power set of $\mathfrak{U}$, plus the null set.

The probability density or mass (set) function, $p(\cdot)$, is defined on the sets of $\mathcal{A}$. Only those sets permitted by the experimental design, such as the $N(\mathcal{E})$ sets of (3.2.4), (p. 108), or (3.2.13), (p. 116), have non-zero selection probability. Label these sets $s=1, 2,$ $\ldots, N(\mathcal{E})$, $N(\mathcal{E})+1, \ldots, N(\mathcal{A})$. In these experimental designs all sets of $\mathcal{A}$ with more than one element of $\mathfrak{U}$ are

assigned zero selection probability, that is, the

experiment is performed but once. In designs which reuse

experimental units, such as crossover designs, only sets

of $\mathcal{Q}$ with two elements of $\mathcal{U}$ may be assigned nonzero

selection probability. All other sets, the vast majority

of the sets in $\mathcal{Q}$, are assigned zero selection probability.

Denote the summation over all elements of $\mathcal{Q}$ as $\sum_{\mathcal{Q}}(\cdot)$.
For $r=1$ and $p=t$, one such $p(\cdot)$ is

$$p(s) = \begin{cases} (p!)^{-b} & \text{for the } N(\mathcal{E})=(p!)^b \text{ elements of } \mathcal{E} \subset \mathcal{Q} \\ 0 & \text{otherwise,} \end{cases}$$

with $\sum_{\mathcal{Q}} p(s) = 1$.

For the randomization model, the probability space is the

triple

$$(\mathcal{U}, \ \mathcal{Q}, \ p(s)). \tag{3.4.2}$$

In the following we consider $s \ni \{1,2, \ldots, N(\mathcal{E})\}$, it

being understood that $p(s)=0$ for the remaining sets of $\mathcal{Q}$.

The random vector of length bp, $\Upsilon(\psi)$, is a set

function which assigns each element, ψ, of $\mathcal{U}$ to a point

in $\mathbb{R}^{bp}$, a bp-dimensional Euclidean space,

$$\Upsilon(\psi) : \mathcal{U} \rightarrow \mathbb{R}^{bp} . \tag{3.4.3}$$

The sets, s, and their probabilities, $p(s)$, define

the experimental design, just as they define the sampling

design in survey sampling. In the terminology of survey

sampling, the selection probability $p(s)$ is the unordered

<u>sample design</u> [Cassel, Särndal and Wretman, 1977, p.10].

3.4.3 Auxiliary Variables and The Data

To each of the elements of any $\mathcal{U}_u$, say ijk, is associated a known <u>concomitant</u> or <u>auxiliary variable</u>(s), such as the covariate Z_{ij}. For each block-plot ij the covariate Z_{ij} is identical for each treatment and each is unaltered by the other.

There are bpt possible responses Y_{ijk}. The vector $\tilde{Y}$ of all of them is the parameter of the finite population. That is, $\tilde{Y}' = (Y_{111}, Y_{112}, \ldots, Y_{11t}, Y_{121}, \ldots, Y_{12t},$

$$\ldots, Y_{ijk}, \ldots, Y_{bpt}).$$

The parameter space is $\mathbb{R}^{bpt}$, a bpt-dimensional Euclidean space; $\tilde{Y} \in \mathbb{R}^{bpt}$.

The data is the random vector Y, of length bp, taking observed values

$$y_{ijk} = \{(ijk, Y_{ijk}), \; ijk \in s \subset \mathcal{U}, \; Y_{ijk} \in \tilde{Y} \in \mathbb{R}^{bpt}\}.$$

A statistic is a function U of the data Y, such as $U(Y) = F(\tau|\mu, b, \wp)$. In the randomization model, $U(Y)$ depends upon the parameter $\tilde{Y}$ only through (a) those Y_{ijk} which have been selected by randomization s, and possibly through (b) known auxiliary information provided by the labels, ijk. The covariates, Z_{ij}'s, are such auxiliary information. Recalling that $\tilde{Y}$ is the vector of parameters, we see that no unknown parameters enter into $U(Y)$. This definition of a statistic is the same as in the

normally distributed random errors model where such statistics as the F-ratio under the null hypothesis depend upon only known parameters, in this case the numerator and denominator degrees of freedom.

Interest in the randomization model is in such statistics, not the entire set of parameters $\tilde{Y}$. The typical statistics of interest (F-ratios) ignore the plot labels in the parameter space $\mathbb{R}^{bpt}$ and replace the plot labels in the data space $\mathbb{R}^{bp}$ with the treatment labels. This treats both spaces as $\mathbb{R}^{bt}$.

The researcher hopes the plots are sufficiently alike (small e_{ij}) and sufficiently unaffected by the treatments (small π_{ijk}) to permit the lower dimensional space to represent the higher dimensional space. The parameter space needs to be "thin" in the plot dimension, for the statistics to be valid for all randomization.

Graph 3.4.1a shows small values of e_{ij} and π_{ijk}. The effect of treatment two reveals itself as positive for any randomization. In Graph 3.4.2a the values of e_{ij} and π_{ijk} are ten times those of Graph 3.4.1a. Now different randomizations lead to different conclusions. One randomization gives $(j,k;Y_{ijk})$ as $(1,1;4.5)$, $(2,3;4.5)$, $(4,3;-3.5)$ and $(3,4;4.5)$, erroneously suggesting that treatment 3 has a negative effect. The base response and treatment effects are identical in both graphs.

Graphs 3.4.1b and 3.4.2b show the range of plot responses projected on the treatment axis. In the former, the spread of responses, the thickness of the plot dimension, is small. In the latter the spread is large. Only when the errors are small relative to the treatment effects (Graph 3.4.1a) are the conclusions free of the randomization used to conduct the experiment.
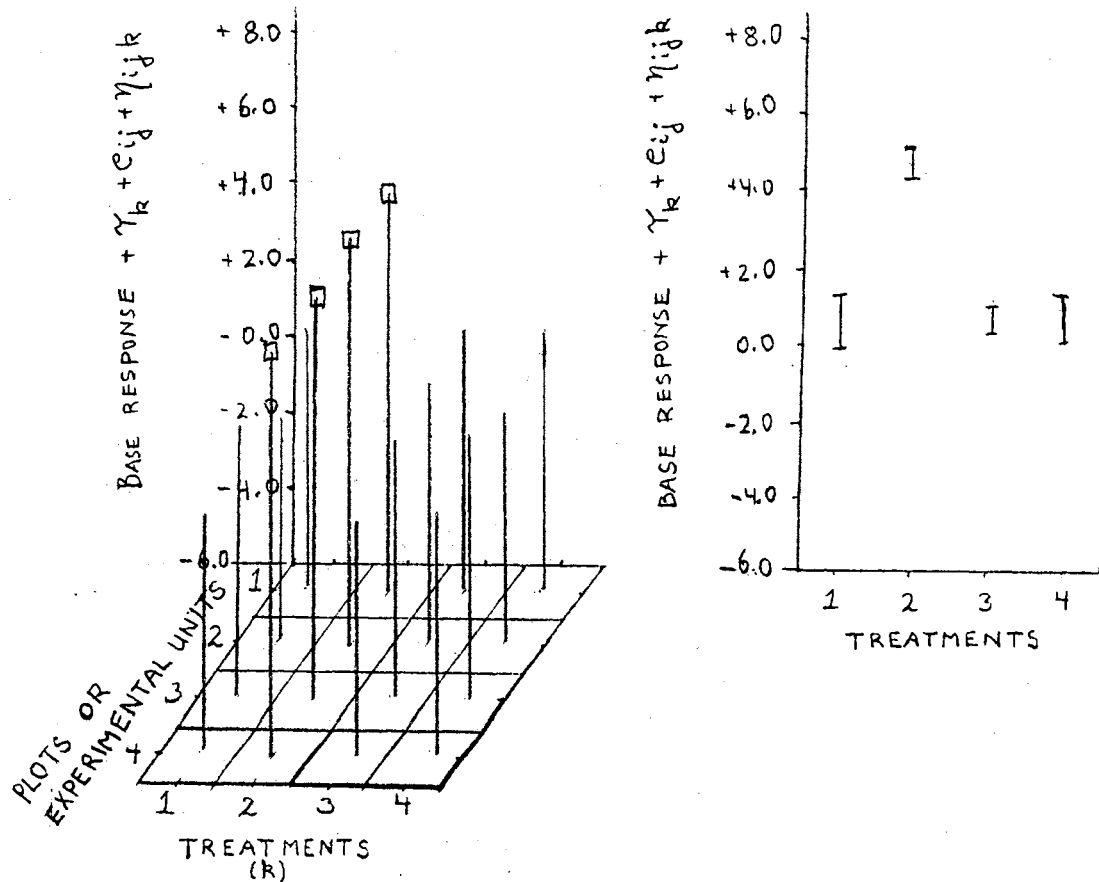
Graphs 3.4.1(a,b)

Near Identical Experimental Units
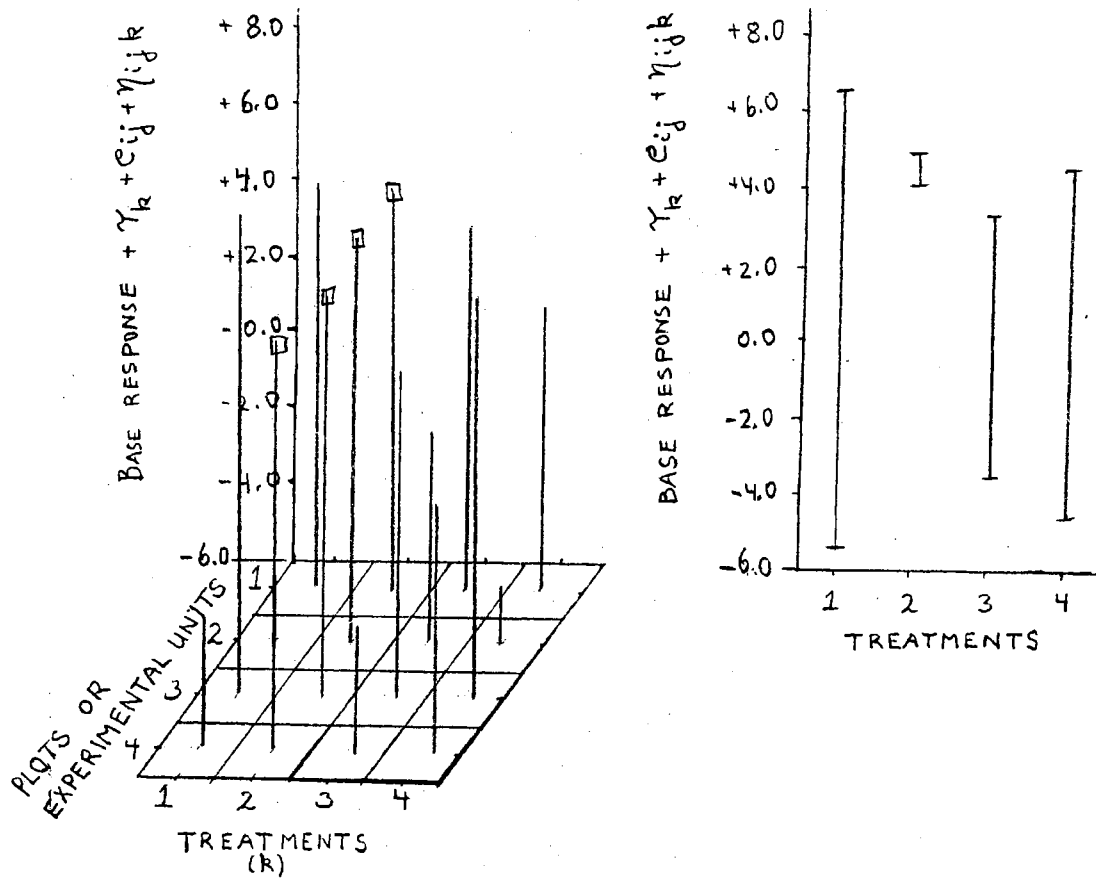
Graphs 3.4.2(a,b)

Nonidentical Experimental Units

3.4.4 Implications and Inferences

Being able to fix the probability of selecting any particular randomization permits a second approach to achieving such desirable attributes in an estimator as unbiasedness and minimum variance. The first approach, applicable in both models, is to alter the functional form of estimators and test statistics. We do not attempt this, but instead hold fixed such functions as

$\hat{F}_{\mu,b,\beta}$ and $F(\tau|\mu,b,\beta)$, as defined in Sections 2.4-2.6.
The second approach, precluded in the normally distribu-
ted random errors model, is to alter the selection
probabilities, the $p(s)$, and hence, the assignment
probabilities, $p(ij:k)$. This alters the expectation of
estimators and test statistics in the randomization
model and hence, their bias and variance. The pair
$\{p(s),$ estimator$\}$ is termed a strategy. Sections 4.7 and
4.8 below discuss alternative strategies, strategies with
unequal selection probabilities.

The randomization model treats its population as
finite and, strictly speaking, must limit its conclusions
to that population. Invariably, one takes a leap of
faith and claims that the new population of experimental
units is sufficiently like the one examined to apply to
it the results obtained from the previous experiment.
The leap of faith occurs earlier in the normally distrib-
uted random errors model. At the onset, one assumes a
distribution, the normal one; a covariance structure, Ψ
or Π; identical experimental units (all $e_{ij}=0$); and no
interaction between experimental unit and treatment (all
$\eta_{ijk}=0$). Having assumed that the experimental units
studied were drawn from an infinite population, the
researcher is free to apply his/her results to future
experimental units assumed to come from the same
population. This model has the advantage of known

distributions for statistics of interest. Underlying

its use is the assumption that the error term in the

model equation arises not from variations among the

experimental units, but rather from a sum of many random

variables. These random variables have unknown distribu-

tions but satisfy some central limit theorem. Thus, the

error term comes to approximate a normal distribution.

As detailed in Section 3.3.3 (p. 124), the randomization

model makes quite different assumptions about its errors.

White [1975] discusses these issues.

It is possible to combine the two models. We do so

in the following section.

Section 3.5

The Normal-Randomization Model

This section develops a matrix notation which allows combining the normally distributed random errors model and the randomization model. Sections 3.5.3 and 3.5.4 use the notation to derive the mean and variance of the vector of responses.   Section 3.5.5 discusses the super-population approach.   The final section discusses exchangeability.   The main results are (3.5.1), (p. 156), (3.5.2), (p. 157) and Theorems 3.5.2(e), (p. 160) and 3.5.5, (p. 163).

3.5.1   Introduction

The notation typically used in the randomization model employs a zero-one indicator random variable.   It has value one if treatment k is assigned to the experimental unit under consideration and zero otherwise. This notation is suitable when all assignment probabilities, $p(ij:k)$, are equal.   It is awkward when they are not. Use of selection probabilities, $p(s)$, equal or not, calls for an alternative notation.   The following notation permits using $p(s)$ as well as combining the normally distributed random errors model and the randomization model.

3.5.2   The Model Equation

This notation builds an alternative model equation from several large, partitioned, matrices.   We define the

matrices Y, X, T, E, H and N. The first two symbols are
redefined from previous use. Assumptions 1-9 continue to
hold. We illustrate the model equation using a design
with two blocks, i=1,2; four treatments, k=1,2,3,4; and
one experimental unit, j=1,2,3,4, per treatment per
block. Changes needed to fit other designs appear to be
straightforward.

The matrix of responses is Y, with dimension bpr x
$N(E)$. Each column, Y_s, represents all bpr responses
arising from one randomization. In the example r=1.
Each column is identically ordered as to blocks, and
within blocks as to treatments. By convention, we
permute the plots starting with the last block. The
first b-1 blocks begin in the first of the $N(B)$ block-
wise randomizations. After building a column for each of
the $N(B)$ block-wise randomizations in the last block, we
move to the second permutation in the penultimate block
and build all $N(B)$ columns for the block-wise randomiza-
tions of the last block. Once all $N(B)$ block-wise
randomization are listed for the penultimate block, we
set the third-to-last block to the second block-wise
randomization and repeat the above. This continues until
the last column has all blocks in their last block-wise
randomization. The Y matrix for the example has elements
Y_{ijk}. Arrows point to some of the permuted subscripts.
Only the plot subscript (j) changes. The matrix is

$$
\mathbf{Y} = \begin{bmatrix}
Y_{111} & Y_{111} & Y_{111} & | & Y_{111} & | & Y_{141} \\
Y_{122} & Y_{122} & Y_{122} & | & Y_{122} & | & Y_{132} \\
Y_{133} & Y_{133} & Y_{133} & | & Y_{143} <- & | & Y_{123} \\
Y_{144} & Y_{144} & \cdots & Y_{144} & | & Y_{134} <- & | & Y_{114} \\
Y_{211} & Y_{211} & Y_{241} & | & Y_{211} & \cdots & | \cdots & Y_{241} \\
Y_{222} & Y_{222} & Y_{232} & | & Y_{222} & | & Y_{232} \\
Y_{233}<->Y_{243} & & Y_{223} & | & Y_{233} & | & Y_{223} \\
Y_{244}<->Y_{234} & & Y_{214} & | & Y_{244} & | & Y_{214}
\end{bmatrix} .
$$

The matrix of model equation parameters, T, has a
block diagonal form with bpr identical blocks. Each
block has dimensions 1 x [1+(b-1)+(t-1)+q]. The matrix T
has dimensions bpr x bpr[1+(b-1)+(t-1)+q]. For the
example, each block is [μ b_1 τ_1 τ_2 τ_3 β], giving T
dimensions [2x4x1] x {[2x4x1][1+(2-1)+(4-1)+1]}, with

$$
\mathbf{T} = \begin{bmatrix}
\mu\ b\ \tau_1\ \tau_2\ \tau_3\ \beta & & & \text{zero} \\
 & \mu\ b\ \tau_1\ \tau_2\ \tau_3\ \beta & & \\
 & & \ddots & \\
\text{zero} & & & \mu\ b\ \tau_1\ \tau_2\ \tau_3\ \beta
\end{bmatrix} .
$$

The matrix $\mathbf{X}$ has one column for each of the
s=1,2,...,N(Ξ) experiment-wise randomizations. Each
column, $\mathbf{X}_s$, has one block for each block-treatment
combination, ik. Each of these blocks, $\mathbf{X}_{s,ik}$ resembles
one (or more, if r>1) transposed row(s) of the $\mathbf{X}^+$ of
(2.3.1), (p. 16). When r>1 experimental units receive
the same treatment in each block, all r rows of $\mathbf{X}^+$
receiving treatment k are included in the column $\mathbf{X}_{s,ik}$.
The indicators and covariates of each $\mathbf{X}_{s,ik}$ are ordered
to conform to the structure of the blocks of T. The
design or incident matrix part of $\mathbf{X}_{s,ik}$ indicates the

block and treatment assigned. These rows are constant for all columns of X. The covariates are tied to the experimental units. They are placed in the X_s to correspond to the experimental units receiving treatment k. The covariate part, Z_{ij}, of X changes in each X_s. In general, $X_{s,ik}$ has dimension $r[1+(b-1)+(t-1)+q]$ x 1 and X has dimension $bpr[(1+(b-1)+(k-1)+1]$ x $N(\mathcal{E})$. For the example, with s=24, X_s is the following 2x4=8 $X_{s,ik}$ stacked one upon the other in the same block-treatment order as the Y_s.

$$X_{24,11}=\begin{bmatrix}1\\1\\1\\0\\0\\Z_{11}\end{bmatrix} \quad X_{24,12}=\begin{bmatrix}1\\1\\0\\1\\0\\Z_{12}\end{bmatrix} \quad X_{24,13}=\begin{bmatrix}1\\1\\0\\0\\1\\Z_{13}\end{bmatrix} \quad X_{24,14}=\begin{bmatrix}1\\1\\-1\\-1\\-1\\Z_{14}\end{bmatrix}$$

$$X_{24,21}=\begin{bmatrix}1\\-1\\1\\0\\0\\Z_{24}\end{bmatrix} \quad X_{24,22}=\begin{bmatrix}1\\-1\\0\\1\\0\\Z_{23}\end{bmatrix} \quad X_{24,23}=\begin{bmatrix}1\\-1\\0\\0\\1\\Z_{22}\end{bmatrix} \quad X_{24,24}=\begin{bmatrix}1\\-1\\-1\\-1\\-1\\Z_{21}\end{bmatrix}.$$

The full X matrix has $N(\mathcal{E})$ columns of X_s as in

$$X=\begin{bmatrix} X_{1,11} & X_{2,11} & X_{s,11} & X_{N(\mathcal{E}),11} \\ X_{1,12} & X_{2,12} & \cdot & X_{N(\mathcal{E}),12} \\ X_{1,13} & X_{2,13} & \cdot & X_{N(\mathcal{E}),13} \\ X_{1,14} & X_{2,14} & X_{s,1t} & X_{N(\mathcal{E}),14} \\ X_{1,21} & X_{2,21} & X_{s,b1} & X_{N(\mathcal{E}),21} \\ X_{1,22} & X_{2,22} & \cdot & X_{N(\mathcal{E}),22} \\ X_{1,23} & X_{2,23} & \cdot & X_{N(\mathcal{E}),23} \\ X_{1,24} & X_{2,24} & X_{s,bt} & X_{N(\mathcal{E}),24} \end{bmatrix}.$$

Note that the treatment subscripts for each $X_{s,ik}$ match the corresponding row of Y. For each block, the product YX will always select τ_1 for its first row, τ_2 for its second row, etc.

The matrix of experimental unit errors, E, has the same form as Y. It has dimensions bpr x $N(E)$ and columns E_s, each with submatrices $E_{i,s(B)}$, where $s(B)$ is the block-wise randomization corresponding to the experiment-wise randomization s. Each element of $E_{i,s(B)}$ is the experimental unit error, e_{ij}, for the $j\underline{th}$ experimental unit appropriate for the randomization. For the example, the submatrices of E_{24} are

$$E_{24} = \begin{bmatrix} E_{1,1} \\ E_{2,24} \end{bmatrix}, \text{ with } E_{1,1} = \begin{bmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{14} \end{bmatrix} \text{ and } E_{2,24} = \begin{bmatrix} e_{24} \\ e_{23} \\ e_{22} \\ e_{21} \end{bmatrix}.$$

The entire E matrix is

$$E = \begin{bmatrix} E_{1,1} & E_{1,1} \cdots E_{1,1} | E_{1,2} \cdots E_{1,2} & | \cdots | & E_{1,\otimes} & \cdots & E_{1,\otimes} \\ E_{2,1} & E_{2,2} \cdots E_{2,\otimes} | E_{2,1} \cdots E_{2,\otimes} & | \cdots | & E_{2,1} & \cdots & E_{2,\otimes} \end{bmatrix}.$$

Note that the experimental unit subscripts for the e_{ij} of each column of E_s match those of Y_s. The same e_{ij} reappear in each column permuted differently. For each block i, all $j=1,2,\ldots,p$ experimental unit errors appear an equal number of times in each row. Thus, the first equality of (3.4.1), Assumption 9, (p. 140), gives as zero the sum of each row of H.

The matrix of experimental unit-treatment interactions, H, has the same form as Y. It has dimensions bpr x $N(E)$ and columns H_s, each with submatrices $H_{i,s(B)}$, where $s(B)$ is the block-wise randomization corresponding to the experiment-wise randomization s. Each element of

$H_{i,s(B)}$ is the experimental unit x treatment interaction

error, n_{ijk}, appropriate for the randomization. For the

example, the submatrices of H_{24} are

$$H_{24}=\begin{bmatrix} H_{1,1} \\ H_{2,24} \end{bmatrix}, \text{ with } H_{1,1}=\begin{bmatrix} n_{111} \\ n_{122} \\ n_{133} \\ n_{144} \end{bmatrix} \text{ and } H_{2,24}=\begin{bmatrix} n_{241} \\ n_{232} \\ n_{223} \\ n_{214} \end{bmatrix}.$$
$$\uparrow$$

The arrow notes the rearranged units (j) of block 2.

The entire H matrix is

$$H=\begin{bmatrix} H_{1,1} & H_{1,1} & \cdots & H_{1,1} & | & H_{1,2} & \cdots & | & \cdots & | & H_{1,\mathcal{B}} & \cdots & H_{1,\mathcal{B}} \\ H_{2,1} & H_{2,2} & \cdots & H_{2,\mathcal{B}} & | & H_{2,1} & \cdots & | & \cdots & | & H_{2,1} & \cdots & H_{2,\mathcal{B}} \end{bmatrix}.$$

Note that the treatment subscripts for each column of H

match the corresponding column of Y. Each row of H

contains the n values for one treatment of one block.

All plots appear an equal number of times in each row.

The center equality of restriction (3.4.1), Assumption

9, (p. 140), gives a zero sum for each row of H.

The matrix N of normally distributed random errors

has the same form and dimensions as Y. Under Assumption

3, (p. 24), these errors are identically distributed, so

they need no subscripts. Every element of N has the

identical symbol. As they represent random variables,

their values differ. Under Assumption 3[*], (p. 16), the

variance-covariance matrix is $\Psi\sigma^2$. In this case, the

errors would need subscripts, perhaps block and plot

subscripts, perhaps only the block subscript, depending

upon Ψ. While the columns of E permute identical

elements, each column of $\mathbb{N}$ has unique elements. Should

each column of $\mathbb{N}$ be identical, $\mathbb{N}$ would be equivalent to

the experimental unit error matrix $\mathbb{E}$.   Section 3.7.3,

below, (p. 198), also discusses this point.

Having defined the necessary matrices, the

normal-randomization model equation is

$$\Upsilon = \mathbb{T}\,\mathbb{X} \; + \; \mathbb{E} \; + \; \mathbb{H} \; + \; \mathbb{N} \qquad\qquad (3.5.1)$$

with dimensions for this example of

$\mathbb{T}$:   bpr x bpr[1+(b-1)+(k-1)+q] ,

$\mathbb{X}$:   bpr[1+(b-1)+(k-1)+q] x $N(\mathbb{E})$ and

Υ, $\mathbb{E}$, $\mathbb{H}$, $\mathbb{N}$:   bpr x $N(\mathbb{E})$.

A model may include one or more of $\mathbb{E}$, $\mathbb{H}$ and $\mathbb{N}$.   Slightly

more compact versions of this notation are possible.

The selection matrix, $\mathbb{S}$, has dimensions $N(\mathbb{E})$ x 1.

It is a column vector with zeros everywhere except for

the $s^{\underline{th}}$ row, which contains a one. By postmultiplying

the two sides of (3.5.1), it selects (a) the responses

actually observed from the matrix Υ, (b) the assignment

of treatments to units from the product $\mathbb{T}\mathbb{X}$ and (c) the

set of experimental unit x treatment interaction errors

from matrix $\mathbb{H}$.   The form of the matrices $\mathbb{E}$ and $\mathbb{N}$ permits

$\mathbb{S}$ to carry along the proper arrangement of these errors.

The value of S selected by (3.2.7), (p. 117), determines

which row of $\mathbb{S}$ has value one. Denote the selected

randomization as $\mathbb{S}_s$, obtained by the method of (3.2.7).

The post-multiplication by $\mathbb{S}_s$ is

$$\Upsilon \, \mathbb{S}_s = [\; \mathbb{T} \, \mathbb{X} \; + \; \mathbb{E} \; + \; \mathbb{H} \; + \; \mathbb{N} \;]\mathbb{S}_s$$

$$\Upsilon_s = (\mathbb{T} \, \mathbb{X})_s + \mathbb{E}_s + \mathbb{H}_s + \mathbb{N}_s . \qquad\qquad (3.5.2)$$

The subscript s identifies the vectors selected from the larger matrices. These vectors determine the responses observed in the experiment actually conducted.

Such a model equation was suggested by one for regression from finite populations developed by Jonrup and Rennermalm [1976].

3.5.3 The Combined Probability Space

When $\mathbb{N}$ is included in the model, the probability space for the normal-randomization model combines the probability spaces of Sections 2.3, (p. 15) and 3.4, (p. 140). It is

$$(\; (\mathbb{U} \times \mathbb{R}^n), \; (\mathcal{A} \times \mathcal{B}^n), \quad \mathbb{P} \times N_n(\mathbb{O}, \Psi\sigma^2) \;),$$

where n=bpr. $\qquad\qquad\qquad\qquad\qquad\qquad (3.5.3)$

Since $\mathbb{S}$ is an indicator vector, its expectation is a vector, $\mathbb{P}$, whose elements are the selection probabilities, $p(s)$, $\mathbb{E}_R(\mathbb{S})=\mathbb{P}$. As indicated in (3.5.3) the probability functions for the two distributions are statistically independent. Since $\mathbb{P}$ is a discrete probability function, the expectation with respect to the normal distribution and the expectation with respect the randomization distribution may be interchanged freely, that is, $E[\cdot] \equiv \mathbb{E}_N\mathbb{E}_R[\cdot] = \mathbb{E}_R\mathbb{E}_N[\cdot]$. This follows from a

theorem by Tonelli [Royden, 1968, p. 270].

In some cases, agricultural studies for example, the variance is known to increase with the magnitude of the response. Thus, the normally distributed random error associated with a large value of b_i, τ_k, βZ_{ij}, e_{ij}, and/or η_{ijk} would have a larger variance than would the error associated with smaller values. We do not examine this case. As indicated by the above Cartesian products, the values of the e_{ij} and/or η_{ijk} do not alter the normal distribution part of the probability space. This is restated as

> <u>Assumption 10</u>.
>
> The randomization distribution and the
> distribution of the normally distributed
> random errors are statistically independent.

3.5.4 Expectations of the Response Vector Y_s

In the matrix notation developed in Section 3.5.2, expectations of the terms of (3.5.2), (p. 157), are straightforward. For example, Theorems 3.5.2(e) and 3.5.4 derive the first two moments of Y_s; Theorem 3.5.5 derives the variance.

<u>Theorem 3.5.1</u> Moments of the Distribution

a. $E_R[\mathbb{S}] = \mathbb{P}$.

> Proof: Each elements of $\mathbb{S}$ is either 0 or 1.
>
> The expectation of a {0,1} indicator variable
>
> is its probability.

b. $E_R[\mathbb{S}\mathbb{S}'] = \sum_{\Omega}\{p(s)[\mathbb{S}_s\mathbb{S}_s']\} = \text{diag}(\mathbb{P})$.

> Proof: Each matrix $\mathbb{S}_s\mathbb{S}_s'$ has dimensions $N(\mathcal{E})\times N(\mathcal{E})$
>
> with a one on a unique diagonal position and zeros
>
> elsewhere. The product $p(s)\mathbb{S}_s\mathbb{S}_s'$ converts the one
>
> on the diagonal to $p(s)$. Summing gives the result.

c. $\text{Var}_R[\mathbb{S}] = \text{diag}(\mathbb{P}) - \mathbb{P}\mathbb{P}'$,

> with diagonal elements $p(s)-[p(s)]^2 = p(s)[1-p(s)]$
>
> and off-diagonal elements $-p(s)p(s')$, $s \neq s'$.
>
> Proof: Apply $\text{Var}(\cdot)=[E(\cdot)^2]-[E(\cdot)]^2$ to (a) and (b).

<u>Theorem 3.5.2</u> First Moments of Equation Parameters

a. $E[(\mathbb{T}\mathbb{X})_s] = E_R[(\mathbb{T}\mathbb{X})\mathbb{S}] = (\mathbb{T}\mathbb{X})E_R[\mathbb{S}] = \mathbb{T}\mathbb{X}\mathbb{P}$.

> Proof of a – d: Apply Theorem 3.5.1(a).

When $\mathbb{P}$ has equal elements, $\mathbb{P}=\mathbb{1}/N(\mathcal{E})$, the covariate
is averaged over the pr experimental units in each block.
For example, the first element of $E(\mathbb{T}\mathbb{X}_s)$ is

$\mu+b_1+\tau_1+\beta(\overline{Z}_1.)$, the second is $\mu+b_1+\tau_2+\beta(\overline{Z}_1.)$, etc.

b. $E[\mathbb{E}_s] = E_R[\mathbb{E}\mathbb{S}] = \mathbb{E}\mathbb{P}$.

c. $E[\mathbb{H}_s] = E_R[\mathbb{H}\mathbb{S}] = \mathbb{H}\mathbb{P}$.

When $\mathbb{P}$ has equal elements, $\mathbb{P}=\mathbb{1}/N(\mathcal{E})$, $\mathbb{P}$ averages each row of $\mathbb{E}$ and $\mathbb{H}$. Each label, ij or ijk, appears an equal number of times in each row. Under this condition, Assumption 9, (p. 140), gives zero as the row sum and $E[\mathbb{E}_S]=E[\mathbb{H}_S]=\mathbb{O}$.

d. $E[\mathbb{N}] = E_R E_N[\mathbb{N}] = \mathbb{O}$.

   Proof: Assumptions 2, (p. 16) and 10, (p. 158).

e. $E[\mathbb{Y}_S] = E_N E_R[\mathbb{Y}] = E_R\{[\mathbb{T}\mathbb{X} + \mathbb{E} + \mathbb{H} + E_N(\mathbb{N})]\ \mathbb{S}\}$

$= E_R\{(\mathbb{X}\mathbb{T}+\mathbb{E}+\mathbb{H}+\mathbb{O})\mathbb{S}\} = (\mathbb{X}\mathbb{T}+\mathbb{E}+\mathbb{H})[E_R(\mathbb{S})]$

$= \begin{cases} [(\mathbb{T}\mathbb{X}) + \mathbb{E} + \mathbb{H}]\mathbb{P} & \text{for general } p(s) \\ (\mathbb{T}\mathbb{X})[\mathbb{1}/N(\mathcal{E})] & \text{when all } p(s) \text{ are equal.} \end{cases}$

   Proof: Combine a - d.

<u>Theorem 3.5.3</u> Second Moments of Equation Parameters

a. $E[\mathbb{E}_S\mathbb{E}_S'] = E_R[\ (\mathbb{E}\mathbb{S})(\mathbb{E}\mathbb{S})'\ ] = E_R[\ \mathbb{E}(\mathbb{S}\mathbb{S}')\mathbb{E}'\ ]$

$= \mathbb{E}[E_R(\mathbb{S}\mathbb{S}')]\mathbb{E}' = \mathbb{E}[\text{diag}(\mathbb{P})]\mathbb{E}'$.

   Proof of a - f: Apply Theorem 3.5.1b.

b. $E[\mathbb{H}_S\mathbb{H}_S'] = \mathbb{H}[\text{diag}(\mathbb{P})]\mathbb{H}'$.

c. $E[\mathbb{E}_S\mathbb{H}_S'] = \mathbb{E}[\text{diag}(\mathbb{P})]\mathbb{H}'$.

When $\mathbb{P}$ has equal elements, $\mathbb{P}=\mathbb{1}/N(\mathcal{E})$ and $\text{diag}(\mathbb{P}) = \mathbb{I}/N(\mathcal{E})$. The equalities above become $\mathbb{E}\mathbb{E}'/N(\mathcal{E})$, $\mathbb{H}\mathbb{H}'/N(\mathcal{E})$, and $\mathbb{E}\mathbb{H}'/N(\mathcal{E})$. Each of these products is block diagonal. Each off-diagonal block is product of a row from one

block times a column from another block. The randomiza-
tion will pair for each block a fixed e_{ij} (or n_{ijk} for

(b)) subscript of one row with all subscripts of the

column-block being multiplied. This sums that column-

block. By Assumption 9, (p. 140), such a sum over all

experimental units in a block is zero. Consequently, all

off-diagonal blocks of the matrix are zero when $P=1/N(E)$.

The same holds for off diagonal elements of the

diagonal blocks. For (a) the diagonal elements are

$(e_{ij})^2$ times p, the number of elements in the block. For

(b) the diagonal elements are $(n_{ijk})^2$ times p. For (c)

the fixed e_{ij} is multiplied by all n_{ijk} for that block.

This sums the n_{ijk}, leaving all elements of (c) as zero.

d. $E[(TX)_S E_S'] = E_R[(TXS)(ES)'] = (TX)[diag(P)]E'$

e. $E[(TX)_S H_S'] = E_R[(TXS)(HS)'] = (TX)[diag(P)]H'$

When P has equal elements, $P=1/N(E)$ and $diag(P) =$
$I/N(E)$. Recall that each row of TX is identical except
for the covariate part. The incident or design parts sum
the columns of the righthand matrix. By Assumption 9
each such sum is zero. Thus, the above two equalities
reduce to the product of a matrix structured like Y, but
with elements $Z'P$, times, respectively, E or H. Without
a covariate these two products would be zero.

f. $E[(TX)_S(TX)_S'] = E_R[(TXS)(TXS)'] = (TX)[diag(P)](TX)'$

g. $E[\mathbb{N}_S\mathbb{N}_S'] = E_N[\mathbb{N}_S\mathbb{N}_S'] = \mathbb{I}\sigma^2$

　　Proof: Assumptions 10, 2 and 3, (p. 158, 16 and 24).

h. $E\{[(\mathbb{TX})_S + \mathbb{E}_S + \mathbb{H}_S]\mathbb{N}_S'\} = E_R[(\mathbb{TX} + \mathbb{E} + \mathbb{H})\mathbb{S}]E_N(\mathbb{N}')$

　　　　$= E_R[(\mathbb{TX} + \mathbb{E} + \mathbb{H})][\mathbb{O}']\} = \mathbb{O}$, by Assumption 10.

Theorem 3.5.4   Second Moment of the Response Vector Υ_S

　　$E[\Upsilon_S\Upsilon_S'] = (\mathbb{TX} + \mathbb{E} + \mathbb{H})[diag(\mathbb{P})][(\mathbb{TX} + \mathbb{E} + \mathbb{H})\mathbb{S}]'$

　　　　$+ \Psi\sigma^2$.

　　Proof:

　　$E[\Upsilon_S\Upsilon_S'] = E_N E_R[\Upsilon_S\Upsilon_S'] = E_N E_R[(\Upsilon\mathbb{S})(\Upsilon\mathbb{S})'\}$

　　$= E_N E_R\{[(\mathbb{TX} + \mathbb{E} + \mathbb{H} + \mathbb{N})\mathbb{S}][(\mathbb{TX} + \mathbb{E} + \mathbb{H} + \mathbb{N})\mathbb{S}]'\}$

　　$= E_R\{[(\mathbb{TX} + \mathbb{E} + \mathbb{H})\mathbb{S}][(\mathbb{TX} + \mathbb{E} + \mathbb{H})\mathbb{S}]'\}$

　　　　$+ E_R E_N \{2[(\mathbb{TXS})(\mathbb{NS})' + (\mathbb{ES})(\mathbb{NS})' + (\mathbb{HS})(\mathbb{NS})']$

　　　　　　$+ (\mathbb{NS})(\mathbb{NS})' \}$

　　$= (\mathbb{TX} + \mathbb{E} + \mathbb{H})[diag(\mathbb{P})](\mathbb{TX} + \mathbb{E} + \mathbb{H})'$

　　　　$+ E_R\{2[\mathbb{O} + \mathbb{O} + \mathbb{O}] + E_N(\mathbb{NS})(\mathbb{NS})' \}$

　　$= (\mathbb{TX} + \mathbb{E} + \mathbb{H})[diag(\mathbb{P})](\mathbb{TX} + \mathbb{E} + \mathbb{H})'$

　　　　$+ 2E_N\{\mathbb{N}[diag(\mathbb{P})]\mathbb{N}'\}$

　　$= (\mathbb{TX} + \mathbb{E} + \mathbb{H})[diag(\mathbb{P})](\mathbb{TX} + \mathbb{E} + \mathbb{H})'$

　　　　$+ \{trace[diag(\mathbb{P})]\}\Psi\sigma^2$.

The triple of zeros follows from $E_N[\mathbb{N}] = \mathbb{O}$.   The

expectation of a quadratic form [Searle, 1971, p. 55] is

applied to each of the identical bp rows of $\mathbb{N}$ and

identical bp columns of $\mathbb{N}'$.   The trace$[diag(\mathbb{P})]\Psi\sigma^2 = \Psi\sigma^2$,

because the elements of $\mathbb{P}$, the $p(s)$, must sum to 1.

When $\mathbb{S}=\mathbb{1}/N(\mathbb{E})$, the individual terms simplify as discussed above.

<u>Theorem 3.5.5</u> Variance of the Response Vector

$$\mathrm{Var}(\mathbf{Y}_S) = (\mathbf{TX} + \mathbb{E} + \mathbb{H})[\mathrm{diag}(\mathbb{P})](\mathbf{TX} + \mathbb{E} + \mathbb{H})' + \Psi_\sigma^2$$

$$- (\mathbf{TX} + \mathbb{E} + \mathbb{H})[\mathrm{diag}(\mathbb{P})]^2(\mathbf{TX} + \mathbb{E} + \mathbb{H})'$$

$$= (\mathbf{TX} + \mathbb{E} + \mathbb{H})[\mathrm{diag}(\mathbb{P})-\mathbb{PP}'](\mathbf{TX} + \mathbb{E} + \mathbb{H})'$$

$$+ \Psi_\sigma^2.$$

Proof: Apply Theorems 3.5.4 and 3.5.2(e) to the variance relationship $\mathrm{Var}(\cdot)=E[(\cdot)^2] - [E(\cdot)]^2$.

With knowledge of the values of e_{ij} and η_{ijk}, one perhaps could alter the elements of $\mathbb{P}$, the $p(s)$, to, say, minimize the variance of $\mathbf{Y}_S$ while leaving it unbiased. The mechanics of this are an extension.

3.5.5 Superpopulation Approach

One may also consider the available experimental units, the ij part of $\mathfrak{y}$, to be a random sample of a larger population, perhaps one now existing, or a conceptual one including experimental units that might exist in the future. Under this viewpoint, the e_{ij} and η_{ijk} are themselves random variables. One approach is to enter them as population means in the model equation and move deviations from these means into the normally distributed random error term. This larger population

suggests the superpopulation approach of survey sampling and focuses on the population to which the inferences of the current experiment will be applied. Expectation with respect to this superpopulation may in fact be what is desired from a designed experiment. See Cassel, Sarndal, and Wretman [1977] for a discussion of the fixed and superpopulation approaches in survey sampling. B. L. Welch [1937, p. 47-48] and Kempthorne and Doerfler [1969, pages 231-235] touch on this concept.

3.5.6 Exchangeability

Random events are exchangeable if and only if their joint distribution is unaltered by any permutation of the labels identifying the individual events. Independent and identically distributed random errors, such as the normally distributed random errors of Chapter 2 are exchangeable random variables. Under the probability space for the randomization model, with $p(s)=[N(\mathcal{E})]^{-1}$, the $N(\mathcal{E})$ randomizations are exchangeable. When the $p(s)$ are equal within a group of randomizations, the randomizations within the group exhibit partial exchangeability. In both cases, all joint distributions have zero probability for two or more randomizations and constant probability of any one randomization within the group.

When $n_s = n_s = 0$, the functions defined below in Theorems 3.7.9, (p. 202), and 3.7.10, (p. 211), have identical values for all randomizations within the

g-group. When the p(s) within a g-group are equal

valued, such random variables exhibit partial

exchangeability.

In the limit, as the number of observations goes to

infinity, the joint probability distribution of

exchangeable random variables is a weighted average of

independent random variables, even if the original joint

probability distribution is not independent. This is "de

Finetti's representation theorem." See de Finetti [1937,

p. 15-17, 81-83, 99-101; 1970, Vol. 2, p. 215; 1977, p.

211], Chow and Teicher [1988, p. 226], or Cassel,

Särndal, Wretman, [1977, p. 72, 85] for this theorem.

The normal-randomization model embeds each randomi-

zation in an infinite sequence of events. Within each

randomization, the events of the sequence differ only by

the normally distributed random error. When $\eta_s = 0$, the Υ_s

of each g-group, (p.197), are so embedded. In this model,

some version of de Finetti's theorem appears to hold. If

so, a tool is available for use in applying limit theorems

to the functions of Theorem 3.7.9, (p. 202). Chow and

Teicher [1988, p. 33, 223-226 and 309-311] illustrate

such uses of exchangeability. Cifarelli, *et al.*, discuss

exchangeability in the experimental setting [1981A], in

the analysis of variance [1979] and in linear regression

[1981B]. See also Lindley and Novick [1981].

We next examine the randomization test with unequal

selection probabilities, p(s).

Section 3.6

The Randomization Test

Section 3.6.1 sketches the randomization test for treatments. Section 3.6.2 discusses the need for zero experimental unit-treatment interaction error. Sections 3.6.3 and 3.6.4 describe the randomization test and power computation for $p(s)$ equal or not. Section 3.6.5 discusses the randomization test and power computation under simulation. Section 3.6.6 shows that the randomization test for the covariate(s) as adjusted for the mean, blocks and treatments will have zero power for some randomizations when the values of the covariates repeat in each block. The final section, 3.6.7, shows the test satisfies the requirements of a hypothesis test. Major results are the OSL (3.6.2), and power (3.6.6), (p. 170-173), of the randomization test and of the test under simulation (3.6.10), (p. 175). See also Edgington [1987].

3.6.1 Overview of the Randomization Test

While the randomization model is a conceptual framework, the randomization test is a statistical test of a null hypothesis. In simulations or in rare cases the randomization test may be extended to measure the power of the test for a given alternative hypothesis. The parameter tested is one the experimentor may randomly assign to experimental units. Herein, this is the part of τ containing the treatments, as in (2.3.4) and (2.3.5), (p. 19).

The basic idea of the randomization test is that if the treatments have no effect, then each observed response, Y_{ijlk}, does not depend upon the treatment k to which experimental unit U_{ijl} was assigned. One computes the value of the test statistics for the randomization actually used to conduct the experiment. One then recomputes it as if the experimental units had been assigned to different treatments. One repeats this for all randomizations and obtains a set of values of the test statistic.

The test statistic is one whose values (a) will tend to be similar if all randomization are equal and (b) will tend to fall into a specified subset of all values, the set S, if the treatments have an effect. Should the test statistic provided by the actual experiment fall into the set S, then either (i) the null hypothesis is false, or (ii) an unusual (unlucky) randomization has been used to conduct the experiment. As stated in Section 3.4.4, (p. 141), the test statistic is fixed as the F-ratios of Chapter Two (2.6.41) and (2.6.45), (p. 63 and 64). The set S is defined so that (b) will tend to occur for the F-ratio, as shown in Section 3.6.7, (p. 181). The set S may differ for other statistics. It is the critical region.

The terms randomization, _rerandomization_ and _permutation_ tests are used with different meanings by different authors [Kotz and Johnson, Vol. 7, p. 524 and

530; Zyskind, *et al.*, 1964, p. 171-175; Kempthorne and Doerfler, 1969, p. 234-235; Gabriel and Hall, 1983; W. Welch, 1990, p. 693]. They usually refer to permuting the experimental units assigned to the treatments. This applies when $r=1$. But, for $r>1$, should the set $\mathbb{G}$ be constrained as in (3.2.12) or (3.2.14), (p. 115), then many permutations will violate the constraints. Thus, the term randomization appears more appropriate.

Cox and Hinkley [1974, p. 180 and 196] differentiate between the terms permutation and randomization. In their usage, the permutation test follows from a physical system generating independent and identical values, such as conditioning on the order statistics. Their randomization test follows from the randomization in the design which makes all permutations equally likely. We use the term randomization to emphasizes not only the reassigning of experimental units to treatments, but also the accompanying changes in experimental unit assignment probabilities, the $p(ij1:k)$'s. The latter is important when the assignment probabilities are not all equal.

3.6.2 Experimental Unit x Treatment Interaction

When the experimental units are artifically reassigned to new treatments, the experimental unit x treatment interactions, $\mathbb{H}_s$, are not, and outside of a simulation, cannot, be reassigned. The interaction

errors of the actual response become assigned to treatments not responsible for them. Because of this, the randomization test assumes the η_{ijk} are zero or negligible, as in Wilk [1955, p. 78-79, (a) and (b)]. This is the assumption of additivity in the strict or in the broad sense of (3.3.11), (p. 135).

3.6.3   Testing the Null Hypothesis

The test begins with the randomization actually used to conduct the experiment, the experiment conducted under ("at") randomization s, @s. The process defined in (3.2.7), (p. 109), is one method of obtaining the value of s. The responses from this randomization provide the actual responses, $Y_{@s}$, read as "the actual responses at randomization s." The actual responses and the actual randomization provide the actual test statistic $t_{@s,s=@s}$.

The next step computes the test statistic for another randomization, $s \neq @s$, using the data observed in the actual experiment, $Y_{@s}$. One pretends the experimental units were assigned to the treatments as specified by randomization s. In matrix notation, the columns of the treatments part of the actual design matrix $X_{T,@s}$, are rearranged within each block to represent randomization s. We denote the rearranged design matrix as X_s. The rows of X_s and the rows of $Y_{@s}$ remain linked to the same experimental units regardless of the randomization.

Repeating this for all $s \neq @s$ produces $N(\mathcal{E})-1$ artificial experiments and $N(\mathcal{E})-1$ artificial test statistics, $t_{@s,s \neq @s}$. These values, plus the actual test statistic, $t_{@s,s=@s}$, form the set

$$\mathcal{T}_{@s} = \{t_{@s,1}, t_{@s,2}, \ldots, t_{@s,s=@s}, \ldots, t_{@s,N(\mathcal{E})}\}. \quad (3.6.1)$$

The design, covariates and/or responses may be such that some artificial test statistics equal others. Theorem 3.7.9(3), (p. 202) specifies such a $t_{@s,s}$. Denoting the random value of $t_{@s,s}$ as $T_{@s,s}$, note that its distribution is known; $T_{@s,s}$ is a statistic.

The final step computes the p-value or the observed significance level, the OSL, for the randomized test as

$$OSL_{R,@s} \equiv \sum_{\mathcal{S}} p(s), \quad (3.6.2)$$

$$\mathcal{S} = \{s \text{ such that } t_{@s,s} \geq t_{@s,s=@s}\} .$$

The summation of $p(s)$ is over those s whose artificial test statistic is greater than or equal to the actual test statistic. Randomizations with values of $t_{@s,s}$ larger than $t_{@s,s=@s}$ and whose small values of $p(s)$ make them unlikely to be selected contribute less to the $OSL_{R,@s}$ than do those randomizations whose large values of $p(s)$ make them more likely to be selected. The actual test statistic, $t_{@s,@s}$, is included in $\mathcal{S}$ [Kempthorne,

1952, p. 130-131]. When all $p(s)$ are equal, one simply counts the randomizations with $t_{@s,s}$ greater than or equal to $t_{@s,s=@s}$, obtaining

$$OSL_{R,@s} = \left\{ \frac{N(t_{@s,s} \geq t_{@s,s=@s})}{N(\Xi)} \right\}. \qquad (3.6.3)$$

The selection of one randomization used to conduct the actual experiment provides an _internal criterion_, $t_{@s,s=@s}$, with which to assess the artificial test statistics, the $t_{@s,s}$. This internal criterion leads to the $OSL_{R,@s}$ and the test of the hypothesis. Note that the value of $OSL_{R,@s}$ depends upon the randomization used to conduct the actual experiment

The minimum value of the $OSL_{R,@s}$ is $p(@s)$, which equals $1/[N(\Xi)]$ if all $p(s)$ are equal. If the $OSL_{R,@s}$ is less than a predetermined α probability of a Type I error, or otherwise indicates a sufficiently small probability or cost of error, then the null hypothesis is rejected. Either the null hypothesis is false or @s is an unusual randomization.

3.6.4 Computing the Power of the Randomization Test

Kempthorne [1952, p. 230-231] details the computation of the sensitivity (power) of the randomization test for a known alternative hypothesis. Although his section title mentions the infinite model approach, the method assumes the presence of experimental unit error. We add

unequal $p(s)$ to his description and consider the possibility of known E, H and/or N. Term the base responses the responses one would obtain if no treatment were applied to the experimental units. A uniformity trial could supply such values. They are

$$\text{base responses} = [\,X_\mu \mid X_b \mid Z\,] \begin{bmatrix} \tau_\mu \\ \tau_b \\ \beta \end{bmatrix} + E., \qquad (3.6.4)$$

combining the notations of Sections 2.3 and 3.5.2, (p. 15 and 157). We use $E.$ to represent any of the equal column vectors of experimental unit error.

Step one assigns the experimental units to the treatments according to randomization $@s = @1$. Then, it adds the known values of the treatment effects, $H_{@1}$ and $N_{@1}$ to the base responses. This gives the actual responses, $Y_{@1}$, for randomization $@s = @1$. For each actual experiment conducted at randomization $@s = @1$, one conducts artificial experiments at randomization $s=1$, $s=2$, ..., $s \neq @s$, ..., $s = N(\Xi)$. For each s, $Y_{@1,s} = Y_{@1,@s} + (H_s - H_{@1}) + (N_s - N_{@1})$. This adjusts the $Y_{@1,s}$ to have the experimental unit x treatment interactions and normally distributed random error of the artificial experiment of randomization s. Proceed as in the Section 3.6.3 with the test of the null hypothesis and compute the set, $\mathcal{T}_{@1}$, of $N(\Xi)$ values of the artificial test statistic derived from the actual

experiment $@s=@1$. With the elements of $\mathfrak{T}_{@1}$, compute the

$OSL_{R,@1}$ by (3.6.2), (p. 168), or (3.6.3), (p. 169).

Repeat this process, making each randomization the

actual one. This gives $N(\Sigma)$ sets of $\mathfrak{T}_{@s}$,

$$\mathfrak{T}_{@1}=\{t_{@1,1} \ , \ \cdots \ , \ t_{@1,s} \ , \ \cdots \ , \ t_{@1,N(\Sigma)}\},$$

$$\mathfrak{T}_{@2}=\{t_{@2,1} \ , \ \cdots \ , \ t_{@2,s} \ , \ \cdots \ , \ t_{@2,N(\Sigma)}\},$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$\mathfrak{T}_{@N(\Sigma)}=\{t_{@N(\Sigma),1} \ , \ \cdots \ , \ t_{@N(\Sigma),N(\Sigma)}\}. \qquad (3.6.5)$$

A total of $N(\Sigma) \times N(\Sigma)$ test statistics need to be

computed. For each set $\mathfrak{T}_{@s}$ compute the $OSL_{R,@s}$ as in the

randomization test. The internal criterion for each $\mathfrak{T}_{@s}$

is the test statistic $t_{@s,@s}$. Using a fixed α, note

those $\mathfrak{T}_{@s}$ with $OSL_{R,@s}$ greater than α. These actual

randomizations improperly accept the false null

hypothesis. The probability of a Type II error, $\check{F}_R$, is

the sum of the probabilities of such randomizations, or

$$\check{F}_R \equiv \sum_{\underset{\sim}{S}} p(@s), \qquad\qquad (3.6.6)$$

$$\underset{\sim}{S} = \{@s \text{ such that } OSL_{R,@s} > \alpha\} \ .$$

The summation of $p(@s)$ is over those $@s$ whose $OSL_{R,@s}$ is

greater than α. The same selection probabilities are

used within each $\mathfrak{T}_{@s}$ and for the set of $N(\Sigma)$ $\mathfrak{T}_{@s}$ itself.

As with the $OSL_{R,@s}$, randomizations which accept the

erroneous null hypothesis and have a large probability

of being selected (large $p(s)$) contribute more to $\check{F}_R$ than

do randomizations with a small selection probability.

When all $p(s)$ are equal, the calculation is

$$\hat{p}_R = \left\{ \frac{N(\mathcal{T}_{@s}) \text{ with } OSL_{R,@s} > \propto}{N(\mathcal{E})} \right\} . \qquad (3.6.7)$$

The power of the randomization test is $1 - \hat{p}_R$.

In theory, the randomization test and power computation may be conducted for any simulation. In practice, $N(\mathcal{T}_{@s})$ and the product $N(\mathcal{T}_{@s}) \times N(\mathcal{T}_{@s})$ are prohibitively large. Gabriel and Hall [1983] and others they mention have devised methods to obtain significance levels and confidence intervals using but a fraction of the calculations required for the full randomization test.

3.6.5 The Randomization Test Under Simulation

In a simulation, no randomization is privileged as the one used to conduct the actual experiment. Instead of a set $\mathcal{T}_{@s}$, one has the set

$$\mathcal{T} = \{ t_{@1}, \; t_{@2}, \; \ldots, \; t_{@s}, \; \ldots, \; t_{@N(\mathcal{E})} \}. \qquad (3.6.8)$$

Each $t_{@s}$ is computed as if randomization s had been used to conduct the actual experiment, that is, computed like the $t_{@s, s=@s}$ of Section 3.6.3. Each randomization provides one $\Upsilon_{@s}$; all randomizations are used. When testing the null hypotheses, there is no privileged $t_{@s}$ to use in computing the $OSL_{R,@s}$ as in (3.6.2) or (3.6.3), (p. 170). Thus, the set $\mathcal{S}$ must be defined by an __external criterion__. One criterion is the point which would define the critical region under the normal distribution theory.

For a test of the null hypothesis with a fixed α, this criterion is the value of $F_{n,d,\lambda=0,(1-\alpha)}$ such that

$$P_F\{t_{@s} \geq F_{n,d,\lambda=0,(1-\alpha)}\} \leq \alpha , \qquad (3.6.9)$$

where $t_{@s}$ is assumed to have the F distribution with indicated parameters. Such $t_{@s}$ improperly reject a true null hypothesis.

The observed significance level for this test is the probability of observing such randomizations,

$$OSL_{R(F)} \equiv \sum_{\mathcal{S}} p(@s), \qquad (3.6.10)$$

$$\mathcal{S}=\{@s \text{ such that } t_{@s} \geq F_{n,d,\lambda=0,(1-\alpha)}\} .$$

The summation of $p(s)$ is over those s whose test statistic $t_{@s}$ is greater than or equal to $F_{n,d,\lambda=0,(1-\alpha)}$. When all $p(s)$ are equal the calculation is

$$OSL_{R(F)} = \left\{ \frac{N(t_{@s} \geq F_{n,d,\lambda=0,(1-\alpha)})}{N(\mathcal{E})} \right\}. \qquad (3.6.11)$$

Strictly speaking, this simulation is not a randomization test, as it relies upon an external distribution for the test criterion. This is indicated by the subscript notation R(F), where the F identifies the external criterion as the F distribution.

The computation of the power of the test, $1 - \hat{\beta}_{R(F)}$, may proceed as in Section 3.6.4. One considers in turn each randomization as the privileged one and obtains $N(\Xi)$ sets of $\mathcal{T}_{@s}$. For each set the external criterion enters via the replacement of (3.6.2), (p. 170), with

$$\text{OSL}_{R(F),@s} \equiv \sum_{S} p(s),$$

$$S = \{s \text{ such that } t_{@s,s} \geq F_{n,d,\lambda=0,(1-\alpha)}\} \; .$$

Then obtain the Type II error probability by replacing (3.6.6), (p. 173), with

$$\hat{\beta}_{R(F)} \equiv \sum_{S} p(@s),$$

$$S = \{@s \text{ such that } \text{OSL}_{R(F),@s} > \alpha\} \; .$$

A second method uses only the set $\mathcal{T}$ of (3.6.8). One defines

$$\hat{\beta}_{R(F)} \equiv \sum_{S} p(@s), \text{ with}$$

$$S = \{@s \text{ such that } t_{@s} < F_{n,d,\lambda=0,(1-\alpha)}\} \; .$$

Here, S is the set of s for which $t_{@s}$ erroneously accepts a false null hypothesis.

A third method includes the magnitude of each Type
II error. One weights the magnitude of each Type II
error by the selection probability of the randomization
committing the error. One defines

$$F_{R(F)} \equiv \sum_{\mathcal{S}} p(@s) F_{F,@s} \quad \text{with}$$

$$\mathcal{S} = \{@s \text{ such that } t_{@s} < F_{n,d,\lambda=0,(1-\alpha)}\}, \text{ and}$$

$$F_{F,@s} \equiv \int_0^{F_{n,d,\lambda=0,(1-\alpha)}} d(F_{n,d,\lambda_{@s} \neq 0,}).$$

The integrand is the noncentral F-distribution for the
true value of F with the noncentrality parameter
appropriate for randomization @s. One integrates from
zero to the point at which one would no longer accept
the erroneous null hypothesis, $F_{n,d,\lambda=0,(1-\alpha)}$. The set
$\mathcal{S}$ is those randomizations for which one commits a Type
II error.

While the randomization distribution will have one
probability of a Type II error, the F-distribution will
have one probability for each randomization due to the
change in its non-centrality parameter from one
randomization to the next. Examining the characteristics
of the above three methods is an extension. Graph 3.6.1
illustrates how the Type II error probabilities differ
with the randomization.
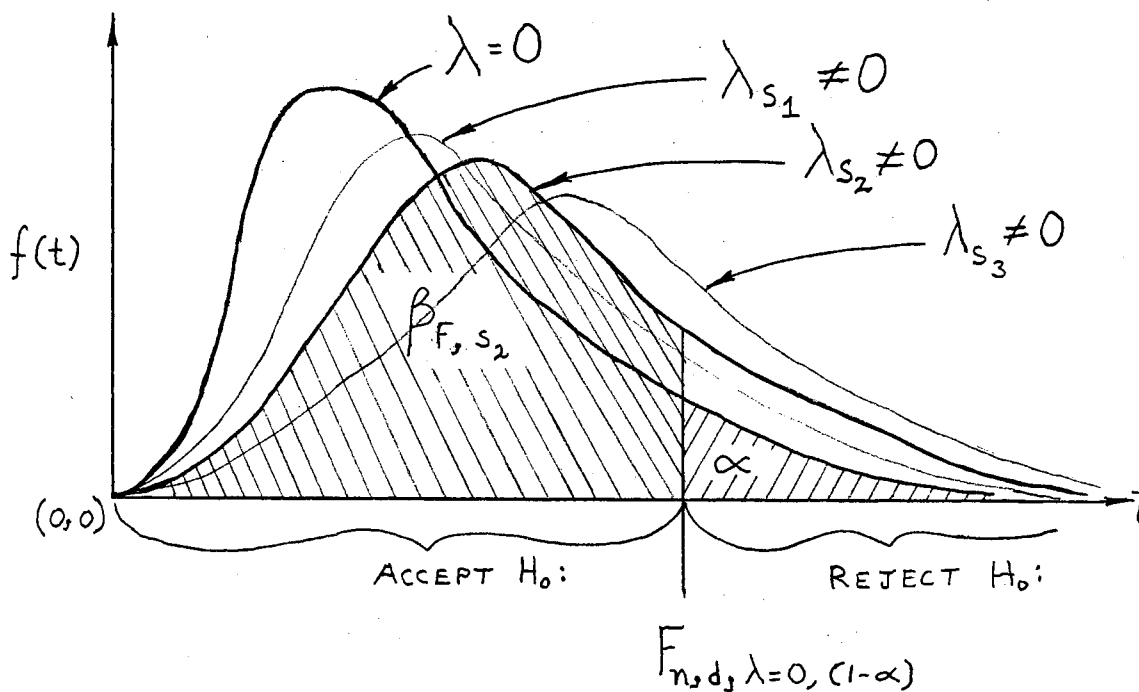
Graph 3.6.1

Type II Error Probabilities

3.6.6 Randomization Applied to $H_{o: \mu, b, \tau, \beta} \beta = 0$.

The randomization test is a test of a variable under the control of the experimentor. As per Section 2.2.1, (p. 8), the covariates are herein not subject to assignment by the experimentor. Consequently, strictly speaking, the randomization test cannot test $H_{o: \mu, b, \tau, \beta} \beta = 0$. Nevertheless, $F(\beta | \mu, b, \tau)$ depends upon the randomization, thus, so does this hypothesis test.

To take an extreme case, suppose Z has but one column whose elements, the covariates Z_{ij}, differ within a block, but repeat exactly in all blocks. An example of this is an industrial experiment, blocked on shifts, which adjust each experimental unit (a lathe) for a known covariate. The covariate, say operating revolutions per minute, differs from lathe to lathe, but remains constant from shift to shift for each lathe. The treatments could be fixed angles of the cutting tool and the responses the yield of machined parts.

We show that in this case, the randomization can force the value of the test statistic to be zero, regardless of the true value of β. The explanation follows from (2.6.10), (p. 52) and (2.6.28), (p. 55). In (2.6.10) X_h is Z and $X_r(X_r'X_r)^-X_r'$ has the form of (2.6.28). For this Z, when the treatments of X repeat in the same order in all blocks, the same treatment is linked to the same covariate value in each block. Such an X provides (2.6.28); in which case, $Z'X(X'X)^-X'=Z'$. Straightforward pre-multiplication of (2.6.28) by such a Z' will show this. Consequently, the idempotency of $X(X'X)^-X'$ gives $Z'Z = Z'X(X'X)^-X'Z$ for such a Z and X. By (2.6.10), (p. 43), λ, in this case $\lambda_{\beta|\mu,b,\tau,\beta}$, equals zero for all β. By Theorem 3.7.6(a), (p. 193), the treatments of X will be linked, in each block, with the identical covariate values in $N(\mathfrak{S})t!$ randomizations.

Such randomizations will order the treatments so that $X_{T,s}$ aligns treatment 1 with the same value of the covariate in every block. And, aligns treatment 2 with a value of the covariate the same in every block and so on for all t treatments. There are $N(\mathcal{G})t!$ ways to do this. In such randomizations, the model X will perfectly predict Z_1 and (2.4.10), (p. 28), will give a zero value for R_Z, and hence $F(\beta|\mu,b,\tau)=0$. Under the normally distributed random error model, the observed significance level, OSL_F, is one for all values of the covariate slope coefficient β. The null hypothesis is never rejected; in fact, $\alpha=0$. Even if $\beta\neq0$, the null hypothesis is always accepted; thus, $\beta_F=1$ and the power, $1-\beta_F$, is zero for these randomizations.

This is what one would expect, as such a covariate, equal in each block, is completely confounded with the treatments, and the hypothesis test adjusts for the treatment. Cox [1982, p. 200, bottom] notes, without proof, that "when the [covariates] identify a system of blocking, there may exist a large number of arrangements with $||\bar{z}_1-\bar{z}_2||_{\Omega_z}=0$ and an exact second moment randomization theory may hold." When the randomizations align the same covariates with the same treatments in each block, the covariates identify a system of blocking. In such randomizations, for the one factor case, Cox's "large number" $= N(\mathcal{G})t!$.

If the covariates in each block are similar instead of exactly equal, the $N(\mathbb{G})t!$ randomizations which align the treatments and similar covariates identically in each block will give the test near zero power. For the same covariate, other randomizations will thoroughly mix the covariates assigned to each treatment. In this case, the model $\mathbb{X}$ will poorly predict $\mathbb{Z}_1$. The absolute value of the difference $\mathbb{Z}_1 - \hat{\mathbb{Z}}_1 = \mathbb{R}_z$ will be large. Since the numerator of the F-ratio is a positive definite quadratic form, $F(\beta|\mu,b,\tau,\beta) > 0$ for such randomizations.

If the covariates take similar values in each block the restriction on randomization of (3.2.13), (p. 116), tends to repeat a similar pattern of covariate values in each block. With such a restriction, different randomizations will greatly differ in their power to detect $\beta \neq \mathbb{0}$. Some will approximate the extreme case described above and have near-zero power.

3.6.7   A Requirement of the Test Statistic

As Kempthorne [1952, p. 128-129] points out, the test statistic "must be such that deviations from the null hypothesis tend to place the value of the [actual test statistic] in a distinctive set of possible values." For any randomization s, the test statistic $t_{@s,s} \equiv F_{@s,s}$ does this, as we now show.

Consider the case when the model equation contains normally distributed random error. By Assumption 9, the

distributions of the normally distributed random errors and the randomizations are independent. Consequently, the distribution of a function of the normally distributed random error alone is independent of the randomization distribution. That is, for randomization S=s,

$$E[F_{@s,S} \mid S=s]$$

$$= E[F_{@s,S}, \; S] \; / \; E_R[S] \quad = \quad E_F[F_{@s,S}] \; E_R[S] \; / \; E_R[S]$$

$$= \begin{cases} \dfrac{d}{d-2} \; [1+2\lambda/n] & \text{when the model contains normal random error,} \\[2ex] F_{@s,s} & \text{otherwise.} \end{cases}$$

$$(3.6.12)$$

See Appendix A.3 , (p. 384), for the mean of the F statistic.

When the model equation has normally distributed random error λ enters the expectation. Under the null hypothesis $\lambda=0$ and under any alternative hypothesis $\lambda>0$. Under any alternative hypothesis each $F_{@s,s}$ has expectation $2\lambda/[n(d-2)]$ greater than $F_{@s,s}$ under the null hypothesis. Thus, the entire distribution of the N(Ξ) test statistics $F_{@s,s}$ is shifted to the right, if the alternative hypothesis is true. This provides the required distinctive set Kempthorne [1952, p. 128-129] discusses.

When the model equation lacks normally distributed random error

$$Y_S = [(TX) + E + H]S_S$$

$$= (T_CX)_S + (T_TX)_S + E_S + H_S \; .$$

Consider the null hypothesis $H_{o;h,r}$ $T_h = 0$. The elements

of T hypothesized to be zero (h) are set to zero in T_o,

while in T_T, only those elements of T hypothesized as

nonzero (r) have nonzero symbolic values; $T = T_o + T_T$. Under

the null hypothesis, $T_T = 0$.

Letting $X(X'X)^- X' = A$, the quadratic form

$$R(h,r) = Y_s'[X(X'X)^- X']Y_s = Y_s'AY_s$$

$$= (T_o X)_s'A(T_o X)_s + (T_T X)_s'A(T_T X)_s + \text{quadratic}$$

and bilinear forms in the other terms

$$\geq (T_o X)_s'A(T_o X)_s + \quad 0 \quad + \quad \text{the same quadratic}$$

and bilinear forms in the other terms.

The inequality follows from $(T_T X)_s'A(T_T X)_s$ being a

positive definite quadratic form. When the alternative

hypothesis is true, T_T is nonzero. This increases the

value of $R(h,r)$, while leaving $R(r)$ unchanged, increasing

the numerator of the F-ratio, $R(h|r)$. The denominator is

reduced. Together, the F-ratio is increased. Thus,

under the alternative hypothesis, the $N(\Xi)$ values of the

$F_{@s,s}$ are shifted to higher values, again providing the

required distinctive set. When the model equation

contains normally distributed random the above two

arguments act together.

We next develop a method of obtaining, for all

randomizations, such values as the model equation

parameter estimates and analysis of variance terms.

Section 3.7

Randomization, Anova Terms and Estimators

Section 3.7.1 develops a matrix method which will produce the design matrix unique to each randomization. Section 3.7.2 applies this method to projectors. Section 3.7.3 that shows these projectors provide a clustering to the distribution of some model equation estimators and analysis of variance terms. Section 3.7.4 shows the clustering of the noncentrality parameter associated with $F(\beta|\mu,b,\tau)$. The final section, 3.7.5, develops, for the single covariate case, a relationship between the noncentrality parameters associated with $F(\beta|\mu,b,\tau)$ and $F(\tau|\mu,b,\beta)$. This relationship allows one to improve the power of both usual hypothesis tests by restricting randomization. The main results are Theorems 3.7.9, (p. 202), and 3.7.15, (p. 219).

3.7.1 Permutation Matrices

The design matrices for two randomizations, s and s', differ only in the columns representing the treatments, the columns of $X_{s,\tau}$. Holding constant from one randomization to the next the order of the experimental units, that is, the rows of X and Z, makes the matrices X_μ, X_b, and Z constant for all randomizations. Within a level of the blocking variable, the rows of X_μ and X_b, are constant. Thus, within a block, one may interchange the rows of X to obtain the configuration of zeros and plus

or minus ones in $\mathbb{X}_\tau$ corresponding to randomization s.
Designate an arbitrary design matrix $\mathbb{X}_{s=1}$ as the one for
randomization number one, s=1. Premultiplication of $\mathbb{X}_{s=1}$
by the proper permutation matrix, $\mathbb{M}_s$, interchanges the
rows of $\mathbb{X}_{s=1}$ to yield the design matrix for randomization
s, $\mathbb{X}_s$. An illustration for a single block follows,

$$
\mathbb{M}_{s,i=1}\ \mathbb{X}_{s=1}\ =
\begin{bmatrix}
0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
1 & | & 1 & | & 1 & 0 & 0 \\
1 & | & 1 & | & 0 & 1 & 0 \\
1 & | & 1 & | & 0 & 0 & 1 \\
1 & | & 1 & | & -1 & -1 & -1
\end{bmatrix}
$$

$$
=\ [\mathbb{X}_\mu | \mathbb{X}_{i=1} | \mathbb{X}_{s,\tau}]\ =
\begin{bmatrix}
1 & | & 1 & | & 0 & 0 & 1 \\
1 & | & 1 & | & 1 & 0 & 0 \\
1 & | & 1 & | & 0 & 1 & 0 \\
1 & | & 1 & | & -1 & -1 & -1
\end{bmatrix}
=\ \mathbb{X}_s .
$$

Effect of a Permutation Matrix – One Block

Example 3.7.1

The full permutation matrix, $\mathbb{M}_s$, is block diagonal,
with one block per level of the blocking factor. All
elements of off-diagonal blocks of $\mathbb{M}_s$ are zero. Each
block submatrix, $\mathbb{M}_{s,i}$, is rt by rt; the entire matrix,
$\mathbb{M}_s$, is brt by brt. As in Example 3.7.1, each row and
each column of the block submatrix, $\mathbb{M}_{s,i}$, has a single
element with the value one; all other elements are zero.

Example 3.7.2 illustrates the full permutation matrix, $\mathbb{M}_s$, for b=2, r=1, t=4 and randomization number s=24. Note how block two of $\mathbb{X}_{s=1}$ changes to block two of $\mathbb{X}_{s=24}$.

$$\mathbb{M}_{s=24}\,\mathbb{X}_{s=1} = \begin{bmatrix} \mathbb{M}_{s=1,\,i=1} & \vline & \text{zero} \\ \hline \text{zero} & \vline & \mathbb{M}_{s=24,\,i=2} \end{bmatrix} \mathbb{X}_{s=1}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 & \vline & & & \\ 0 & 1 & 0 & 0 & \vline & & & \\ 0 & 0 & 1 & 0 & \vline & & \text{zero} & \\ 0 & 0 & 0 & 1 & \vline & & & \\ \hline & & & & \vline & 0 & 0 & 0 & 1 \\ & \text{zero} & & & \vline & 0 & 0 & 1 & 0 \\ & & & & \vline & 0 & 1 & 0 & 0 \\ & & & & \vline & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & \vline & 1 & \vline & 1 & 0 & 0 \\ 1 & \vline & 1 & \vline & 0 & 1 & 0 \\ 1 & \vline & 1 & \vline & 0 & 0 & 1 \\ 1 & \vline & 1 & \vline & -1 & -1 & -1 \\ \hline 1 & \vline & -1 & \vline & 1 & 0 & 0 \\ 1 & \vline & -1 & \vline & 0 & 1 & 0 \\ 1 & \vline & -1 & \vline & 0 & 0 & 1 \\ 1 & \vline & -1 & \vline & -1 & -1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \vline & 1 & \vline & 1 & 0 & 0 \\ 1 & \vline & 1 & \vline & 0 & 1 & 0 \\ 1 & \vline & 1 & \vline & 0 & 0 & 1 \\ 1 & \vline & 1 & \vline & -1 & -1 & -1 \\ \hline 1 & \vline & -1 & \vline & -1 & -1 & -1 \\ 1 & \vline & -1 & \vline & 0 & 0 & 1 \\ 1 & \vline & -1 & \vline & 0 & 1 & 0 \\ 1 & \vline & -1 & \vline & 1 & 0 & 0 \end{bmatrix} = \mathbb{X}_{s=24} = [\mathbb{X}_\mu | \mathbb{X}_b | \mathbb{X}_{s=24,\,\tau}]$$

(3.7.1)

Effect of A Permutation Matrix - Two Blocks

Example 3.7.2

The block diagonal structure of $\mathbb{M}_s$ allows each block sub-matrix to switch only the rows of one block of the design matrix, $\mathbb{X}_s$. The label s is the block-wise randomization in $\mathbb{M}_{s,i}$, and the experiment-wise randomization in $\mathbb{M}_s$. In Example 3.7.2 the block-wise randomizations are 1 and 24; the experiment-wise randomization is 24. For brevity, we shorten $\mathbb{X}_{s=1}$ to $\mathbb{X}_1$.

3.7.2 Permutation and Projection Matrices

There are $N(\mathcal{E})$ permutation matrices forming the set

$$\mathfrak{M} = \{\mathbb{M}_1 = \mathbb{I}, \ \mathbb{M}_2, \ \cdots, \ \mathbb{M}_s, \ \cdots, \ \mathbb{M}_{N(\mathcal{E})}\} . \qquad (3.7.2)$$

Each permutation matrix, except the first, is the product of one or more type one elementary matrices, the "row switcher" matrix, $\mathbb{R}$, times the identity matrix, $\mathbb{I}$. Recall that such elementary matrices are orthogonal; thus $\mathbb{R}^{-1}$ exists and equals $\mathbb{R}'$. The product of two row switcher matrices is also a row switcher matrix, but not necessarily an elementary one, as it may switch more than one row. We prove the following theorems.

Theorem 3.7.1.

$\mathbb{M}_s$ is orthogonal, that is, $\mathbb{M}_s' \mathbb{M}_s = \mathbb{M}_s^{-1} \mathbb{M}_s = \mathbb{M}_s \mathbb{M}_s' = \mathbb{M}_s \mathbb{M}_s^{-1} = \mathbb{I}$.

Proof: For $m \geq 1$ and $\mathbb{R}$ a type one elementary matrix or $\mathbb{I}$,

$$\mathbb{M}_{s,i}' = (\mathbb{R}_1 \ \mathbb{R}_2 \ \cdots \mathbb{R}_{m-1} \ \mathbb{R}_m)'$$

$$= \mathbb{R}_m' \ \mathbb{R}_{m-1}' \ \cdots \ \mathbb{R}_2' \ \mathbb{R}_1'$$

$$= \mathbb{R}_m^{-1} \mathbb{R}_{m-1}^{-1} \ \cdots \ \mathbb{R}_2^{-1} \ \mathbb{R}_1^{-1}$$

$$= (\mathbb{R}_1 \ \mathbb{R}_2 \ \cdots \ \mathbb{R}_{m-1} \ \mathbb{R}_m)^{-1}$$

$$= (\mathbb{M}_{s,i})^{-1} \qquad (3.7.3)$$

Therefore, by definition, $\mathbb{M}_{s,i}$ is orthogonal. As this holds for all blocks, $\mathbb{M}_s$ is orthogonal.

Theorem 3.7.2.

    When the set $\mathbb{G}$ is unrestricted, the set $\mathfrak{m}$, (3.7.2),

and the operation matrix multiplication form a group.

The set of $\mathbb{M}_{s,i}$ corresponding to $\mathfrak{m}$ and matrix

multiplication likewise form a group.

Proof: By writing $\mathbb{M}_s$ as a product of elementary row

switcher matrices, one may directly verify the conditions

defining the pair ($\mathfrak{m}$, matrix multiplication) as a group.

One list of these conditions is

    a) $\mathfrak{m}$ is closed under matrix multiplication,

    b) matrix multiplication is associative,

    c) matrix multiplication has an identity element, $\mathbb{I}$,

and d) each element of $\mathfrak{m}$, $\mathbb{M}_s$, has an inverse under the

       operation of matrix multiplication, $\mathbb{M}_s'$, which is

       also an element of $\mathfrak{m}$.

See, for example, Burton [1971, p. 48-53 and 69-74].

    Conditions (b) and (c) follow from matrix multipli-

cation. Condition (d) follows from (i) the transpose of

each $\mathbb{M}_s$ has the proper form and thus is a member of $\mathfrak{m}$,

and (ii) by Theorem 3.7.1 the transpose is the inverse.

Condition (a) follows from the fact that the product of

any row of one $\mathbb{M}_s$ times the columns of another $\mathbb{M}_{s'}$ will

yield exactly one element of that row of the product,

$\mathbb{M}_s\mathbb{M}_{s'}$, with a nonzero value. This one nonzero value must

have one as its value. This holds for all rows of the

product. The value one is in a different column in each
row of the premultiplier; consequently, the value one
will be in a different column of the product $M_S M_{S'}$.
Thus, the product will have the form of a permutation
matrix and be a member of $\mathfrak{M}$. One may fix the labels
$1,2,\ldots,N(\mathcal{E})$ to the elements of $\mathfrak{M}$ in an arbitrary manner.
This proves Theorem 3.7.2.

Restricted randomization destroys the group charac-
teristic of the pair ($\mathfrak{M}$, matrix multiplication). The
product of two permutation matrices, each included among
the restricted set, would not necessarily be among the
restricted set, violating condition (a) of Theorem 3.7.2.
One may preserve the group nature of the pair ($\mathfrak{M}$, matrix
multiplication) by including all randomizations and then
assigning a zero value to the experiment-wise randomiza-
tion selection probability, $p(s)$, corresponding to those
randomizations which would have been restricted. Herein
$N(\mathcal{E})$ is the number of all possible randomizations.

Theorem 3.7.3.

$[X_S'X_S]$ and $[X_S'X_S]^-$ are constant for all
randomizations.

When needed, the same generalized inverse is assumed for
all randomizations.

Proof: $[\Sigma_S{}'\Sigma_S] = [(\mathbb{M}_S\Sigma_1)'(\mathbb{M}_S\Sigma_1)] = [\Sigma_1{}'(\mathbb{M}_S{}'\mathbb{M}_S)\Sigma_1]$

$$= [\Sigma_1{}'(\mathbb{I})\Sigma_1], \quad \text{by Theorem 3.7.1.} \qquad (3.7.4)$$

The assumed uniqueness of the inverse provides the same

result for $[\Sigma_S{}'\Sigma_S]^-$.

Theorem 3.7.4.

$$[\Sigma_S(\Sigma_S{}'\Sigma_S)^-\Sigma_S{}'] = \mathbb{M}_S \, [\Sigma_1(\Sigma_1{}'\Sigma_1)^-\Sigma_1{}'] \, \mathbb{M}_S{}' \; .$$

Proof: By Theorem 3.7.3 and that $\Sigma_S = \mathbb{M}_S\Sigma_1$,

$$[\Sigma_S(\Sigma_S{}'\Sigma_S)^-\Sigma_S{}'] = (\mathbb{M}_S\Sigma_1)(\Sigma_1{}'\Sigma_1)^-(\mathbb{M}_S\Sigma_1)'$$

$$= \mathbb{M}_S \, [\Sigma_1(\Sigma_1{}'\Sigma_1)^-\Sigma_1{}'] \, \mathbb{M}_S{}' \; . \qquad (3.7.5)$$

Theorem 3.7.5.

For projectors $[\Sigma_1(\Sigma_1{}'\Sigma_1)^-\Sigma_1{}']$, all of whose

diagonal blocks have, within each block, identical

diagonal elements and identical off-diagonal elements,

the diagonal blocks of $[\Sigma_S(\Sigma_S{}'\Sigma_S)^-\Sigma_S{}']$, or equivalently

of $\mathbb{M}_S[\Sigma_1(\Sigma_1{}'\Sigma_1)^-\Sigma_1{}']\mathbb{M}_S{}'$, are constant for all $\mathbb{M}_S$ in $\mathbb{m}$,

i.e., for all randomizations. $\qquad (3.7.6)$

Proof: Each diagonal block of $\Sigma_1(\Sigma_1{}'\Sigma_1)^-\Sigma_1{}'$ is pre- and

post-multiplied by the same block of $\mathbb{M}_S$, transposed for

the post-multiplication. As seen in Example 3.7.3 below,

for a specific $\mathbb{M}_S$, pre-multiplication by $\mathbb{M}_S$ moves a given

diagonal element vertically w rows. Post-multiplication

by the transpose, $\mathbb{M}_s'$, moves the element horizontally w

columns, returning it to the diagonal, albeit in a new

position. This applies to all diagonal elements and

holds for all $\mathbb{M}_s$. Thus, all off-diagonal elements remain

in off-diagonal locations. By assumption, all diagonal

elements have the same value; thus the resulting diagonal

is identical to the original diagonal. By assumption,

all off-diagonal elements have the same value; thus the

resulting off-diagonal elements are identical to the

original off-diagonal elements, albeit also relocated.

This proves Theorem 3.7.5.

Section 2.6.3, below (2.6.28),(p. 55), discusses

this structure in a projector. The projectors for

$[\Sigma_\mu | \Sigma_b]$, (2.6.20), (p. 50); the effects version of Σ_T,

(2.6.24), (p. 52); and $[\Sigma_\mu | \Sigma_b | \Sigma_T]$, (2.6.28), (p. 55)

all have this structure. The projectors for matrices

related to these three matrices by the conditions of

Theorem 2.8.3, (p. 87), also have this structure. Note

that the coding version of Σ_T is not so related to the

effects version of Σ_T; thus, its projector does not have

this structure, nor does the projector for Σ_T with more

than one factor. Consideration of the effects of

randomization upon projectors derived from Σ_T with

multiple factors and interactions is an extension.

While $\Sigma_1(\Sigma_1'\Sigma_1)^-\Sigma_1'$ is symmetric, for clarity

Example 3.7.3 uses distinct elements. Each element is

its original row, column location. All elements remain

in their original blocks.

$$M_s[\Sigma_1(\Sigma_1'\Sigma_1)^-\Sigma_1']M_s'$$

$$=
\begin{bmatrix}
0 & 0 & 1 & 0 & & & & \\
1 & 0 & 0 & 0 & & \text{zero} & & \\
0 & 1 & 0 & 0 & & & & \\
0 & 0 & 0 & 1 & & & & \\
\hline
 & & & & 0 & 0 & 1 & 0 \\
 & \text{zero} & & & 0 & 0 & 0 & 1 \\
 & & & & 0 & 1 & 0 & 0 \\
 & & & & 1 & 0 & 0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 \\
21 & 22 & 23 & 24 & 25 & 26 & 27 & 28 \\
31 & 32 & 33 & 34 & 35 & 36 & 37 & 38 \\
41 & 42 & 43 & 44 & 45 & 46 & 47 & 48 \\
\hline
51 & 52 & 53 & 54 & 55 & 56 & 57 & 58 \\
61 & 62 & 63 & 64 & 65 & 66 & 67 & 68 \\
71 & 72 & 73 & 74 & 75 & 76 & 77 & 78 \\
81 & 82 & 83 & 84 & 85 & 86 & 87 & 88 \\
\end{bmatrix}$$

$$\times
\begin{bmatrix}
0 & 1 & 0 & 0 & & & & \\
0 & 0 & 1 & 0 & & \text{zero} & & \\
1 & 0 & 0 & 0 & & & & \\
0 & 0 & 0 & 1 & & & & \\
\hline
 & & & & 0 & 0 & 0 & 1 \\
 & \text{zero} & & & 0 & 0 & 1 & 0 \\
 & & & & 1 & 0 & 0 & 0 \\
 & & & & 0 & 1 & 0 & 0 \\
\end{bmatrix}$$

$$=
\begin{bmatrix}
31 & 32 & 33 & 34 & 35 & 36 & 37 & 38 \\
11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 \\
21 & 22 & 23 & 24 & 25 & 26 & 27 & 28 \\
41 & 42 & 43 & 44 & 45 & 46 & 47 & 48 \\
\hline
71 & 72 & 73 & 74 & 75 & 76 & 77 & 78 \\
81 & 82 & 83 & 84 & 85 & 86 & 87 & 88 \\
61 & 62 & 63 & 64 & 65 & 66 & 67 & 68 \\
51 & 52 & 53 & 54 & 55 & 56 & 57 & 58 \\
\end{bmatrix}
\begin{bmatrix}
0 & 1 & 0 & 0 & & & & \\
0 & 0 & 1 & 0 & & \text{zero} & & \\
1 & 0 & 0 & 0 & & & & \\
0 & 0 & 0 & 1 & & & & \\
\hline
 & & & & 0 & 0 & 0 & 1 \\
 & \text{zero} & & & 0 & 0 & 1 & 0 \\
 & & & & 1 & 0 & 0 & 0 \\
 & & & & 0 & 1 & 0 & 0 \\
\end{bmatrix}$$

$$=
\begin{bmatrix}
33 & 31 & 32 & 34 & 37 & 38 & 36 & 35 \\
13 & 11 & 12 & 14 & 17 & 18 & 16 & 15 \\
23 & 21 & 22 & 24 & 27 & 28 & 26 & 25 \\
43 & 41 & 42 & 44 & 47 & 48 & 46 & 45 \\
\hline
73 & 71 & 72 & 74 & 77 & 78 & 76 & 75 \\
83 & 81 & 82 & 84 & 87 & 88 & 86 & 85 \\
63 & 61 & 62 & 64 & 67 & 68 & 66 & 65 \\
53 & 51 & 52 & 54 & 57 & 58 & 56 & 55 \\
\end{bmatrix}
= \Sigma_s(\Sigma_s'\Sigma_s)^-\Sigma_s' \ .$$

Effect of a Permutation Matrix - Projector Matrix

Example 3.7.3

The diagonal elements of the off-diagonal blocks of $X_1(X_1'X_1)^-X_1'$ are not returned to the diagonal. When the off-diagonal blocks have elements with different values in their on- and off-diagonal positions, randomization alters the projector, as with the projectors for X_T (effects version) (2.6.24), (p. 52), and for $[X_\mu|X_b|X_T]$, (2.6.28), (p. 55). When they are the same, as in the projector for $[X_\mu|X_b]$, (2.6.20), (p. 50), randomization does not alter the projector. This is as one would expect; randomization occurs within the block, affecting only the treatments, not the blocks.

Theorem 3.7.6.

For projectors $X_S(X_S'X_S)^-X_S'$, all of whose diagonal blocks have, within each block, identical diagonal elements and identical off-diagonal elements (as in Theorem 3.7.5), and for a number of randomizations equal to $[N(\mathfrak{G})t!]^b$, there are

a) exactly $[N(\mathfrak{G})t!]^{b-1}$ groups of randomizations, each comprised of exactly $[N(\mathfrak{G})t!]$ randomizations. Within each group all projectors $X_S(X_S'X_S)^-X_S'$ are identical.

b) at most $[N(\mathfrak{G})t!]^{b-1}$ groups of randomizations, each having an integer multiple (1, 2, ...) of $[N(\mathfrak{G})t!]$ randomizations. That is, having $1[N(\mathfrak{G})t!]$ or $2[N(\mathfrak{G})t!]$ or $3[N(\mathfrak{G})t!]$, etc. number of randomizations. Within each

of these groups all $\mathbb{X}_S(\mathbb{X}_S{}'\mathbb{X}_S)^-\mathbb{X}_S{}'\mathbb{Z}$ are identical.

Consequently, within each group, all values are identical

for $\mathbb{R}_{S,z}$, $\mathbb{R}_{S,z}(\mathbb{R}_{S,z}{}'\mathbb{R}_{S,z})^-\mathbb{R}_{S,z}{}'$ and $(\mathbb{R}_{S,z}{}'\mathbb{R}_{S,z})^-\mathbb{R}_{S,z}{}'$.

Proof of Part (a):

By Theorem 3.7.2, (p. 188), each of the following
b-1 products is a member of $\mathfrak{M}$

$$\mathbb{M}_{S,1}(\mathbb{M}_{S,2})', \quad \mathbb{M}_{S,1}(\mathbb{M}_{S,3})', \quad \ldots, \quad \mathbb{M}_{S,1}(\mathbb{M}_{S,b})'. \qquad (3.7.7)$$

Replacing each product by the label for the member of $\mathfrak{M}$
which the product equals, yields an ordered sequence of
labels,

$$t_1, \; t_2, \; \ldots, \; t_{b-1}. \qquad (3.7.8)$$

The identical sequence of labels may be obtained for any
of the other $[N(\mathfrak{s})t!]-1$ choices of $\mathbb{M}_{S,1}$ by replacing

$$(\mathbb{M}_{S,2})', \quad (\mathbb{M}_{S,3})', \ldots, \quad (\mathbb{M}_{S,b})'$$

with the appropriate

$$(\mathbb{M}_{S*,2})', \quad (\mathbb{M}_{S*,3})', \ldots, \quad (\mathbb{M}_{S*,b})'.$$

For each block, s and s* are two necessarily different
block-wise randomizations. The group nature of $(\mathfrak{M},$
matrix multiplication) ensures that this can be done.
There are $\{[N(\mathfrak{s})t!]^b\}/\{[N(\mathfrak{s})t!]\} = [N(\mathfrak{s})t!]^{b-1}$ such
sequences of labels, (3.7.8). Label each unique sequence
of labels of (3.7.8) by g,

$$g = 1, \; 2, \; \ldots, \; [N(\mathfrak{s})t!]^{b-1}. \qquad (3.7.9)$$

Designate as $\mathbb{F}_{1,1}$, $\mathbb{F}_{1,2}$, $\ldots$, $\mathbb{F}_{1,b}$, the first row of

b blocks of $[\mathbb{X}_1(\mathbb{X}_1'\mathbb{X}_1)^-\mathbb{X}_1']$. The first row of blocks of

the product $\mathbb{M}_s[\mathbb{X}_1(\mathbb{X}_1'\mathbb{X}_1)^-\mathbb{X}_1']\mathbb{M}_s'$ is

$$\mathbb{M}_{s,1}\mathbb{F}_{1,1},(\mathbb{M}_{s,1})', \quad \mathbb{M}_{s,1}\mathbb{F}_{1,2},(\mathbb{M}_{s,2})',$$

$$\ldots, \quad \mathbb{M}_{s,1}\mathbb{F}_{1,b},(\mathbb{M}_{s,b})'. \quad (3.7.10)$$

The permutation matrices $\mathbb{M}_{s,1}$ and $(\mathbb{M}_{s,2})'$ only switch,

not multiply, the rows and columns of $\mathbb{F}_{1,2}$. Thus, the

products $\mathbb{M}_{s,1}\mathbb{F}_{1,2},(\mathbb{M}_{s,2})'$ and $\mathbb{M}_{s*,1}\mathbb{F}_{1,2},(\mathbb{M}_{s*,2})'$ are

equal since the products $\mathbb{M}_{s,1}(\mathbb{M}_{s,2})'$ and $\mathbb{M}_{s*,1}(\mathbb{M}_{s*,2})'$

are equal. Being equal, they have the same t label of

(3.7.8). This holds for all b blocks of row one and

for all $[N(\mathfrak{S})t!]^{b-1}$ sequences (3.7.7) with identical

sequences of t labels (3.7.8). Thus, all randomizations

with the same value of g (3.7.9) have the identical first

row of blocks of $\mathbb{M}_s[\mathbb{X}_1(\mathbb{X}_1'\mathbb{X}_1)^-\mathbb{X}_1']\mathbb{M}_s'$.

We now show that any row of blocks of one

$\mathbb{X}_s(\mathbb{X}_s'\mathbb{X}_s)^-\mathbb{X}_s'$ is identical to the same row of any another

$\mathbb{X}_{s*}(\mathbb{X}_{s*}'\mathbb{X}_{s*})^-\mathbb{X}_{s*}'$, provided s and s* are in the same

g-group. Without loss of generality, consider the second

row of blocks. Two sequences of products are

$$\mathbb{M}_{s,2}(\mathbb{M}_{s,1})', \quad \mathbb{I}, \quad \mathbb{M}_{s,2}(\mathbb{M}_{s,3})', \quad \ldots, \quad \mathbb{M}_{s,2}(\mathbb{M}_{s,b})' \text{ and}$$

$$\mathbb{M}_{s*,2}(\mathbb{M}_{s*,1})', \quad \mathbb{I}, \quad \mathbb{M}_{s*,2}(\mathbb{M}_{s*,3})', \ldots, \mathbb{M}_{s*,2}(\mathbb{M}_{s*,b})'.$$

The assumption that g is the same for both sequences gives

$$M_{s,1}(M_{s,2})' = M_{s*,1}(M_{s*,2})'. \qquad (3.7.11)$$

Premultiplication and Theorem 3.7.1, (p. 187), give

$$[M_{s,1} \ (M_{s,2})'] \ [\ M_{s,2} \ (M_{s,3})']$$

$$= \ M_{s,1} \ I \ (M_{s,3})' \ \text{and} \qquad (3.7.12)$$

$$[M_{s*,1}(M_{s*,2})'] \ [\ M_{s*,2}(M_{s*,3})']$$

$$= \ M_{s*,1} \ I \ (M_{s*,3})' \ . \qquad (3.7.13)$$

By the assumption of equal values of g, the right-hand side of (3.7.12) equals the right-hand side of (3.7.13). Thus,

$$[M_{s,1} \ (M_{s,2})'] \ [\ M_{s,2} \ (M_{s,3})']$$

$$= [M_{s*,1}(M_{s*,2})'] \ [\ M_{s*,2}(M_{s*,3})']. \qquad (3.7.14)$$

Premultiplication by the transpose of $[M_{s,1}(M_{s,2})']$,

which by (3.7.11) equals the transpose of $M_{s*,1}(M_{s*,2})'$,

gives

$$M_{s,2}(M_{s,3})' = M_{s*,2}(M_{s*,3})'. \qquad (3.7.15)$$

Thus, the third block of row two of the first projector matrix (s) equals the same block of the second projector matrix (s*).

This same argument applies to all blocks of row two and all rows of blocks. Thus, all randomizations with the same value of g (3.7.9) have identical values of $M_s[X_1(X_1'X_1)^-X_1'](M_s)'$. This proves Part (a).

The permutation matrix of Example 3.7.3, (p. 192), has block-wise randomizations labeled 13 and 18. The experiment-wise randomization is s =(13-1){[N($\mathbb{G}$)t!]} + 18 =(12)(24) + 18 = 306, while g=6. The permutation matrix with block-wise randomizations 1 and 6, has experiment-wise randomization of 6, also with g=6. Thus, randomizations s=306 and s=6 yield identical projector matrices, when the initial projector matrix has the special form discussed in Theorem 3.7.5. We use the term g-group for the groups of randomizations with the same value of g in (3.7.9), (p. 194).

For the $\mathbb{X}_S$ of Example 3.7.2, (p. 186), t=4, b=2, and r=1, thus N($\mathbb{G}$)=1, giving [N($\mathbb{G}$)t!]$^{b-1}$ = [(1)(4!)]$^{2-1}$ = 24 possible values for the projector $\mathbb{X}_S(\mathbb{X}_S'\mathbb{X}_S)^-\mathbb{X}_S'$, not (4!)^2 = 576, the total number of randomizations.

Proof of Part (b) of Theorem 3.7.6:

By definition $\mathbb{R}_{S,z} = \mathbb{R}_{S,x}\mathbb{Z} = [\mathbb{I} - \mathbb{X}_S(\mathbb{X}_S'\mathbb{X}_S)^-\mathbb{X}_S']\mathbb{Z}$, (2.4.11), (p. 28). Thus, randomizations with equal $\mathbb{X}_S(\mathbb{X}_S'\mathbb{X}_S)^-\mathbb{X}_S'$ also have equal $\mathbb{R}_{S,z}$. By part (a) each equal-valued $\mathbb{R}_{S,z}$ appears in [N($\mathbb{G}$)t!] randomizations, as do $\mathbb{R}_{S,z}(\mathbb{R}_{S,z}'\mathbb{R}_{S,z})^-\mathbb{R}_{S,z}'$ and $(\mathbb{R}_{S,z}'\mathbb{R}_{S,z})^-\mathbb{R}_{S,z}'$. However, each column of $[\mathbb{X}_S(\mathbb{X}_S'\mathbb{X}_S)^-\mathbb{X}_S']\mathbb{Z}$ is a weighted sum of the elements of a column of $\mathbb{Z}$. For some $\mathbb{Z}$, two or more of the g-groups will yield the same sum. The covariate of

Table 2.7.1, (p. 69), does this. For that Z there are

$$[N(\mathbb{G})t!]^{b-1} - 1 = [(1)(4!)]^{2-1} - 1 = 24-1 = 23$$

randomizations groups. One value of the projector

$\mathbb{R}_{s,z}(\mathbb{R}_{s,z}{}'\mathbb{R}_{s,z})^{-}\mathbb{R}_{s,z}{}'$ appears $(2)[N(\mathbb{G})t!] = (2)[24] = 48$

times. This proves Part (b) of Theorem 3.7.6.

The functions of Part (b) are grouped, but these

groups may be fewer than $[N(\mathbb{G})t!]^{b-1}$ in number and

contain an integer multiple of $[N(\mathbb{G})t!]$ randomizations.

Equality of the $X_s(X_s{}'X_s)^{-}X_s{}'$ implies equality of the

$\mathbb{R}_{s,z}(\mathbb{R}_{s,z}{}'\mathbb{R}_{s,z})^{-}\mathbb{R}_{s,z}{}'$ for randomizations within the same

g-group, but not the converse. Such duplication is

unlikely with multiple covariates, but will occur when

multiple observations have the same covariate values.

Possibly this result can be exploited in the analysis of

covariance analog of the measurement of pure error.

3.7.3 Permutation, Anova Terms and an Estimator

The repetition in multiple randomizations of

identical values of $X_s(X_s{}'X_s)^{-}X_s{}'$ and functions with Z

(Theorem 3.7.6) imparts a cluster pattern to the values

of those Anova terms in which they appear. Theorem 3.7.9

presents the clustering and dispersing parts of these

terms.

The clustering term involves $(Z\beta + e)$, where e is

the nx1 vector of experimental unit error. Alone, it is

constant for all randomizations. The dispersing term

involves ($\eta_S + n_S$), the sum of the nx1 vector for

experimental unit-treatment interaction errors (η_S) plus

the nx1 vector of normally distributed random errors

(n_S).   Typically, both of these vectors differ with each

randomization.   In each Anova term, ($Z\beta + e$) is multi-

plied by one or more projectors which are constant for

all randomizations in the same randomization group.   The

product fixes a cluster center for the values of the

Anova term for those randomizations in the group.   Within

the group, the constant projector multiplies ($\eta_S + n_S$).

This product disperses the values about the cluster

center, providing the η_S and/or n_S change with the

randomizations, as is the typical case.

The relative magnitude of the elements of ($Z\beta + e$)

and ($\eta_S + n_S$) determine the degree of intermingling of

values from different clusters, $i.e.$, from different

randomization groups.   When β is near zero, the elements

of ($Z\beta + e$) may be nearly equal to those of ($\eta_S + n_S$).

In this case, the clusters merge one into another.

Typically, $Z\beta \neq 0$ and thus, the elements of ($Z\beta + e$) are

large relative to those of ($\eta_S + n_S$).   In this case,

clusters are distinct with little intermingling of

values.   Under strict additivity (3.3.11), (p. 125),

$e \neq 0$, ($\eta_S + n_S$) = 0 and there is no dispersion; each

cluster collapses to a single value identical for all

[N($\mathfrak{S}$)t!] randomizations in the group.   The randomization

test (Section 3.6.3, p. 169), holds constant the value

of Υ_s, treating n_s and n_s as negligible for all

randomizations (Section 3.6.2, (p. 170)). Again, each

cluster collapses to a single value. When the clusters

are distinct, a continuous distribution is a poor approx-

imation to this discrete randomization distribution.

The clustering imparts a stair step shape to the

cumulative distribution function when $p(s)$ is constant

for all randomizations, s. As clusters become less

distinct, the steps collapse, making the cumulative

distribution function more like the sigmoid curve of a

continuous density function, albeit discrete.

There are at most brt^2 distinct elements of n_s, one

for each experimental unit (with r replications for t

treatments in b blocks) times one per treatment. They

are reselected to form each of the $[N(s)t!]^b$ n_s for all

randomizations. There are $(brt)[N(s)t!]^b$ distinct

elements of the $[N(s)t!]^b$ n_s. An extension is the

examination of the different dispersion effects, if any,

of the two error terms n_s and n_s. Note that should n_s be

constant for all randomizations, it would have the same

effect as e. Hence, the assumption here and in Section

3.5.2, (p. 155), that the columns of N of (3.5.1),

(p. 156), the n_s, differ with each randomization.

The main result follows two preliminary theorems.

<u>Theorem 3.7.7.</u>

$$\mathbb{Z}'\{\mathbb{R}_{S,z}(\mathbb{R}_{S,z}'\mathbb{R}_{S,z})^-\mathbb{R}_{S,z}'\}$$

$$= \mathbb{Z}'\{\mathbb{I} - \mathbb{X}_S(\mathbb{X}_S'\mathbb{X}_S)^-\mathbb{X}_S'\} \qquad (3.7.16)$$

Proof: By the definition of $\mathbb{R}_{S,z}$ (2.4.7), (p. 26),

and the idempotency and symmetry of $\mathbb{R}_{S,x}$,

$$\mathbb{Z}'\{\mathbb{R}_{S,z}(\mathbb{R}_{S,z}'\mathbb{R}_{S,z})^-\mathbb{R}_{S,z}'\}$$

$$= \mathbb{Z}'\mathbb{R}_{S,x}\mathbb{Z}[(\mathbb{R}_{S,x}\mathbb{Z})'(\mathbb{R}_{S,x}\mathbb{Z})]^-(\mathbb{R}_{S,x}\mathbb{Z})'$$

$$= (\mathbb{R}_{S,x}\mathbb{Z})'(\mathbb{R}_{S,x}\mathbb{Z})[(\mathbb{R}_{S,x}\mathbb{Z})'(\mathbb{R}_{S,x}\mathbb{Z})]^-(\mathbb{R}_{S,x}\mathbb{Z})',$$

which, by the Corollary of p. 20 of Searle [1971] and the

definition of $\mathbb{R}_{S,z}$, becomes

$$= (\mathbb{R}_{S,x}\mathbb{Z})' = \mathbb{Z}'\{\mathbb{I} - \mathbb{X}_S(\mathbb{X}_S'\mathbb{X}_S)^-\mathbb{X}_S'\}.$$

<u>Theorem 3.7.8.</u>

For full rank $\mathbb{R}_{S,z}$, $(\mathbb{R}_{S,z}'\mathbb{R}_{S,z})^-\mathbb{R}_{S,z}'\mathbb{Z} = \mathbb{I}$ (3.7.17)

Proof: By the definition of $\mathbb{R}_{S,z}$ (2.4.7), (p. 26),

$$(\mathbb{R}_{S,z}'\mathbb{R}_{S,z})^-\mathbb{R}_{S,z}'\mathbb{Z})$$

$$= [(\mathbb{R}_{S,x}\mathbb{Z})'(\mathbb{R}_{S,x}\mathbb{Z})]^-(\mathbb{R}_{S,x}\mathbb{Z})'\mathbb{Z},$$

which by the idempotency and symmetry of $\mathbb{R}_{S,x}$

$$= [(\mathbb{R}_{S,x}\mathbb{Z})'(\mathbb{R}_{S,x}\mathbb{Z})]^-(\mathbb{Z}'\mathbb{R}_{S,x}')(\mathbb{R}_{S,x}\mathbb{Z})$$

$$= [(\mathbb{R}_{S,x}\mathbb{Z})'(\mathbb{R}_{S,x}\mathbb{Z})]^-[(\mathbb{R}_{S,x}\mathbb{Z})'(\mathbb{R}_{S,x}\mathbb{Z})].$$

By Assumption 4, (p. 18), $\mathbb{R}_{S,z}$ has full column rank,

and thus, so does $[(\mathbb{R}_{S,x}\mathbb{Z})'(\mathbb{R}_{S,x}\mathbb{Z})]$. Consequently, the inverse is a true inverse, giving the result. See Section 2.4.1, (p. 25), for details on the rank of $\mathbb{R}_{S,z}$.

Theorem 3.7.9.

When (a) the design matrix $\mathbb{X}_S$ has the projector $[\mathbb{X}_S(\mathbb{X}_S'\mathbb{X}_S)^-\mathbb{X}_S']$, all of whose diagonal blocks have, within each block, identical diagonal elements and identical off-diagonal elements (as in Theorem 3.7.6(b)), and (b) for a number of randomizations equal to $[N(\mathbb{g})t!]^b$, there are at most $[N(\mathbb{g})t!]^{b-1}$ clusters of numerical values, with each cluster composed of an integer multiple of $[N(\mathbb{g})t!]$ randomizations for

1. The sum of squares for the covariate(s) as adjusted for the mean, blocks and treatments, $R_S(\beta|\mu,b,\tau)$,

2. The residual sum of squares, SSR_S,

3. The F-ratio, and hence the OSL_F, for the covariate(s) as adjusted for mean, blocks and treatments, $F_S(\beta|\mu,b,\tau)$,

4. The estimate for the regression slope coefficient, as adjusted for the mean, blocks and treatments, $\hat{\beta}_{S;\mu,g,\tau,\beta}$, when $\mathbb{Z}$ has full column rank as per Assumption 4, (p. 18),

5. The F-ratio, and hence the OSL_F, for the treatments as adjusted for the mean, blocks and covariate(s), $F_S(\tau|\mu,b,\beta)$, and the F-ratio (and OSL_F) for the entire model equation, when the null hypothesis of $\tau=0$ is true.

Graph 4.5.9, (p. 288), and the points marked "u" in Graphs 4.6.5 and 4.6.7, (p. 299-302), illustrate Part 3. Graphs 4.5.3 and 4.5.4, (p. 277-278), illustrate Part 4. Graphs 4.5.5, (p. 280), and the top left of Display 4.6.7, (p. 350), illustrate Part 5.

To begin the proof, write the vector of observations as

$$\mathbf{Y}_S = \mathbf{X}_S\tau + \mathbb{Z}\beta + e + \eta_S + n_S . \qquad (3.7.18)$$

Proof of Part 1.

$$R_S(\beta|\mu,b,\tau) \equiv \mathbf{Y}_S'\{\mathbb{R}_{S,z}(\mathbb{R}_{S,z}'\mathbb{R}_{S,z})^-\mathbb{R}_{S,z}'\}\mathbf{Y}_S$$

By (2.4.13b), (p. 194), all terms in $\mathbf{X}_S$ are zero, leaving

$$(\mathbb{Z}\beta + e)'\{\mathbb{R}_{S,z}(\mathbb{R}_{S,z}'\mathbb{R}_{S,z})^-\mathbb{R}_{S,z}'\}(\mathbb{Z}\beta + e)$$

$$+ 2(\mathbb{Z}\beta + e)'\{\mathbb{R}_{S,z}(\mathbb{R}_{S,z}'\mathbb{R}_{S,z})^-\mathbb{R}_{S,z}'\}(\eta_S + n_S)$$

$$+ (\eta_S + n_S)'\{\mathbb{R}_{S,z}(\mathbb{R}_{S,z}'\mathbb{R}_{S,z})^-\mathbb{R}_{S,z}'\}(\eta_S + n_S). \qquad (3.7.19)$$

By Theorem 3.7.6(b) the first term is constant for all $[N(\mathfrak{G})t!]$ randomizations having the same value of the g of (3.7.9), (p. 194). It sets at most $[N(\mathfrak{G})t!]^{b-1}$ cluster centers. Within each g-group, the other two terms of

(3.7.19) disperse values about each cluster center. When

$(\eta_s + n_s) = 0$ for all s, as in strict additivity, or is

negligible as in the randomization test, each of the at

most $[N(\mathbb{G})t!]^{b-1}$ clusters collapses to a single value.

Otherwise, the relative values of $(\mathbb{Z}\beta + e)$ and $(\eta_s + n_s)$

determine the extent of cluster dispersion and overlap.

Duplication of values for clusters and/or individual

randomizations may occur via the weighted sum of certain

$(\mathbb{Z}\beta + e)$ and/or $(\eta_s + n_s)$, in the manner of Theorem 3.7.6(b)

or by the addition operation of $(\mathbb{Z}\beta + e)$ and/or $(\eta_s + n_s)$.

Proof of Part 2:

From (2.5.2), (p. 36),

$$SSR_s = \mathbf{Y}_s'\{\mathbb{I} - \mathbf{X}_s(\mathbf{X}_s'\mathbf{X}_s)^{-}\mathbf{X}_s' - \mathbb{R}_{s,z}(\mathbb{R}_{s,z}'\mathbb{R}_{s,z})^{-}\mathbb{R}_{s,z}'\}\mathbf{Y}_s$$

$$= \{(\tau'\mathbf{X}_s' + \beta'\mathbb{Z}') + e' + (\eta_s' + n_s')\}$$

$$\times \{\mathbb{I} - \mathbf{X}_s(\mathbf{X}_s'\mathbf{X}_s)^{-}\mathbf{X}_s' - \mathbb{R}_{s,z}(\mathbb{R}_{s,z}'\mathbb{R}_{s,z})^{-}\mathbb{R}_{s,z}'\}\mathbf{Y}_s.$$

Since $\mathbb{I} - \mathbf{X}_s(\mathbf{X}_s'\mathbf{X}_s)^{-}\mathbf{X}_s' = \mathbb{R}_{s,x}$, terms involving $\mathbf{X}_s'$ are

zero, by (2.4.13(b)). By Theorem 3.7.7, terms involving

$\mathbb{Z}'$ cancel. The idempotency provides the same

cancellation on the right for $\mathbf{Y}_s$, leaving,

$$SSR_s = e'\{\mathbb{I} - \mathbf{X}_s(\mathbf{X}_s'\mathbf{X}_s)^{-}\mathbf{X}_s' - \mathbb{R}_{s,z}(\mathbb{R}_{s,z}'\mathbb{R}_{s,z})^{-}\mathbb{R}_{s,z}'\}e$$

$$+ 2e'\{\mathbb{I} - \mathbf{X}_s(\mathbf{X}_s'\mathbf{X}_s)^{-}\mathbf{X}_s' - \mathbb{R}_{s,z}(\mathbb{R}_{s,z}'\mathbb{R}_{s,z})^{-}\mathbb{R}_{s,z}'\}(\eta_s + n_s)$$

$$+ (\eta_s + n_s)\{\mathbb{I} - \mathbf{X}_s(\mathbf{X}_s'\mathbf{X}_s)^{-}\mathbf{X}_s' - \mathbb{R}_{s,z}(\mathbb{R}_{s,z}'\mathbb{R}_{s,z})^{-}\mathbb{R}_{s,z}'\}(\eta_s + n_s).$$

$$(3.7.20)$$

Values of the first term are constant for all $[N(\mathbb{G})t!]$ randomizations having the same value of g, as in (3.7.9). Within each g-group, the last two terms disperse values about the value set by the first term. When $(n_s+n_s)=\mathbb{O}$ for all s, as in strict additivity, or is negligible as in the randomization test, each of the $[N(\mathbb{G})t!]^{b-1}$ clusters collapses to a single value. Otherwise, the relative values of e and (n_s+n_s) determine the extent of cluster dispersion and overlap. As with $R_s(\mathbb{P}|\mu,b,\tau)$, clusters and/or individual randomizations may have identical values of SSR_s.

Proof of Part 3:

   By Parts 1 and 2,

$$F_s(\mathbb{P}|\mu,b,\tau) = \{ [1/q] / [1/(n-p-q)] \} \times$$

$$\frac{[(\mathbb{Z}\mathbb{P}+e)+(n_s+n_s)]'\{\mathbb{R}_{s,z}(\mathbb{R}_{s,z}'\mathbb{R}_{s,z})^-\mathbb{R}_{s,z}'\}[(\mathbb{Z}\mathbb{P}+e)+(n_s+n_s)]}{[e+n_s+n_s]'\{\mathbb{I}-\mathbb{X}_s(\mathbb{X}_s'\mathbb{X}_s)^-\mathbb{X}_s'-\mathbb{R}_{s,z}(\mathbb{R}_{s,z}'\mathbb{R}_{s,z})^-\mathbb{R}_{s,z}'\}[e+n_s+n_s]},$$

   where n=the number of observations,

   p=rank($\mathbb{X}_s$) and q=rank($\mathbb{Z}$).   (3.7.21)

When $(n_s+n_s)=\mathbb{O}$ for all s, as in strict additivity, or is negligible, as in the randomization test, each of the $[N(\mathbb{G})t!]^{b-1}$ clusters collapses to a single value. Otherwise, the relative values of $(\mathbb{Z}\mathbb{P}+e)$ and (n_s+n_s) determine the extent of cluster dispersion and overlap.

As with $R_S(\digamma|\mu,b,\tau)$, clusters and/or individual

randomizations may have identical values of $F_S(\digamma|\mu,b,\tau)$.

Equation (3.7.21) provides a formula for the $t_{@S,S}$

of (3.6.1), (p. 170), and (3.6.5), (p. 173), and the

$t_{@S}$ of (3.6.8), (p. 174). This result proves the

assertion in Pitman of $(t!)^{b-1}$ unique values of the test

statistic and details the conditions under which it

holds [Pitman, 1937, bottom six lines of p. 323]. (We

state his assertion in our notation.) His model lacks a

covariate. Without one, the first projector of

$R_S(\tau|\mu,b)=R_S(\mu,b,\tau)-R(\mu,b)$ has the special form giving

it identical values for randomizations having the same

value of g, as in (3.7.9), (p. 194), The second projector

is constant for all randomizations. Thus, one obtains

Pitman's result.

Graph 4.5.9, (p. 288), illustrates Part 3 via the

OSL_F's associated with the $F_S(\digamma|\mu,b,\tau)$. The automatic

scaling has grouped clusters with similar values, but in

each case there are an integral multiple of $N(\mathfrak{S})t! =$

$1(4!)=24$ randomization per bar of the histogram. The

points marked "u" on Graphs 4.6.5 and 4.6.6, (p. 299),

also illustrate Part 3. Each "u" is one cluster of

$N(\mathfrak{S})t! = 24$ points when $\eta_S=n_S=\mathbb{O}$.

Proof of Part 4:

By the definition of $\hat{\beta}_{S;\mu,b,\tau,\beta}$, (2.4.17), (p. 29), and Theorem 3.7.8, (p. 201),

$$\hat{\beta}_{S;\mu,b,\tau,\beta}$$

$$= (\mathbb{R}_{S,z}{}'\mathbb{R}_{S,z})^{-}\mathbb{R}_{S,z}{}'[\mathbb{Z}\beta + e + (\eta_S + n_S)]$$

$$= \beta + (\mathbb{R}_{S,z}{}'\mathbb{R}_{S,z})^{-}\mathbb{R}_{S,z}{}'[\ e\ ]$$

$$+ (\mathbb{R}_{S,z}{}'\mathbb{R}_{S,z})^{-}\mathbb{R}_{S,z}{}'[\ \eta_S + n_S\ ]. \qquad (3.7.22a)$$

The first term, β, is constant for all randomizations. The second term (in e) is constant within each randomization group. The third term (in $\eta_S + n_S$) differs with each randomization and disperses values about the $[N(\mathbb{G})t!]^{b-1}$ cluster centers set by the second term. When $(\eta_S + n_S) = \mathbb{O}$, as in strict additivity, or is negligible as in the randomization test, each of the $[N(\mathbb{G})t!]^{b-1}$ clusters collapses to a single value. Otherwise, the relative values of e and $(\eta_S + n_S)$ determine the extent of cluster dispersion and overlap. As with $R_S(\beta|\mu,b,\tau)$, clusters and/or individual randomizations may have identical values of $\hat{\beta}_{S:\mu,b,\tau,\beta}$. Equation (3.7.22a) is a generalization to multiple covariates and additional error types of a result found in Cochran [1957, the second to last equation of p. 272].

The last parts of (3.7.22a) are the least squares estimator of $\beta_{S,\varepsilon}$ in the model equation

$$[e + \eta_S + n_S] = \mathbb{R}_{S,z}\beta_{S,\varepsilon} + \text{error} \qquad (3.7.22b)$$

$$= (\mathbb{R}_{S,x}\mathbb{Z})\beta_{S,\varepsilon} + \text{error}$$

$$= \{[\mathbb{I} - \mathbb{X}_S(\mathbb{X}_S{}'\mathbb{X}_S)^{-}\mathbb{X}_S{}']\mathbb{Z}\}\beta_{S,\varepsilon} + \text{error}.$$

This equation should be contrasted with (2.4.8), (p. 27). For the one covariate case, a near zero correlation of $[e+\eta_S+n_S]$ with $\{[\mathbb{I}- \mathbb{X}_S(\mathbb{X}_S{}'\mathbb{X}_S)^{-}\mathbb{X}_S{}']\mathbb{Z}\}$ implies $\beta_{S,\varepsilon} \approx 0$ and a small bias of $\hat{\beta}_{S:\mu,b,\tau,\beta}$.

Graph 4.5.3, (p. 277), illustrates Part 4 when $e \neq \mathbb{O}$ and $\eta_S = n_S = \mathbb{O}$. The automatic scaling has grouped clusters with similar values, but in each case, there is an integral multiple of $N(\mathbb{G})t! = 1(4!) = 24$ randomizations per group. Graph 4.5.4, (p. 278), illustrates Part 4 when $e = n_S = \mathbb{O}$ and $\eta_S \neq \mathbb{O}$. It is similar to the case $e = \eta_S = \mathbb{O}$ and $n_S \neq \mathbb{O}$.

Proof of Part 5:

When $\tau = \mathbb{O}$, (3.7.18), (p. 203), lacks the $\mathbb{X}_S\tau$ term. Consequently, $R_S(\tau|\mu,b,\beta)$ is

$$(\mathbb{Z}\beta + e)' \{\mathbb{R}_S - \mathbb{B}\} (\mathbb{Z}\beta + e)$$

$$+ 2(\mathbb{Z}\beta + e)' \{\mathbb{R}_S - \mathbb{B}\} (\eta_S + n_S)$$

$$+ (\eta_S + n_S)' \{\mathbb{R}_S - \mathbb{B}\} (\eta_S + n_S), \qquad (3.7.22c)$$

where $\mathbb{R}_S$ is the projector for $[\mathbb{X}_\mu|\mathbb{X}_b|\mathbb{X}_{S,\tau}|\mathbb{Z}]$ and $\mathbb{B}$ is the projector for $[\mathbb{X}_\mu|\mathbb{X}_b|\mathbb{Z}]$.

Equation (3.7.22c) is (3.7.19) for $\tau = 0$. By Theorem

3.7.6(a), (p. 193), the first term is constant for all

$[N(\Xi)t!]$ randomizations having the same value of the g of

(3.7.9), (p. 194). It sets at most $[N(\Xi)t!]^{b-1}$ cluster

centers. Within each g-group, the other two terms of

(3.7.22c) disperse values about each cluster center.

When $(\tau_s + n_s) = 0$ for all s, as in strict additivity, or is

negligible as in the randomization test, each of the at

most $[N(\Xi)t!]^{b-1}$ clusters collapses to a single value.

Otherwise, the relative values of e and $(\tau_s + n_s)$ determine

the extent of cluster dispersion and overlap. As with

$R_S(\beta | \mu, b, \tau)$, clusters and/or individual randomizations

may have identical values of $F_S(\tau | \mu, b, \beta)$ when $\tau = 0$. As

the other parts of the sum of squares for the entire

model are constant or clustered within a g-group, when

the τ part is zero, then the model sum of squares is also

constant or clustered.

Graphs 4.5.5, (p. 280), 4.6.7, (p. 302), 4.6.8, (p.

303) and the top left part of Display 4.8.7, (p.350-351),

illustrate Part 5. In Graph 4.6.21, $\tau = 0$, and all $N(\Xi)t!$

$= 1(4!) = 24$ values of the OSL for $F_S(\tau | \mu, b, \beta)$ of this

g-group have collapsed to the single point marked "X".

This completes the proof of Theorem 3.7.9.

The projector for $[X_\mu | X_b | Z]$ appearing in the sum of squares for the treatments as adjusted for the mean, blocks and covariate, lacks the special structure which gives multiple randomizations identical values for the sum of squares, $R_S(\tau | \mu, b, \beta)$. When $\tau \neq 0$, its distribution has $[N(\mathcal{G})t!]^b$ distinct values, as does the corresponding F-ratio, even with strict additivity or in the randomization test. The clustering of values of the denominator, (3.7.20), of the F-ratios may impart clustering to the ratio itself. Similarly, the estimate of τ as adjusted, $\hat{\tau}$, will have $[N(\mathcal{G})t!]^b$ distinct values. The clustering of the $\hat{\beta}$ may impart a clustering to the $\hat{\tau}$. In both cases, the degree of clustering depends upon the relative values of the elements of the involved vectors and matrices. As with $R_S(\beta | \mu, b, \tau)$, the sums $(Z\beta + e)$ and/or $(\eta_s + n_s)$ may give different randomizations identical values.

3.7.4 Permutation and the Noncentrality Parameters

The following two theorems detail the effect randomization has upon the noncentrality parameters.

Theorem 3.7.10.

For (a) those design matrices $\mathbf{X}_S = [\mathbf{X}_\mu | \mathbf{X}_b | \mathbf{X}_{S,\tau}]$ with projectors $[\mathbf{X}_S(\mathbf{X}_S{}'\mathbf{X}_S)^{-}\mathbf{X}_S{}']$, all of whose diagonal blocks have, within each block, identical diagonal elements and identical off-diagonal elements (as in Theorem 3.7.5), and (b) a constant variance σ^2 and (c) for a number of randomizations equal to $[N(\mathbb{G})t!]^b$, there are at most $[N(\mathbb{G})t!]^{b-1}$ unique values for the noncentrality parameter

$$\lambda_{S,\beta|\mu,b,\tau,\beta} \tag{3.7.23}$$

$$= \{1/(2\sigma^2)\} \times \beta'\{\mathbf{Z}'\mathbf{Z} - \mathbf{Z}'\mathbb{M}_S[\mathbf{X}_1(\mathbf{X}_1{}'\mathbf{X}_1)^{-}\mathbf{X}_1{}']\mathbb{M}_S{}'\mathbf{Z}\}\beta.$$

Each unique value will occur in an integral multiple of $[N(\mathcal{E})t!]$ randomizations for the g-groups of (3.7.9). When experimental unit-treatment interaction, η_S, is included in the model equation, one may consider the variance to change with the randomization. This will disperse the values of this noncentrality parameter about the cluster centers provided by a constant variance. Note that none of the error terms directly enters into the value of the noncentrality parameters.

When Σ_τ has all columns summing to zero within each block, as in the effects version of Σ_τ, (3.7.23) becomes

$$\lambda_{s,\beta|\mu,b,\tau,\beta} = \{1/(2\sigma^2)\} \times \{ \beta'Z' [\mathbb{I}-\mathbb{W}] Z\beta \}$$

$$- \{1/(2\sigma^2)\}$$

$$\times \{\beta'Z' \mathbb{M}_S[\Sigma_{1,\tau}(\Sigma_{1,\tau}'\Sigma_{1,\tau})^-\Sigma_{1,\tau}']\mathbb{M}_S' Z\beta \},$$

where $\mathbb{W}=(1/t)\{diag[\mathbb{J}_{txt}]\}_{bxb}$ as in (2.6.20).   (3.7.24)

Proof: Apply Theorem 3.7.6, (p. 193), to (2.6.37), (p. 42), with $\Sigma_\tau=\Sigma_S$. Use (3.7.5) to obtain (3.7.23). Apply Theorem 2.6.16, (p. 56), to (3.7.23), grouping the $\mathbb{I}$ and $\mathbb{W}$ matrices, to obtain (3.7.24). Theorem 3.7.6(b) provides the number of groups and the number of randomizations within the groups. The values of Z may be such that two or more randomizations yield identical values of $\lambda_{s,\beta|\mu,b,\tau,\beta}$. The covariates of Table 2.7.1, (p. 65), have but 24-1=23 unique values, one value appearing 2(24)=48 times. The vertical lines of points of Graph 4.8.1, (p.338_), illustrate this theorem. This graph led to the theorems of Section 3.7.

The abbreviated notation used subsequently for (3.7.24) is

$$\lambda_{s,\beta|\mu,b,\tau,\beta} = \lambda_{f,\beta} - \lambda_{g,\beta} , \qquad (3.7.25)$$

where $\lambda_{f,\beta} = \{1/(2\sigma^2)\} \times \{\beta'Z' [I-W] Z\beta\}$,

which is fixed for all randomizations, and

where $\lambda_{g,\beta} = \{1/(2\sigma^2)\}$

$\times \{\beta'Z' M_s[X_{1,\tau}(X_{1,\tau}'X_{1,\tau})^{-}X_{1,\tau}']M_s' Z\beta\}$, which

is constant within subgroups, g, of randomizations having identical $X_s(X_s'X_s)^{-}X_s'$ matrices.

This abbreviated notation holds when (a) X gives its projector the necessary structure, as discussed in Theorem 3.7.5, (p. 190), also below (2.6.28), (p. 55), and when (b) the columns of X_τ sum to zero within each block, as discussed in Theorem (2.6.16), (p. 56). The single factor case with X_τ in its effects version satisfies these conditions. These conditions are not satisfied when X_τ is in its coding version.

Theorem 3.7.11.

There are at most $[N(\mathfrak{S})t!]^b$ unique values for the noncentrality parameter

$$\lambda_{s,\tau|\mu,b,\tau,\beta} = \{1/(2\sigma^2)\} \times \tau'\{X_{1,\tau}'X_{1,\tau}\}\tau$$

$- \{1/(2\sigma^2)\} \times \tau'\{ X_{1,\tau}'M_s' [X_r(X_r'X_r)^{-1}X_r'] M_sX_{1,\tau}\}\tau$,

where $X_r = [X_\mu|X_b|Z]$, constant for all s. \qquad (3.7.26)

When the treatment matrix Σ_τ is orthogonal to both the blocks matrix Σ_b and the mean vector Σ_μ, the second term of (3.7.26) may be reexpressed giving

$$\lambda_{s,\tau|\mu,b,\tau,\beta} = \{1/(2\sigma^2)\} \times \tau'\{\Sigma_{1,\tau}'\Sigma_{1,\tau}\}\tau$$

$$-\{1/(2\sigma^2)\} \times \tau'\{\Sigma_{1,\tau}'M_s' [\mathbb{Z}(\mathbb{Z}'(\mathbb{I}-\mathbb{W})\mathbb{Z})^{-1}\mathbb{Z}'] M_s\Sigma_{1,\tau}\}\tau,$$

where $\mathbb{W}=(1/t)\{diag[\mathbb{I}_{txt}]\}_{bxb}$ as in (2.6.20).   (3.7.27)
The vertical spread of the points of Graph 4.8.1, (p.338), illustrates this theorem its form of (3.7.28) below.

Proof: Applying Theorem 3.7.3, (p. 189), to the first term of (2.6.39), (p. 62), gives the first term of (3.7.26). The second term follows directly from (2.6.39) with $\Sigma_{s,\tau} = M_s\Sigma_{1,\tau}$. Application of Theorem 2.6.22, (p. 51), to the second term of (3.7.26) provides the second term of (3.7.27). Note that $\Sigma_r=[\Sigma_\mu|\Sigma_b|\mathbb{Z}]$ is constant for all randomizations. Neither version of the projector, $[\Sigma_r(\Sigma_r'\Sigma_r)^{-}\Sigma_r']$ or $[\mathbb{Z}(\mathbb{Z}'(\mathbb{I}-\mathbb{W})\mathbb{Z})^{-1}\mathbb{Z}']$, has the special form of its diagonal blocks which gives identical values to g-groups of randomizations. It is possible for a given Σ and $\mathbb{Z}$ to have different individual randomizations with identical values for $\lambda_{s,\tau|\mu,b,\tau,\beta}$. The covariates of Table 2.7.1, (p. 65), do this. This appears to regularly occur within some randomization groups, but not in others. An extension is to discover why.

The abbreviated notation used subsequently for (3.7.27) is

$$\lambda_{s,\tau|\mu,b,\tau,\beta} = \lambda_{f,\tau} - \lambda_{s,\tau}, \tag{3.7.28}$$

where $\quad \lambda_{f,\tau} = \{1/(2\sigma^2\} \times \{\tau'\Sigma_{1,\tau}'\Sigma_{1,\tau}\tau\}$

which is fixed for all randomizations,

and $\quad \lambda_{s,\tau} = \{1/(2\sigma^2)\} \times$

$$\tau'\{ \Sigma_{1,\tau}'\mathbb{M}_s' [\Sigma_r(\Sigma_r'\Sigma_r)^-\Sigma_r'] \mathbb{M}_s\Sigma_{1,\tau} \}\tau,$$

$$= \{1/(2\sigma^2)\} \times$$

$$\{\tau'\Sigma_{1,\tau}' \mathbb{M}_s'[\mathbb{Z} (\mathbb{Z}'(\mathbb{I} - \mathbb{W})\mathbb{Z})^{-1} \mathbb{Z}']\mathbb{M}_s \Sigma_{1,\tau} \tau\},$$

which changes with each randomization.

As in (2.6.20), $\mathbb{W}=(1/t)\{diag[\mathbb{J}_{txt}]\}_{bxb}$

and $\Sigma_r=[\Sigma_\mu | \Sigma_b | \mathbb{Z}]$.

For each $\lambda_{g,\beta}$ with the same value, there are $[N(\mathbb{S})t!]$ $\lambda_{s,\tau}$. Typically, each $\lambda_{s,\tau}$ has a unique value.

This abbreviated notation holds for Σ_τ orthogonal to both the blocks and mean matrices. This is the case when both Σ_b and Σ_τ are in their effects version.

An extension is to derive, if possible, the fixed and random parts of the noncentrality parameters with fewer restrictions on Σ_τ, such as, multiple factors and/or the coding version.

Note that the random part of both noncentrality parameters can only decrease the value of the entire parameter. See the discussion of positive semi-definite matrices below (2.6.11), (p. 43).

3.7.5 Restricting Randomizations to Maximize Power

Selecting the randomization which maximizes the noncentrality parameter maximizes the power of the F-test under the normally distributed random errors model, as discussed in (2.6.6a), (p. 41).

When the noncentrality parameters may be divided into fixed and random parts, as in Theorems 3.7.10, (p. 211), and 3.7.11, (p. 213), and there is but one covariate, Theorem 3.7.15, below, allows one to simultaneously obtain

    a) a near maximum power for the F-test of alterna-
       tive, nonzero, values of the treatment effect, as
       adjusted for the mean, blocks and covariate, and

    b) the maximum power for the F-test of an
       alternative, nonzero, univariate, covariate slope
       coefficient, β, as adjusted for the mean, blocks
       and treatments.

Furthermore, one retains a degree of randomness in the assignment of experimental units to treatments. The randomness obtained satisfies goal two of Section 3.2.8, protection against favoring a particular treatment, (p. 118), but not necessarily goal one, freedom from the same treatment being applied to adjacent experimental units.

Further work is needed on the effect of this restriction on goal one and on model equation robustness, goal three.

For use in the proofs which follow we state the result,

Theorem 3.7.12.

For $A = B'B$ of full rank, $(x'y)^2 \leq (x'Ax)(y'A^{-1}y)$,

with equality when $x \propto A^{-1}y$,

where x and y are column vectors. (3.7.29)

Citation: C. R. Rao, 1973, p. 54, Section 1e.1,(ii)(b), Equation 1e.1.4. This is a version of the Cauchy-Schwarz inequality. Here, $\propto$ designates a constant proportion for each element of the vectors x and $A^{-1}y$.

Theorem 3.7.13.

For X_τ with full column rank

$$(\beta'Z'M_S X_{1,\tau}\tau)^2 \leq$$

$$\{\beta'Z'M_S X_{1,\tau}(X_{1,\tau}'X_{1,\tau})^{-1}X_{1,\tau}'M_S'Z\beta\}$$

$$\times \ \{\tau'(X_{1,\tau}'X_{1,\tau})\tau\} \ . \tag{3.7.30}$$

Proof:

Let $A = (X_{1,\tau}'X_{1,\tau})^{-1}$,

$x = [X_{1,\tau}' \ M_S' \ Z]\beta$ and

$y = \tau$.

Theorem 3.7.12 gives the desired result. Only most unusual τ, β and Z would make the two vectors of Theorem

3.7.12, $x=[\Sigma_{1,\tau}{}' \mathbb{M}_s Z]\beta$ and $\mathbb{A}^{-1}y=(\Sigma_{1,\tau}{}'\Sigma_{1,\tau})\tau$, proportional

($\propto$) to one another. Thus, equality in (3.7.30) or

(3.7.31), below, is unlikely.

In the abbreviated notations of (3.7.25), (p. 213),

and (3.7.28), (p. 215), and under those conditions,

(3.7.30) becomes

$$L_s(2\sigma^2)^2 \lessapprox (\lambda_{g,\beta})(\lambda_{f,\tau}), \qquad (3.7.31)$$

where $L_s = (\beta' Z' \mathbb{M}_s \Sigma_{1,\tau}\tau)^2$ with the subscript s indicating

that the left hand side changes with each randomization.

Theorem 3.7.14.

For Σ_τ orthogonal to Σ_μ and Σ_b, with all columns of

Σ_τ summing to zero within each block, and with Σ_τ and Z

having full column rank,

$$(\tau' \Sigma_{1,\tau}{}' \mathbb{M}_s' Z\beta)^2 \quad \lessapprox$$

$$\{\tau' \Sigma_{1,\tau}{}' \quad \mathbb{M}_s' Z [Z'(\mathbb{I}-\mathbb{W})Z]^{-1} Z' \mathbb{M}_s \quad \Sigma_{1,\tau}\tau\}$$

$$x \quad \{\beta' [Z'(\mathbb{I}-\mathbb{W})Z] \beta\} , \quad \text{where}$$

$\mathbb{W}=(1/t)\{diag[\mathbb{J}_{txt}]\}_{bxb}$ as in (2.6.20).   (3.7.32)

Proof:

Let $\mathbb{A} = [Z'(\mathbb{I}-\mathbb{W})Z]^{-1}$,

$x = [Z' \mathbb{M}_s \Sigma_{1,\tau}]\tau$ and

$y = \beta$.

Theorem 3.7.12 gives the desired result. With but one

covariate, the two vectors of Theorem 3.7.12,

$x=[Z'M_s\Sigma_{1,s}]\tau$ and $A^{-1}y=[Z'(I-W)Z]^{-1}\beta$, are scalers and thus are always proportional. Therefore, equality holds in (3.7.32) and (3.7.33), below.

In the abbreviated notations of (3.7.25), (p. 213), and (3.7.28), (p. 215), and under those conditions, (3.7.32) becomes

$$L_s(2\sigma^2)^2 \leq (\lambda_{s,\tau})(\lambda_{f,\beta}). \qquad (3.7.33)$$

where $L_s=(\beta'Z'M_s\Sigma_{1,\tau}\tau)^2$, with the subscript s indicating that the left hand side changes with each randomization. Note that L_s has the same value in Theorems 3.7.13 and 3.7.14.

Theorem 3.7.15.

Consider an arbitrary group, g, of $[N(\Xi)t!]$ randomizations, s, with identical values of $\lambda_{g,\beta}$. When there is only one covariate and the conditions on Σ_τ of Theorems 3.7.10, (p. 211), and 3.7.11, (p. 213), hold, for any such group, g,

$$\lambda_{s,\tau} \leq \left\{\frac{\lambda_{f,\tau}}{\lambda_{f,\beta}}\right\} \lambda_{g,\beta} , \qquad (3.7.34)$$

independently of β, for all τ and for all s in g.

Proof: For a single covariate, (3.7.33) becomes $L_s(2\sigma^2)^2 = (\lambda_{s,\tau})(\lambda_{f,\beta})$. Applying this to (3.7.31) gives $(\lambda_{s,t})(\lambda_{f,\beta}) \leq (\lambda_{g,\beta})(\lambda_{f,\tau})$. Division by $\lambda_{f,\beta}$ gives the above inequality.

Examination of the detailed form of $\lambda_{g,\beta}$ and $\lambda_{f,\beta}$

for the single covariate case, shows each to be scalers

multiplied by β^2. Thus, β cancels in (3.7.34) and

(3.7.36), below. Equality in (3.7.34) holds for only

those most unusual combinations of τ, β and $\mathbb{Z}$ which give

equality in (3.7.31). Graph 4.8.1, (p. 338), illustrates

this theorem. Graph 4.8.2, (p. 339), illustrates the

case where the variance σ^2 changes with each randomi-

zation s. This disperses the columns of values as seen

in Graph 4.8.1.

The end of Section 2.6.1, (p. 41), and of Section

3.6.7, (p. 181), discuss increasing the power of the

hypothesis test when using the F-statistic under the

normally distributed random errors model. For given

degrees of freedom, one wishes to maximize the non-

centrality parameter. Since $\lambda_{s,\tau}$ can only reduce the

value of $\lambda_{\tau|\mu,b,\tau,\beta}$, large values of $\lambda_{s,\tau}$ should be

avoided. Theorem 3.7.15 shows how to do this. Selecting

the randomization set g_{min} with the minimum value of $\lambda_{g,\beta}$

simultaneously selects the $N(\mathbb{S})t!$ randomizations

whose theoretical maximum $\lambda_{s,\tau}$ is the minimum. Larger

values of $\lambda_{s,\tau}$ and $\lambda_{g,\beta}$ are thus avoided. One conducts

the actual experiment using one of the randomizations in

g_{min}. This set has $[N(\mathbb{E})t!]$ randomizations from which to

choose the one used to conduct the actual experiment.

Theorem 3.7.15 does not ensure selecting the randomization with the minimum $\lambda_{s,\tau}$, nor selection from among a set of randomizations with the smallest $\lambda_{s,\tau}$. Values of $\lambda_{s,\tau}$ smaller than the smallest $\lambda_{s,\tau}$ in the group g_{min} may exist in other groups of larger $\lambda_{g,\beta}$. Theorem 3.7.15 merely allows avoiding large values of $\lambda_{s,\tau}$. By minimizing the random part, $\lambda_{s,\tau}$, one maximizes $\lambda_{s,\tau|\mu,b,\tau,\beta}$. Theorem 3.7.15 shows how to avoid $\lambda_{s,\tau|\mu,b,\tau,\beta}$ which are far from the maximum.

The ratio $\{\lambda_{f,\tau}/\lambda_{f,\beta}\}$ of (3.7.34) approximates the regression slope coefficient for the least squares line fit to the $[N(\mathbb{G})t!]$ points, (X,Y), where $X=\lambda_{g,\beta}$ and Y=observed maximum $\lambda_{s,\tau}$ in group g. This line is

(observed maximum $\lambda_{s,\tau}$)=(slope)($\lambda_{g,\beta}$)+error. (3.7.35)

When $\lambda_{f,\tau} \approx \lambda_{f,\beta}$, the approximated slope is near horizontal (one) and large reductions in $\lambda_{g,\beta}$ are needed to appreciably reduce the maximum of the $\lambda_{s,\tau}$ for those randomizations in group g. The range of values of $\lambda_{g,\beta}$ and the ratio $\lambda_{f,\tau}/\lambda_{f,\beta}$ indicates the sensitivity of the maximun within g-group noncentrality parameter, $\lambda_{s,\tau}$, to different randomizations.

From (2.6.34) and (2.6.38), $0 \leq \lambda_{s,\tau} \leq \lambda_{f,\tau}$ and $0 \leq \lambda_{g,\beta} \leq \lambda_{f,\beta}$. Thus, dividing both sides of (3.7.34)

by $\lambda_{f,\tau}$ gives, for all s in a fixed g,

$$0 \leq \frac{\lambda_{s,\tau}}{\lambda_{f,\tau}} \leq \frac{\lambda_{g,\beta}}{\lambda_{f,\beta}} \leq 1 . \qquad (3.7.36)$$

For the one-covariate, one-factor case, the noncentrality parameters of the F-test for the two usual hypotheses of interest are interconnected.

Theorem 3.7.15 does not hold with multiple covariates. None of the noncentrality parameters depend upon the data. With sufficient resources, all values of $\lambda_{s,\tau}$ and $\lambda_{g,\beta}$ could be computed and ranked for those alternative τ and β of interest. Randomization could be restricted to those randomizations with low ranks for both $\lambda_{s,\tau}$ and $\lambda_{s,\beta}$, should such exist. Widely differing alternatives of τ and β would necessitate recomputing the rankings. Scaler multiplication of the τ or β would not alter the rankings. The randomization group supplying g_{min} appears to be the one with the smallest elements of $\mathbf{Z}'\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-}\mathbf{X}_1'$. When s=1 orders the covariates from low to high within each block, g_{min} occurs in the randomization in which the unit-valued elements of the off-diagonal block of $\mathbf{X}_g(\mathbf{X}_g'\mathbf{X}_g)^{-}\mathbf{X}_g'$ are on the upper-right to lower-left diagonal. This randomization has the least squaring of the elements of $\mathbf{Z}$; it is the most balanced. An extension is to consider consider multiple factors, interactions, and other such $\mathbf{X}_\tau$ as the coding version.

Section 3.8

Expected Value of the F-Ratio

Section 3.8.1 reviews known results for the expected value of the F-ratio in the randomized block design without a covariate. Section 3.8.2 discusses results for the design with a covariate. Section 3.8.3 discusses possible further work on this expectation. The review of literature is the major contribution of this section.

3.8.1 The Randomized Block Design without a Covariate

For the balanced randomized block design with b blocks and t treatments, but without a covariate, the F-ratio has a numerator degrees of freedom n=t-1 and a denominator degrees of freedom d=[(number of observations-1)-(b-1)-(t-1)] and is

$$F_s = \frac{(\text{Sum of Squares for Treatments})_s /n}{(\text{Sum of Squares for the Residuals})_s/d} \ . \qquad (3.8.1)$$

The transformation of Appendix A.4, (p. 384), yields

$$R_s = \frac{(\text{Sum of Squares for Treatments})_s}{\text{Total Sum of Squares}} \ , \qquad (3.8.2)$$

which is the coefficient of determination, usually denoted as R^2. Under the assumption of normally distributed random errors, R_s has the non-central Beta distribution, as defined in Appendix A.4.

The advantage of R_s over F_s is, that under the null hypothesis of no treatment effects, the denominator of R_s is constant for all randomizations. The constant

denominator makes possible deriving the moments of R_s

under the null hypothesis and the randomization model.

Bartlett [1935] derived the mean and variance of R_s for

the case of two treatments. Plackett [1960, p. 155-156]

presents a clear proof with discussion. B. L. Welch

[1937] and Pitman [1937] extend the results to any number

of treatments. Pitman adds the third and fourth moments.

All authors consider only experimental unit error; both

experimental unit-treatment interaction error and

normally distributed random error are zero. All assign

equal probability to each randomization and consider only

the null hypothesis case. Under the alternative

hypothesis of treatment effects, the denominator of R_s is

not constant for all randomizations [Atiqullah, 1963, p.

337] If the σ^2 of the non-centrality parameter is

considered to change with the randomization, for example,

when experimental unit-treatment interaction is present,

then the non-centrality parameter also changes with the

randomization.

When the null hypothesis of no treatment effects is

true, the first moment is 1/b under *both* the normally

distributed errors and the randomization models. That

is, one obtains the same first moment of the statistic R_s

when it is averaged over these two quite different proba-

bility spaces. This may help explain why the normally

distributed random errors model provides a good

approximation to the randomization model for this design.

The variance of the central-Beta distribution is

$$\sigma_B^2 = \frac{2(b-1)}{b^2[b(t-1)+2]} \approx \frac{2b}{b^3[t-1]} = \frac{2}{b^2[t-1]} . \qquad (3.8.3)$$

The variance of R_S under the randomization model is

$$\sigma_R^2 = \frac{2(1-A)}{b^2(t-1)}, \quad \text{where } A = \frac{\sum\limits_{i=1}^{b} \left[\sum\limits_{j=1}^{t} e_{ij}^2 \right]^2}{\left[\sum\limits_{i=1}^{b} \sum\limits_{j=1}^{t} e_{ij}^2 \right]^2} . \qquad (3.8.4)$$

When the variances of the experimental units are equal

within each block, A takes its minimum value, 1/b, in

which case $(1-A) = (b-1)/b$ and $\sigma_B^2 \approx \sigma_R^2$. When the

variances of the e_{ij} are zero in all blocks but one, then

A takes its maximum value of 1. In this case the two

variances differ. When the number of blocks is large,

and the block variances approximately equal, then the

effect of A is minimal and the two variances are near

equal. Pitman [1937, p. 331-335] discusses conditions

yielding near-equality for the third and fourth moments.

B.L. Welch [1937, p.47] and Pitman [p.335] conclude that

the randomization test may be approximated by the normal

theory test, especially in the upper tail, for the rando-

mized block design. This conclusion does not hold for

all designs, such as the Latin Square [B.L. Welch, p.47].

Atiqullah [1963] considers the randomized block

design with one factor with t levels. He considers only

experimental unit error (additivity) and assigns equal

probability to all randomizations.   Instead of the

F-ratio, he uses Fisher's z transformation of the mean

squares (MS),

$$Z=(1/2)\ln(MS\text{-hypothesis}/MS\text{-residual}). \qquad (3.8.5)$$

Atiqullah uses a Taylor's series expansion and cumulants

to derive approximate limiting moments for the null and

non-null distribution of Fisher's Z [Atiqullah, 1963, Eq.

21, 22, and 35, 36].   The limit is taken as the number of

treatment levels t goes to infinity.

Atiqullah [1962A] derives similar approximate limit-

ing moments for Z under an infinite population model with

a fixed allocation of treatments to experimental units

but an unknown distribution of the errors [his equations

25 and 26].   For the null hypothesis, applying the normal

distribution to these results gives expressions identical

to those derived under the randomization model, provided

the within-block variances of the experimental unit

errors are equal [See Atiqullah, 1963, p. 338].

Robinson [1973A] shows that the critical value of

the F-ratio tends to a constant in probability as the

number of blocks, b, becomes large.   Further, as b goes

to infinity, the sequence of F-ratios, $\{F_b\}$, tends in

distribution to a central chi-square distribution, when

$\{F_b\}$ goes to zero, or tends to a non-central chi-square

distribution, when $\{F_b\}$ goes to a constant.   Thus, the

randomization test of the null hypothesis of no treatment

effects, is asymptotically as powerful as the test using

the F ratio under the normally distributed random errors

model as the number of blocks becomes large. Robinson

includes only experimental unit error in his model and

assigns equal probability to all randomizations.

3.8.2 The Randomized Block Design with a Covariate

When a covariate is included in the randomized block

design, Robinson [1973B, p. 368] writes, "it is difficult

to obtain results on the usual test statistic of the

[F-Ratio]. In fact, is has not been possible to obtain

moments of this statistic under the randomization model,

as was done by Pitman [1937] and B. L. Welch [1937] for

the analysis of variance case. However, it is possible

to obtain asymptotic results"

D. R. Cox [1956, p. 1147] explains the source of

this difficulty,

> If we try to calculate the randomization
>
> expectations of [the residual and treatment
>
> mean square, both adjusted for the covariate],
>
> there is the difficulty that $[SSR_{yz}]^2/[SSR_{zz}]$
>
> is a ratio of random variables so that no
>
> simple exact expression of the form of its
>
> expectation can be written down.

Table 2.5.2, (p. 38) and Section 2.5.2, (p. 35) show

this. Different randomizations tally different covariate

values to obtain the sums SST_{yz} and SST_{zz}. Hence, SSR_{yz}

and SSR_{zz} differ with each randomization.

Atiqullah [1962B] uses a mixed model equation and derives moments for the Z of (3.8.5), (p. 226). This Z corresponds to the Z of the fixed model for the "random" coefficients as adjusted for the fixed coefficients. While such results could apply to the analysis of covariance, his results are limited to the case where the projector for the adjusted sums of squares, such as $\mathbb{R}_S(\mathbb{R}_S'\mathbb{R}_S)^{-1}\mathbb{R}_S'$ of Table 2.5.1, (p. 37), has identical elements along its diagonal. This condition he terms quadratic balance [p. 141]. In only unusual cases will a continuous covariate be quadratically balanced.

The 1964 paper by Atiqullah discusses the sensitivity of the F-test in the analysis of covariance for a design with one factor and an arbitrary number of covariates but without blocks. The errors of his model resemble those of the normally distributed random errors model, but have some non-normal distribution. The model is not a randomization model. One finding is that the lack of quadratic balance in the projector for the covariates determines the sensitivity of the F-test to non-normality [p. 365]. Another is that non-normality of the errors has little effect on the F-test provided the covariates are normally distributed [p. 368]. He uses Fisher's Z, (3.8.5), for testing the treatments as adjusted by the covariate. His last three sections present results for random covariates, interaction

between slope and treatment and model equation terms
quadratic in the covariate.

The model equation of Robinson [1973B] has blocks,
one factor with t levels (treatments) and one covariate.
His estimator of the regression slope coefficient is
within each block, thus, it is invariant to the
randomization. When the estimator is over all blocks, as
herein, its value changes with each group of randomi-
zations, as proved in Theorem 3.7.9(4), (p. 202). He
considers only experimental unit error (additivity) and
assigns equal probability to all randomizations.

Robinson focuses on the treatments sum of squares as
adjusted by the covariate. His Theorem 2 proves the
asymptotic distribution of the associated test statistic,
S, under the null hypothesis, to be a central chi-square
with t-1 degrees of freedom as the number of blocks goes
to infinity. He constructs an increasing function of the
usual F-ratio (his Eq. 3). He then proves that a
function of the difference between these two test
statistics goes to zero in probability as the number of
blocks goes to infinity (his Theorem 3). Thus, [Rao,
1973, p. 122, (x)(d)], the increasing function of the
usual F-ratio has the same asymptotic distribution as S,
the central chi-square with t-1 degrees of freedom.
Rejecting the null hypothesis using the increasing
function of the usual F-ratio is equivalent to rejecting
using the F-ratio itself [Robinson, 1973B, p.372].

Robinson proves (his Theorem 5) that when the null hypothesis is false (the treatment effects are not all zero) and the sequence of the usual non-centrality parameter λ_b has a limit as the number of blocks b goes to infinity, then a second statistic tends in distribution to a non-central chi-square distribution with non-centrality parameter the limiting value of λ_b. As the number of blocks goes to infinity, the usual F-ratio, under the normally distributed random errors model, has the same chi-square limiting distribution.

Robinson concludes, "at least for a large number of blocks, the usual normal theory test is appropriate under a randomization model with the quite weak conditions of Theorem 1" [p. 372]. These conditions are that as the number of blocks goes to infinity (i) the sequence of the correlation coefficient of the treatments and covariate has a limit and (ii) the sequence of a sum of $\{SST_{xx}/SSTot_{xx} + SST_{zz}/SSTot_{zz}\}$ goes to zero.

3.8.3 Possible Approaches

Hooper [1989] groups nine published articles into four categories by the techniques used. These categories are (a) real and simulated data, (b) moment calculations, (c) Edgeworth expansions and (d) limit theorems. Hooper adds (e), assigning a model to the experimental unit errors. There appears to be no results for models including experimental unit-treatment interaction errors.

Consolidation of existing Edgeworth expansion and limit
theorems (c and d) may suggest additional results.
Gayen [1950] presents Edgeworth expansions. Some
possible approaches to categories (a) and (b) follow.

The availability of personal computers appears to
permit inexpensive calculation of the F-ratio for all
randomizations for small real-data uniformity trials.
The Eden and Yates [1933] study computed a sample of 1000
of the trial's possible 24^8 randomizations. While a bit
large, multiple personal computers could compute all
randomizations for this data in otherwise unused
time over the course of several months. See also
Center [1982].

Hsu's [1938] results on the moments of a quadratic
form, $x\mathbb{A}x'$, have been simplified for special cases by C.
R. Rao [1952], Plackett [1960, p. 40-41] and Atiqullah
[1962A]. These results assume the projector $\mathbb{A}$ is fixed
and the vector x has a random, but unknown continuous
distribution. The permutation matrix can be shifted from
the projector, here $\mathbb{A}$, to the vectors $\mathbb{Y}$ and $\mathbb{Z}$, making $\mathbb{A}$
fixed and the vectors random, albeit with discrete
distributions. More useful would be results on the
randomization moments of $x\mathbb{A}_S x'$ and $[x\mathbb{A}_S x']/[x\mathbb{B}_S x']$ for x
a fixed vector and $\mathbb{A}_S$ and $\mathbb{B}_S$ determined by randomization
s under a randomization model.

Atiqullah [1963] uses infinite population cumulants
to express the randomization moments. Finite population

moments are more appropriate, but need development in the
multivariate, mixed moment case. The tensor techniques
of McCullagh [1987] and the finite population results of
Irwin and Kendall [1943-45] and Boullion, Seaman and
Young [1990] appear to be one route to such development.

Hooper [1989] posits two models [his 1.8 and 2.2]
for the experimental unit errors. With the second model,
his Theorem 2.1 proves the power of a randomization test
converges to the asymptotic normal-theory power, as the
number of experimental units goes to infinity with the
number of treatments fixed. He discusses conditioning on
the experimental units or on the randomization used to
conduct the experiment. Several experimental designs are
presented, all without a covariate. While his proofs
appear to call for orthogonality of the columns of the
design matrix, perhaps proofs can be obtained which
include a covariate.

Zuskind, *et al.*, [1960, 1964, 1968] and Kempthorne,
et al., [1967] discuss related topics. See also the
results of Roux, discussed above on page 127.

In summary, the expected value of the F-ratio
under the randomization model has not been, and appears
never to become, an easy problem.

We next examine estimating the unit error terms by
reusing the experimental units.

Section 3.9

Experimental Procedures when Both Errors are Present

Section 3.9.1 sketches a two-stage procedure capable

of handling phenomena in which both experimental unit

and experimental unit-treatment interaction errors are

present. Section 3.9.2 discusses other experimental

procedures allowing control of both errors. All proce-

dures reuse experimental units, which is not always pos-

sible. These sections present some of the difficulties in

experiments when such errors are large. See Good [1979].

3.9.1 A Two-Stage Procedure

This two-stage procedure permits detection of

treatment effects when the experimental units may exhibit

both experimental unit and experimental unit-treatment

interaction errors. This procedure requires experimental

units capable of repeatedly receiving, in succession, all

treatments.

Stage one is a uniformity trial designed to insure

repeatability of measurement and to estimate the

experimental unit errors. No treatments are applied.

The initial model equation is

$$S_{ij\_1} = \mu + r_1 + b_i + \wp z_{ij} + e_{ij} + \eta_{ij\_1} \,. \qquad (3.9.1)$$

The term r_1 denotes the replication effect; the subscript

is the letter "1". The dash (_)in two subscripts reserves

a place for the treatment subscript which enters at stage

two. The model parameters may be defined as the least
squares "estimators" for some conceptual, but finite,
population of observations fitted to an appropriate model
matrix $[X \mid Z]$.

Alternatively, one may explicitly define the
parameters in terms of averages over the conceptual
population of observations. We use the averaging version
of the dot notation. First, adjust for the covariate,
giving $V_{ij\_1} = S_{ij\_1} - \beta Z_{ij}$. Then

$$\mu = V_{..\_.} , \qquad \text{the overall mean,}$$

$$r_1 = V_{..\_1} - V_{..\_.} , \text{ the replication effect,}$$

$$b_i = V_{i.\_.} - V_{..\_.} , \text{ the block effect,}$$

$$e_{ij} = V_{ij\_.} - V_{i.\_.} , \text{ the experimental unit error,}$$

$$\eta_{ij\_1} = (V_{ij\_1} - V_{..\_1}) - (V_{ij\_.} - V_{..\_.}), \text{ the}$$

experimental unit-replication interaction.
As β is not known, the interpretation of estimates of the
other model parameters will incorporate an uncertainty
from the estimate of β. Note that η is here interaction
with replication; there are no treatments in Stage One.

Stage One ends when the experimental technique gives
r_1 and η_{ij_1} near zero values for all repetitions and
reasonably stable estimates for the experimental unit
errors, e_{ij}. If these conditions cannot be achieved,
then the treatment effects will be confounded.

The Stage Two model equation is

$$W_{ijkl} = \mu + b_i + e_{ij} + \beta Z_{ij} + t_k + n_{i \cdot k} + n_{ijk} + \varepsilon_{ijkl}.$$

Adjust the response for the experimental unit errors and covariate slope coefficient estimated in Stage One by

$$Y_{ijkl} = W_{ijkl} - e_{ij} - \beta Z_{ij} ,$$

and obtain,

$$Y_{ijkl} = \mu + b_i + t_k + n_{i \cdot k} + n_{ijk} + \varepsilon_{ijkl}. \qquad (3.9.2)$$

The analysis of the actual experiment uses (3.9.2). As treatments are administered and possibly measured separatly within each block, say a hospital, we retain a block term in the model equation. The definitions of the model equation parameters in terms of averages over a conceptual population are

μ $\qquad$ $Y_{\ldots\ldots}$, $\qquad$ the overall mean,

b_i $\qquad$ $Y_{i\ldots} - Y_{\ldots\ldots}$, the block effect,

t_k $\qquad$ $Y_{\cdot\cdot k\cdot} - Y_{\ldots\ldots}$, the treatment effect,

$n_{i \cdot k}$ $\qquad$ $Y_{i \cdot k \cdot} - Y_{i\ldots} - Y_{\cdot\cdot k\cdot} + Y_{\ldots\ldots}$,

$\qquad$ the block-treatment interaction.

n_{ijk} $\qquad$ $(Y_{ijk\cdot} - Y_{ij\cdot\cdot}) - (Y_{i \cdot k\cdot} - Y_{i\ldots})$, the

$\qquad$ the within-block experimental unit-

$\qquad$ treatment interaction and

ε_{ijkl} $\qquad$ $Y_{ijkl} - Y_{ijk\cdot}$, the replication effect

$\qquad$ for experimental unit ij when it receives

$\qquad$ treatment k.

Estimates of these defined quantities will incorporate uncertainty from the estimates of e_{ij} and β; Z_{ij} is assumed known.

Alternatively, the model parameters may be defined as least squares estimates derived from a conceptual population of responses. As Stage One showed the replication effect ε_{ijkl} to be small, we assign it an error status with a normal distribution, mean zero, variance σ^2. Here we combine the randomization model with the normally distributed random errors model.

If the experimental units are homogeneous within blocks, responding to each treatment similarly, then η_{ijk} will be near zero. If the experimental units of each block respond to treatments similarly, then $\eta_{i \cdot k}$ will also be near zero.

Zhou, et al. [1989] provide an actual example of experimental unit-treatment interaction error. Subjects from two racial groups showed differences in response to metabolic clearance of the drug propranolol. Heretofore, patients had not been blocked on racial heritage for this drug, thus the interaction was within the blocks. In (3.9.2), had the blocks been hospitals with patients of mixed racial heritage and the treatment been propranolol vs a placebo, then the block-treatment interaction action ($\eta_{i \cdot k}$) would have been zero, while the experimental unit-treatment interaction (η_{ijk}) would have been non-zero.

As indicated by the subscripts, each experimental unit needs to be subjected to each treatment r (replication) times, a total of t x r measurements per experimental unit. This provides the error degrees of freedom needed for the statistical analysis. Such a design is applicable only in special cases. Possibly an incomplete block design would allow estimation of base response effects and a sufficiently broad screening of the experimental units for treatment interaction effects.

3.9.2 Other Experimental Procedures

The crossover experimental design permits one to partially separate the experimental unit-treatment interaction error from the measurement error. In this design, half of the experimental units are given first treatment A and later treatment B, perhaps a control. The other half of the experimental units receive the two treatments in the reverse order, B first, then A. If this cycle is repeated several times, those experimental units unusually susceptible to one or both treatments will repeatedly manifest themselves. With sufficient cycles one drives quite low the probability of so many repeatedly poor (or superior) performances exhibited under only one treatment being due to measurement errors alone. In the example of penicillin, repeated adverse reactions to the drug would indicate an experimental unit-treatment interaction error, not a measurement error. Of course, such experimental designs are not

always possible, as in the penicillin example, or in cases where changes in uncontrollable variables are too great between the cycles, such as growing-season-long agriculture experiments, or in those resulting in the permanent altering or sacrifice of the experimental unit.

When the experimental unit is reused in different experiments, as are plots at an agricultural experiment station or panelists in a series of food taste tests, information can be accumulated on the experimental unit over several different experiments. Some experimental units may show consistent superior yields or preferences across a single class of treatments (experimental unit-treatment interaction error) or a wide variety of classes of treatments (experimental unit error). Such historical data can provide an estimate of the experimental unit differences.

Should the same experimental units be available for a replication of the experiment, it may be advantageous to select a randomization, s', within the same g-group as the randomization, s, used in the original experiment. If only experimental unit error is present the two estimates of the covariate slope coefficient, β, should be equal. Any difference between the two estimates, $\hat{\beta}_s$ and $\hat{\beta}_{s'}$, suggests the magnitude of other error types. Such considerations are an extension. See Richards [1980].

We next illustrate the results of of Chapter III with numerical simulations.

CHAPTER IV

NUMERICAL SIMULATIONS

Section 4.1

Introduction

Chapter Four presents and discusses numerical simulations which provide examples for several of the theorems of Chapter Three. The simulations examine the three types of errors, experimental unit error, experimental unit-treatment interaction error and normally distributed random error. For each type of error, the simulations examine four cases: the combinations of (a) all treatments zero vs one specific set of nonzero treatments and (b) equal (1 to 1) vs one set of unequal (1 to 4) within-block error variances. The 1 to 1 cases have equal error variances in the two blocks; the 1 to 4 cases have the variance in the second block four times that in the first block. The 1 to 1 and 1 to 4 cases have, for each randomization, the same error variance when computed over all experimental units. Twelve sets of randomizations are examined.

Attention focuses on bias in estimation, distribution of estimators and anova terms, the closeness of the F distribution to the distribution of the calculated F-ratios, and the use of a ratio and of the noncentrality parameter $\lambda_{\beta|\mu,b,\tau}$ to weight randomizations.

The randomization test is discussed, but being a simulation, no actual experiment was performed. Thus, no randomization is privileged as the one against which to assess those not performed. The simulation is a pseudo-randomization test employing an external criterion as discussed in Section 3.6.5, (p. 174-178).

Major conclusions are that when the null hypothesis for the adjusted treatments, $H_{o:\mu,b,\tau,\beta}\tau=0$, is true, its mean square does not equal that of the error mean square for any of the three error types. The two mean square errors are close for the normally distributed random error. In Bailey's [1987, p. 712] terminology, none of these models is valid, (p. 114). Standard deviations of the F-ratios exceed their average by a factor of four for the two unit errors and a factor of 20 for normally distributed random error. For all three error types, non-equality of within-block error variance greatly reduces the means and standard deviations of the F-ratios, but has little effect upon means and standard deviations of parameter estimates. Only the experimental unit errors provide unbiased estimates for the adjusted treatments. Only this error type estimated each treatment level with

the same standard deviation. All error types provide biased estimates for the adjusted covariate slope parameter, although the experimental unit-treatment interaction error's average estimate is close to β. Only this error's estimates for the mean and block parameters are close to unbiased.

Distributions of estimators derived using the experimental unit errors do not have the normal distribution, nor do the F-ratios for this error type have the usual F distributions. Nonequality of the within-block error variances affects the F-ratios for this error type. The respective distributions for the other two errors are closer to the normal and the F distribution and are little affected by nonequal within-block error variances.

The weighting suggested by D. R. Cox [1956, p. 1148] has but modest effect for this simulation. For this single covariate case, his ratio is a linear function of the noncentrality parameter for the adjusted covariate.

Restricting randomization to the g-group with the maximum value for the noncentrality parameter of the adjusted covariate improves the power to detect a nonzero β and, modestly, a nonzero τ for the experimental unit-treatment interaction and normally distributed random errors. There is little improvement in the cases of experimental unit errors. Averaging model equation parameters over this g-group yielded means with less bias

and smaller standard deviation than did averaging over

other g-groups.   This holds for all three error types.

The singular normal distribution and related

F-distribution should be considered more closely as the

proper source of the degrees of freedom when computing

the observed significance levels, especially in the

experimental unit error cases.

Subsequent chapter subheadings are the following:

Section 4.2   Construction of Random Errors

Section 4.3   Method of Simulation

Section 4.4   Expectation and Standard Deviation

of Parameter Estimates and Anova Terms

Section 4.5   Distribution of Estimators and Anova

Terms

Section 4.6   F-Distribution Probability Plots

Section 4.7   Weighted Selection Probabilities

Section 4.8   G-groups, Noncentrality Parameters,

and OSL's

Section 4.2

Construction of Random Errors

Sections 4.2.1-4.2.3 describe the construction of
simulated values for three types of errors, experimental
unit error, experimental unit-treatment interaction error
and normally distributed error. Section 4.2.4 discusses
combining error types. Of importance is the method of
constructing the first two types of error so that
they are orthogonal to the covariate.

4.2.1 Construction of the Experimental Unit Errors

The experimental unit errors are attached to the
experimental units and are unaltered by the treatment
assigned to the experimental unit and by whatever value
the assigned treatment has. Thus, only one set of these
errors need be constructed. The following sketches
their construction; computer program meu in Appendix
F.3 provides the details.

The method of construction is such that the set of
experimental unit errors,

$$\{e_{ij}; \ i=1,2, \ \ldots, \ b, \ j=1,2, \ \ldots, p\}$$

(a) sums to zero, (b) has variance equal to 30, (c) in
vector form is orthogonal to the vector of covariates,
and (d) gives $\hat{\beta} = \beta$ in (4.2.2) below. Conditions (a) and
(b) impose Assumptions 2, $E_R(e_{ij})=0$, (p. 16), and 3,
$Var_R(e_{ij})= \sigma^2$, (p. 24) on each randomization. Condition

(a) satisfies a definition of the experimental unit

error, e_{ij}^K, of (3.3.2) which sums to zero over the exper-

imental units for each block and Assumption 9, (p. 140),

that the e_{ij} also sum to zero. For models with a random

covariate, condition (c) ensures a zero covariance

between the errors and the covariates, but not

necessarily statistical independence. Throughout, we

consider the covariates to be fixed.

Construction begins in block $i=1$ with the equation

$$Y_{1j} = \beta Z_{1j} \; . \qquad\qquad (4.2.1)$$

The intercept term, $\mu + b_i + \tau_k$, does not affect the

errors, so it is omitted. In effect, all experimental

units receive the same, $k\underline{th}$, treatment. To each point of

(4.2.1) add a random error, e_{1j}. These simulations used

the SAS function UNIFORM to generate the initial values

of all e_{ij}. Compute the initial $\hat{\beta}$ using the usual least

squares method. The steps in the remainder of this

paragraph are performed repeatedly until the obtained

slope $\hat{\beta}$ differs from the desired slope of $\beta=2.0$ by less

than 10^{-8} radians. This was achieved in fewer than 100

iterations. Compute the errors as the difference $\beta Z_{1j} -$

$\hat{\beta} Z_{1j}$. Center the errors about zero by subtracting their

mean. This ensures that they will sum to zero, (a)

above. Center the vector of Y_{1j} values about zero. Since

both β and $\hat{\beta}$ pass through the same point, $(\bar{Y}, \bar{Z})$, here
$(0,0)$, they may be treated as angles, measured in radians
in a counter-clockwise manner, from the line $Z=0$ through
the origin. Compute the usual least squares estimator of
β using the model

$$Y_{1j} = \beta Z_{1j} + e_{1j} \ . \qquad\qquad (4.2.2)$$

Subtract the arc tangent (in radians) of $\hat{\beta}$ from that of
β, $\tan^{-1}(\beta) - \tan^{-1}(\hat{\beta})$. Rotate the set of points
(Y_{1j}, Z_{1j}) by a number of radians equal to this difference.
Rotation is clockwise for positive differences and
counterclockwise for negative differences. Retain the
rotated Y_{1j} values, but replace the rotated covariates by
their original values, Z_{1j}. The least squares estimate
of the slope of this hybrid set of points is between the
original estimate and the desired value (2.0). Return to
recompute the errors.

Once a slope sufficiently close to β has been
obtained, (d) above, multiply the errors by a constant to
obtain the desired equal within-block variance, (b) above.
Repeat the iteration for block 2. Notice that these
errors are the residuals of a least squares estimation of
a regression parameter. Thus, they are orthogonal to the
covariates, (c) above. When considered as residuals from
a model whose error term has the multivariate normal
distribution, these residuals have the singular normal
distribution discussed below (2.3.3) (p. 17).

Multiplication by the appropriate constants, one per block, provides a second set of errors having unequal within-block variances.   The chosen ratio is 1:4; the variance of the errors of block two is four times that of block one.   For all three types of errors the unequal within-block variances are 12 and 48.   The large ratio makes the effects of unequal within-block variances more visible in graphic displays.   The variance over all blocks is 30 for all sets of errors.   Defining the variance to have the number of observations as its denominator eases the construction of such unequal within-block variance errors.

Table 4.2.1 displays the experimental unit errors. The magnitude of these errors is of the order of the treatment effects, $\tau_k =$ {-6.5, -3.5, 2.5 7.5}.   The goal is to present the simulation with "loud noise."   The summary statistics are based upon more decimals than are shown in the table.   Note that all means are zero, satisfying the first equality of Assumption 9, (p. 140). Also, within-block variances are 30, or 12 and 48, while both over-all variances are 30.

TABLE 4.2.1

EXPERIMENTAL UNIT ERRORS

Block	Plot	Experimental Unit Errors Within-Block Variances:	
(i)	(j)	Equal	Not Equal
1	1	5.78	3.66
1	2	-7.44	-4.70
1	3	4.70	2.97
1	4	-3.04	-1.92
Within-Block Mean		0.00	0.00
Within-Block Variance		30.00	12.00
2	1	-5.99	-7.57
2	2	8.92	11.29
2	3	-1.81	-2.29
2	4	-1.12	-1.42
Within-Block Mean		0.00	0.00
Within-Block Variance		30.00	48.00
All Units Mean		0.00	0.00
All Units Variance		30.00	30.00

Graph 4.2.1 plots the errors for block 1. The letters on the line mark the values of Y_{1j} of (4.2.1), without error. The "A" represents j=1, "B" represents j=2, and so on. Above or below the line are the Y_{1j}'s of (4.2.2) with error for the equal (E) and not equal (N) within-block variance cases. The slopes of the least squares fit of the E's and of the N's each equal 2.0, the same slope as the line through the letters A through D. The variance of the E's is 30, while that of the N's is, for this block, only 12. Thus, the N's are closer to the line than are the E's. These E's and N's become the

e_{ij}'s of (4.2.2), (p. 245) and the elements of the identical columns of E, (3.5.1), (p. 156). Section 4.6.4, (p. 315), discusses their variance-covariance matrix should the errors be considered a random vector.
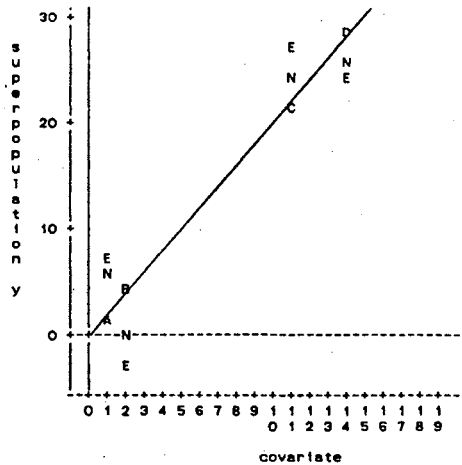
Construction of Experimental Unit Errors - Block One

Graph 4.2.1

4.2.2 Construction of the Experimental Unit-Treatment Interaction Errors

These experimental unit errors are altered by the treatment assigned to the experimental unit, but not by whatever value the assigned treatment has. Thus, one set of these errors needs to be constructed for each level of the treatment factor. With two exceptions, the process

is as described in Section 4.2.1. The conditions (a)-(d) are met for each treatment. Here, condition (a) satisfies the identity definitions of the experimental unit-treatment interaction error, η^{N1}_{ijk}, of (3.3.5), (p.27), and η^{K}_{ijk} (3.3.7), (p. 128), both of which sum to zero across the experimental units for each treatment. Condition (a) also satisfies the second equality of Assumption 9, (p. 140). That is, the η_{ijk} sum to zero over the experimental units (j) for each treatment (k). Summation to zero over the treatments for each unit was not achieved as discussed below. The computer program meuxti in Appendix F.4 provides the details.

The first exception is to repeat, within each block, the process of Section 4.2.1 for all except the last treatment level. The errors for the last treatment level are the negative of the sum of the errors for the first t-1 treatment levels. All are then multiplied by a constant to obtain the desired variance. This ensured that all sums over experimental units, one per treatment level, equal zero and that all sums over treatments, one per experimental unit, are close to zero, with the two exceptions shown in Table 4.2.2 below. The sum over both experimental units and treatments is zero. We did not find a method to provide zero sums and equal variances both over all experimental units and over all treatments.

Table 4.2.2 displays the experimental unit-
interaction errors for the equal within-block variance
case. Note that experimental unit 1 of block 1 lowers
its response by 2.1 units when treatment 1 is applied,
but increases its response by 5.24 units when treatment
4, the far right of row one, is applied. This is the
interaction effect. The errors are of the magnitude of
the treatment effects, again with the goal of presenting
the simulation with "loud noise." For each treatment,
within each block and over all experimental units, all
means equal zero and all variances equal 30, all based
upon more decimals than are shown in the table.

TABLE 4.2.2

EXPERIMENTAL UNIT-TREATMENT
INTERACTION ERRORS

Block (i)	Plot (j)	Treatments			
		k=1	k=2	k=3	k=4
1	1	-2.10	-3.44	-1.20	5.24
1	2	4.27	1.71	-0.86	-3.98
1	3	-7.98	8.07	8.64	-6.79
1	4	5.81	-6.34	-6.58	5.53
2	1	0.99	-4.18	6.21	-1.68
2	2	1.22	7.68	-8.52	-0.21
2	3	-8.69	-6.08	-0.59	8.54
2	4	6.48	2.58	2.90	-6.65

Unlike the experimental unit errors, no column of errors will appear in any one randomization. Hence, the mean and variance imposed here will not carry over to those errors appearing in any one randomization. Within each randomization, the mean of the η_{ijk} will not equal zero, nor will the variance equal 30. The variance term of the noncentrality parameters will differ for each randomization. The imposition of a variance of 30 on the columns of Table 4.2.2 is an attempt to maintain a similarity of the magnitude of the experimental unit and experimental unit-treatment interaction errors.

Alternatively, one could standardize within each randomization those errors which did appear. This would force each randomization to have a different set of experimental unit-treatment interaction errors, which violates their definition, (p. 127). On the other hand it would impose Assumptions 2, $E_R(\text{error})=0$, (p. 16), and 3, $\text{Var}_R(\text{error})= \sigma^2$, (p. 24), on each randomization. Either method gives, for this method of construction, $E_R(\text{errors})=0$ and $\text{Var}_R(\text{errors})=30$ over all randomizations.

Table 4.2.3 displays the means and variances for each experimental unit as summed across all treatments. Ideally, each would have zero mean. Across all (16) experimental unit-treatment interaction errors of any one block, the mean is zero and the variance is 30.

TABLE 4.2.3

EXPERIMENTAL UNIT-TREATMENT
INTERACTION ERRORS
SUMMED ACROSS TREATMENTS

Block (i)	Plot (j)	Mean	Variance
1	1	-1.50	11.15
1	2	1.14	9.36
1	3	1.94	62.11
1	4	-1.58	26.77
2	1	1.34	14.86
2	2	0.17	33.28
2	3	-6.82	43.52
2	4	5.31	23.56

For each treatment within each block there is a graph corresponding to Graph 4.2.1. Such graphs are similar to Graph 4.2.1 and hence are omitted.

For each treatment, a second set of experimental unit-treatment interaction errors was constructed with unequal block variances (12 and 48) as computed over all experimental units. There errors are multiples of the ones shown in Table 4.2.2 and are omitted.

4.2.3 Construction of the Normally Distributed Random Errors

The normally distributed random errors are attached to the experimental units and are unaffected by the treatment assigned to the experimental unit and by the

value of such treatment.   If these errors were identi-
cal for each randomization, they would merely add to the
experimental unit errors.   These simulations use a
different set of normally distributed random errors for
each randomization.   The SAS function RANNOR generates
the random errors which are then standardized to mean
zero and variance 30 by the SAS procedure Standardize.
The standardization is done anew for each randomization.
As with the other errors, a second set with unequal block
variances (12 and 48) was constructed.   Step Two of the
computer program cterms in Appendix F.5, provides the
details.   The program computes additional, unused, error
sets.

4.2.4   Combination of Errors

The sets of errors may be added together.   The
overall mean will remain zero while the variance will
increase.   This increase in variance makes it difficult
to compare simulations using a single type of error.
Standardization of the variances is possible, but it
destroys the definitions of the experimental unit and
experimental unit-treatment interaction errors.   Such
combined errors are an extension.

We next describe how these errors are combined to
form the observed responses for all randomizations.

Section 4.3

Method of Simulation

Section 4.3.1 describes the generation of twelve sets of all possible randomizations. Section 4.3.2 discusses the generation of analysis of variance terms and model equation estimates.

4.3.1 Generation of Randomizations and Attachment of Treatments and Error

All simulations use two blocks, each with four experimental units, one for each of the four levels of the single factor, and one covariate. Since there is but one factor, each level is also a treatment. As per (3.2.6), (p. 109), there are $N(\Xi)= (t!)^{b}= (4!)^{2}= 576$ possible randomizations, each with $bt= (2)(4) = 8$ observations, for a total of 4,608 observations. The values for the mean, blocks, covariate and, when present, the experimental unit errors, are constant for each randomization. With each new randomization, at least two of the eight experimental unit-treatment interaction errors are replaced by others from among the 16. All eight normally distributed random errors differ with each randomization. When the treatments are not zero, at least two experimental units receive different treatments as one changes from one randomization to another.

In the computer data file, each of the 4,608 observations contains the 11 variables: identification numbers for (1) randomization s, (2) block i, and (3) experimental unit j; (4) a value equal to the sum of the effects for the mean, block and covariate; (5) a value adding the effect of the assigned treatment; and (6-11) values for the three types of errors, experimental unit, experimental unit-treatment interaction and normally distributed random error, each in the two versions (a) equal within-block variances and (b) unequal within-block variances. From these, one computes twelve observed Y_{ijk} values, one for each combination of the six error terms and zero or nonzero treatment effects. Thus, twelve simulated Y_{ijk}'s appear in each of the 4,608 observations. These twelve Y_{ijk}'s are the observations for the twelve simulations discussed in Section 4.1, (p. 239). The SAS program cterms, of Appendix F.5, provides the details.

Values of the model equation parameters are $\mu=10$, $b_1=-1.5$, $b_2=1.5$, $\tau_1=-6.5$, $\tau_2=-3.5$, $\tau_3=2.5$, $\tau_4=7.5$ (or all $\tau_k=0.0$), and $\beta=2.0$. The covariate values are

$$Z' = (1,\ 2,\ 11,\ 14,\ \ 5,\ 9,\ 16,\ 19),$$

in the order Z_{ij}, (i=1,2; (j=1,2,3,4)).

4.3.2 Generation of Analysis of Variance Terms

The following process is applied to each of the twelve sets of randomizations. The SAS procedure GLM reads the eight observations for one randomization,

outputting the normally printed analysis of variance
table to a temporary computer file. This is repeated for
the remaining 575 randomizations. A SAS program then
reads the 576 analysis of variance tables, extracting
numbers of interest. Among them are (a) the mean square
error for the treatments adjusted for the mean, blocks,
and covariate, (b) the mean square error for the
covariate adjusted for the mean, blocks, and treatments,
(c) the residual mean square error, (d) the calculated
F-ratios and observed significance levels (OSL's) or
p-values for (a) and (b), and (e) the estimated values of
the adjusted k treatment effects and of the adjusted
covariate slope coefficient. These estimates use the
coding form of the X matrix,, (p. 74). A subsequent
step adds the estimates provided by the usual constraints
of Assumption 5, (p.20), $\hat{\tau}_{k;\mu,b,\tau,\beta}$ and $\hat{\beta}_{\mu,b,\tau,\beta}$. The
conversion is by (2.8.14) of Theorem 2.8.1, (p. 81).
These are the estimates herein discussed. The SAS
program glm of Appendix F.6 is typical; the other eleven
differ in the observed Y_{ijk} offered, via macros, to the
GLM procedure as the dependent variable, and a minor
change for nonzero τ.

The following sections discuss these anova terms
and estimators for the twelve cases, beginning with
their means and standard deviations.

<u>Section 4.4</u>

Expectation and Standard Deviation

of Parameter Estimates and Anova Terms

The first three subsections examine randomizations
with equal within-block error variances. The first,
4.4.1, presents expectations of estimates of the model
equation parameters. Section 4.4.2 presents the associ-
ated standard deviations. Section 4.4.3 presents means
and standard deviations for terms from the analysis of
variance tables. The final section, 4.4.4, examines
randomizations with unequal within-block error variances.
We use the shorter $\hat{\tau}$ and $\hat{\beta}$ for the estimated estimators.

Major findings for this simulation are as follows:

1. For the randomization expectation, the experimental
unit errors provide unbiased estimates for the treatment
effects as adjusted for the mean, blocks, and covariate,
but not for the adjusted covariate. The other two error
types provide nearly unbiased estimates for the adjusted
covariate, but not for the adjusted treatments effects.
The expectation is over all randomizations, $E_R(\cdot)$.

2.   Only the experimental unit errors provide equal
randomization variance for the adjusted treatment effects.

3.   The randomization average mean square for the
adjusted treatments under the null hypothesis ($\tau|\mu,b,\beta=0$)
does not equal the average for the residual mean square.

For the cases in which the errors have a variance $\sigma^2=30$ for all randomizations (experimental unit and normally distributed random errors), the average residual mean square does not equal 30.

4.   Randomization means and standard deviations for the F ratios are smaller when the within-block error variances are unequal than when they are equal. The means and standard deviations for the mean squares and model equation parameter estimates are similar in the equal and unequal within-block error cases.

4.4.1 Expectation of Parameter Estimates - Equal Block-wise Error Variances

Table 4.4.1 displays each model equation parameter value and its expectation over all randomizations. Two values are shown for the treatments. To the left of the slash is the value when the τ_k equal zero; to the right is the value when the τ_k equal the nonzero values shown in the column titled "Value". The values of the treatment effects do not affect the value of the estimates of other parameters. The selection probability is $p(s)=1/576$ for all randomizations s. For the cases of experimental unit and normally distributed random errors, the error variance is 30 for each block of each randomization. For the case of experimental unit-treatment interaction error, the error variance differs in each block and each randomization.

TABLE 4.4.1

RANDOMIZATION EXPECTED VALUE OF PARAMETER ESTIMATES
EQUAL WITHIN-BLOCK ERROR VARIANCES

Parameter:		Three Types of Error: Experimental		Normally
Name	Value	Unit	Unit-Treat. Interaction	Distributed
μ	10.0	6.05	9.94	9.36
b_1	-1.5	-.42	-1.48	-1.33
b_2	1.5	.42	1.48	1.33
τ_1	0/-6.5	0/-6.5	-.68/-7.18	.28/-6.22
τ_2	0/-3.5	0/-3.5	.88/-2.62	.23/-3.27
τ_3	0/ 2.5	0/ 2.5	-.46/ 2.04	.01/ 2.51
τ_4	0/ 7.5	0/ 7.5	.26/ 7.76	-.52/ 6.98
β	2.0	2.41	2.01	2.07

For the experimental unit errors, the estimator for
the adjusted treatments is unbiased, with respect to the
randomization distribution, both when the treatments are
zero and when they take the above nonzero values. Esti-
mates for all other parameters, μ, b and β, are biased.
Relative to the other two types of errors, the biases are
large. The bias of .41 for $\hat{\beta}$ equals the randomization

expectation of $(\mathbb{R}_{s,z}'\mathbb{R}_{s,z})^{-}\mathbb{R}_{s,z}'[e]$ as shown in (3.7.22a) of Theorem 3.7.9(4), (p. 202). In this equation, the vector e is constant for all randomizations; only the $\mathbb{R}_{s,z}$ change. The biases of .01 and .07 for the other two types of error also follow (3.7.22a).

With experimental unit-treatment interaction error, the mean estimate for the treatment effects suffers a shift, constant for each of treatment k's two values. For example, for τ_1, the mean of the estimates is too small by .68 when $\tau_1=0$ and when $\tau_1=-6.5$. Similar shifts occur for the other τ_k. The cause of this shift is unknown, but is not believed to be a consequence of the location of the errors relative to the line (above or below) in Graph 4.4.1, (p. 248). The two reasons for this belief are (a) a similar shift occurs with experimental unit-treatment interaction errors, which differ with each randomization and (b) the shift does not occur with experimental unit error. The bias for other model equation parameters is quite small, less than .07. Tests suggest that even such a small value well exceeds the accumulated computer round-off error.

With normally distributed random error, the mean estimate for the treatment effects again suffers a shift, constant for each of treatment k's two values. For example, for τ_1, the mean is too large by .28 when $\tau_1=0$ and when $\tau_1=-6.5$. The randomization bias for other model

equation parameters is small, smaller than with experimental unit error, but larger than with experimental unit-treatment interaction error.

Some yet unanswered questions are the following:

1. Why the normally distributed errors yielded biased estimates,

2. Why one error type yielded unbiased estimates for the treatments but not the other model equation parameters, while the other error types yielded biased estimates for the treatments, but nearly unbiased estimates for other parameters and

3. The origin of the constant shift observed in each treatment mean estimates using experimental unit-treatment interaction and normally distributed random error.

4.4.2 Standard Deviation of Parameter Estimates - Equal Within-Block Error Variances

Table 4.4.2 displays the standard deviation of the parameter estimates. The standard deviations are unaffected by the value of the treatment effects, hence only one value is shown for each pair of τ_k.

In all cases, the standard deviations are large, often larger than the value of the parameter being estimated. Such large values suggest that the magnitude of the simulation errors is excessive relative to the magnitude of the parameter values. It is known that

under the normally distributed random errors model, a covariate destroys the otherwise spherical variance of the estimates of the adjusted treatments. Thus, the unequal treatment standard deviations in the normal error column are to be expected. Interestingly, spherical, equal variance, estimates are provided by the experimental unit errors. We suggest that this is due to $\beta Z_{ij} + e_{ij}$, for a fixed experimental unit ij, being constant for all randomizations, while $\beta Z_{ij} + \eta_{ijk}$ and $\beta Z_{ij} + n_{ij}$ differ with each randomization. All three error types preserve the spherical variance of the estimates of the two levels of the blocking variable. This is expected, as the estimates for the b_i are not adjusted for the covariate, as are the estimates for the τ_k.

TABLE 4.4.2

RANDOMIZATION STANDARD DEVIATION OF PARAMETER ESTIMATES
EQUAL WITHIN-BLOCK ERROR VARIANCES

Parameter:		Three Types of Error Experimental:		Normally Distributed
Name	Value	Unit	Unit-Treat. Interaction	
μ	10.0	18.6	7.51	10.86
b_1	-1.5	5.07	3.19	2.96
b_2	1.5	ditto	ditto	ditto
τ_1	0/-6.5	11.84	5.18	7.55
τ_2	0/-3.5	ditto	5.23	7.21
τ_3	0/ 2.5	ditto	5.01	6.93
τ_4	0/ 7.5	ditto	4.83	6.90
β	2.0	1.93	.79	1.13

4.4.3 Expectation and Standard Deviation of ANOVA Terms— Equal Within-Block Error Variances

Table 4.4.3 displays the expectation and standard deviation of several ANOVA terms for each of the three types of errors. Shown are the cases when all treatment effects are (a) zero ($\tau=0$) and (b) when they are those of this simulation as listed in Table 4.4.1 ($\tau\neq0$). All three types of errors have a variance of 30 within each block and randomization.

TABLE 4.4.3

RANDOMIZATION EXPECTATION AND STANDARD DEVIATION
OF ANOVA TERMS
EQUAL WITHIN-BLOCK ERROR VARIANCE

ANOVA Term Expectation:	Three Types of Error Experimental:		
	Unit	Unit-Treat. Interaction	Normally Distributed
MS Residual	45.01	26.71	40.03
MS $\tau=\mathbb{O}\mid\mu,b,\beta$	49.99	24.96	39.23
MS $\tau\neq\mathbb{O}\mid\mu,b,\beta$	144.99	91.17	100.56
MS $\beta\neq0\mid\mu,b,\tau$	527.5	525.0	561.1
F $\tau=\mathbb{O}\mid\mu,b,\beta$	38.93	7.23	52.86
F $\tau\neq\mathbb{O}\mid\mu,b,\beta$	70.41	24.86	112.27
F $\beta\neq0\mid\mu,b,\tau$	136.77	107.79	248.83
Standard Deviation:			
MS Residual	39.39	22.10	32.10
MS $\tau=\mathbb{O}\mid\mu,b,\beta$	26.26	18.20	22.20
MS $\tau\neq\mathbb{O}\mid\mu,b,\beta$	71.29	52.19	65.01
MS $\beta\neq0\mid\mu,b,\tau$	336.8	379.4	400.5
F $\tau=\mathbb{O}\mid\mu,b,\beta$	150.6	41.1	1081.4
F $\tau\neq\mathbb{O}\mid\mu,b,\beta$	320.0	158.0	2292.3
F $\beta\neq0\mid\mu,b,\tau$	459.8	479.0	4172.6

Under the normally distributed random errors model
the residual mean square and, under the null hypothesis,
the mean square for $\tau=\mathbb{O}\mid\mu,b,\tau$, both divided by σ^2, have
the χ^2 distribution with mean equal to 1 and variance
equal to 2. (See Appendix A.2, (p. 382) or Searle
[1971], Corollaries 2.2-2.4, p. 58.) For $\sigma^2=30$ and the
randomization expectation, the mean values are
40.03/30.0= 1.33 and 39.23/30.0= 1.31 for normally
distributed error. Neither equals 1. Nor are the

randomization variances what one would expect;

$(32.10)^2/(30)^2 = 1030.41/900 = 1.14$ and $(22.20)^2/(30)^2 = 492.84/900 = .44$; neither equals 2. For these two mean squares, the randomization mean and variance are both smaller than the theoretical mean and variance as derived under the normally distributed random errors model. Section 4.6.4, (p. 315), discusses a possible reason for this.

Under the normally distributed random errors model, the mean, variance, and higher moments of the mean squares are defined when the null hypothesis does not hold. All involve a constant noncentrality parameter. As shown in Theorems 3.7.10 and 3.7.11 (p. 211 and 213), under the randomization model, the noncentrality parameters change with the randomization. The proper comparison of moments and variances obtained under simulation with those derived from the normal distribution model involving a noncentrality parameter is not clear. Possibly one could compute the normally distributed random errors model moments and variance for each randomization using its particular noncentrality parameter, average them using $p(s)$ and compare the result with simulation values such as those above. Such considerations are an extension.

Comparison with the mean of the F-distribution is not possible because the mean is undefined when, as is the case here, the denominator degrees of freedom is less than or equal to 2. (See Appendix A.3, (p. 383).)

When the errors represent experimental unit error or experimental unit-treatment interaction error, the randomization expectations of the residual and null hypothesis mean squares are close, 45.01 - 49.99 and 26.71 - 24.96, but clearly not equal. For these two types of errors, and zero treatment effects, the mean square for the adjusted treatments is a (randomization) biased estimator for the residual mean square error, at least for this simulation. For normally distributed random error, the expectations are nearly equal, 40.03 vs 39.23; the difference exceeds computer roundoff error.

The randomization expected value of the residual mean square does not equal the variance of the errors, 30, for any of the three error types. In particular, they do not equal 30 for the experimental unit and normally distributed random errors, both of which are standardized to have $\sigma^2 = 30$ in each randomization. See Sections 4.2.1 and 4.2.3, (p. 243 and 252); in the terminology of Bailey, (p.114), these processes are not valid.

When the treatments are nonzero, the randomization expectation of the mean square for the adjusted treatments is at least 2.5 that when the treatments are zero. Likewise, the randomization expectation of the F-ratios for nonzero treatments is about twice that for zero treatments. This suggests that, on average, the F-ratio is distinguishing between the null and alternative hypotheses, as discussed in Section 3.6.7, (p. 178).

The large standard deviation for all F-ratios indicates the instability of the F-ratios from one randomization to another. The standard deviations exceed the randomization means for the F-ratios by a factor of 3 to 4, even when the treatments are zero. This suggests that it is important to condition the conclusions of an analysis of covariance upon the randomization actually used to conduct the experiment.

4.4.4 Expectation and Standard Deviation of Anova Terms- Unequal Within-Block Error Variances

The cases for unequal within-block error variances use a ratio of 1 to 4 for the variance of block one ($\sigma^2_{i=1}=12$) to that of block two ($\sigma^2_{i=2}=48$), with a variance of 30 overall. The denominator for the variance is the number of observations, not the number of observations minus one. These values of the within-block error variances hold in each randomization for experimental unit error and normally distributed random error, but only approximately for experimental unit-treatment interaction error. For the last error type see the discussion below Table 4.2.2, (p. 250).

Unequal within-block variances have little effect on the mean or standard deviation of the estimated parameters and mean squares. Thus, there is no table for these cases. However, the mean and standard deviation of the F-ratio are affected. Table 4.4.4 displays values for these terms when the within-block error variances are unequal.

TABLE 4.4.4

RANDOMIZATION MEAN AND STANDARD DEVIATION
OF ANOVA TERMS
UNEQUAL WITHIN-BLOCK ERROR VARIANCES

Anova Term	Three Types of Error Experimental:		Normally
Expectation:	Unit	Unit-Treat. Interaction	Distributed
F $\tau=0|\mu,b,\beta$	3.39	4.25	5.88
F $\tau\neq0|\mu,b,\beta$	6.48	18.13	14.84
F $\beta\neq0|\mu,b,\tau$	20.84	96.87	62.42
Standard Deviation:			
F $\tau=0|\mu,b,\beta$	4.27	14.23	34.40
F $\tau\neq0|\mu,b,\beta$	8.68	63.07	106.52
F $\beta\neq0|\mu,b,\tau$	20.67	381.69	323.71

Both means and standard deviations of the F-ratios are smaller than for the equal within-block error variance cases. Yet, the means of the numerators and denominators of this ratio are similar to those for the equal within-block error variances cases, shown in Table 4.4.3, (p. 264). The implication is that less frequently

unequal than in the equal within-block error cases does a
large numerator appear over a small denominator in the
F-ratio.

Note that both means and standard deviations are
smaller for the experimental unit error case than for the
other two types of errors.

We now turn from the moments of these terms to
their full distributions.

Section 4.5

Distribution of Estimates and Anova Terms

Section 4.5 discusses the distribution of various parameter estimates and analysis of variance terms as provided by this simulation. With one observation per randomization, these are discrete distributions and are presented in histogram form. Histograms for the equal and unequal within-block error variance cases are similar; only the former are shown. Section 4.5.1 displays the distribution of $\hat{\tau}$; Section 4.5.2 presents that of $\hat{\beta}$. Section 4.5.3 discusses the distribution of the OSL's associated with the F-ratio for $\tau = 0$ and $\tau \neq 0$ for the experimental unit errors. Section 4.5.4 covers the experimental unit-treatment interaction errors. Section 4.5.5 describes the OSL's distribution for $\beta = 0$ for both unit errors.

Major findings for experimental unit errors follow:

1. The distributions of $\hat{\tau}$ and $\hat{\beta}$ are non-normal.

2. For $H_{0; \mu, b, \beta} \tau = 0$, the actual Type I error rate is larger than α when α is derived from the central F distribution. This suggests the F-distribution does not well approximate the randomization distribution of the F-ratios for these errors, at least in the upper tail.

3. For $H_{0; \mu, b, \beta} \tau \neq 0$ and $\alpha = .05$, the Type II error rate is high. This suggests that the magnitude of the errors has overwhelmed the model equation parameters.

For experimental unit-treatment interaction and normally distributed random errors major, findings are,

1. The distributions of $\hat{\alpha}$ and $\hat{\beta}$ are near normal.

2. For $H_{o; \mu, b, \beta} \tau = 0$, the actual Type I error rate is close to α as derived from the F distribution. This suggests that the F-distribution is a close approximation to the randomization distribution of the F-ratios for these two error types, at least in the upper tail.

3. For $H_{o; \mu, b, \beta} \tau \neq 0$ and $H_{o; \mu, b, \tau} \beta \neq 0$ with $\alpha = .05$, the Type II error rates are high, as with experimental unit errors (3 above).

Several graphs illustrate parts of Theorem 3.7.9, (p. 202); part (3) is illustrated by Graphs 4.5.9 and 4.5.10, (p. 289). Part (4) is illustrated by Graphs 4.5.3 and 4.5.4, (p. 277). Part (5) is illustrated by Graph 4.5.5, (p. 280).

4.5.1 Distribution of $\tau_1 - \hat{\tau}_1$

The distributions for the four treatment effects, the $\hat{\tau}_k$, are each similar; only the distribution of $\hat{\tau}_1$ is discussed. Further, the distribution of $\tau_k - \hat{\tau}_k$ is identical, axis values excepted, for $\tau_k = 0$ or $\tau_k \neq 0$. Hence, the only distributions presented are those of the differences $\tau_1 - \hat{\tau}_1$. These discrete distributions are shown in histogram form for the cases of experimental unit error (Graph 4.5.1) and of experimental unit-treatment interaction

error (Graph 4.5.2). The latter histogram is similar to
the one for normally distributed random error, which is
omitted.

Graph 4.5.1 displays the distribution of $\tau_k - \hat{\tau}_k$ for
treatment k=1, with the values of $\tau_1 = \{0, -6.5\}$. Points
adjacent to the 8 on the vertical, lefthand, axis
represent estimates of τ_1 in error by a deficit of
between 4 and 12 units, that is, estimates from -4 to -12
when $\tau_1 = 0$ or estimates from -10.5 to -18.5 when $\tau_1 = -6.5$.


```
                       The Randomization Model with Blocks and a Covariate
                                 Data set is  sasglm.glm002

S-err trihata                                                    Cum.              Cum.
Midpoint                                                Freq   Freq  Percent  Percent

  -72    |                                                0     0    0.00     0.00
  -64    |**                                              6     6    1.04     1.04
  -56    |                                                0     6    0.00     1.04
  -48    |                                                0     6    0.00     1.04
  -40    |**                                              6    12    1.04     2.08
  -32    |                                                0    12    0.00     2.08
  -24    |                                                0    12    0.00     2.08
  -16    |                                                0    12    0.00     2.08
   -8    |*************************************          114   126   19.79    21.88
    0    |******************************************************  270   396   46.88    68.75
    8    |***********************************           162   558   28.13    96.88
   16    |**                                             6   564    1.04    97.92
   24    |                                                0   564    0.00    97.92
   32    |**                                             6   570    1.04    98.96
   40    |                                                0   570    0.00    98.96
   48    |                                                0   570    0.00    98.96
   56    |                                                0   570    0.00    98.96
   64    |**                                             6   576    1.04   100.00
   72    |                                                0   576    0.00   100.00
   80    |                                                0   576    0.00   100.00
         ---+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--
            2  4  6  8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46
                                    Percentage
```

Graph 4.5.1
The Distribution of $\tau_1 - \hat{\tau}_1$ for
Experimental Unit Error

As hinted at by the unbiasedness of $\hat{\tau}_1$, (Table 4.4.1, p. 259), this distribution is centered at zero.

Each group of six points in the far upper and lower tails represents 1.04 percent of the distribution; they are five standard deviation units (11.84) from the mean. Each of the two groups of six points in the near upper and lower tails are another 1.04 percent of the distribution. They are some 3 standard deviation units from the mean. This distribution has more probability in its far tails than one would find in the normal distribution; it is tail heavy relative to the normal distribution. These twenty-four randomizations yielded highly erroneous estimates of τ.

From two to three standard deviations, (23.6 to 35.4) the density is zero, less than a normal distribution. Some 95% of these observations are within one standard deviation of the mean, while for a normal distribution one would find only 84% of the observations in this region. Relative to the normal distribution, this distribution is overly peaked, and tail heavy with sizeable voids in the shoulders. It is reasonably symmetric and unimodal about zero.

We believe the repetition of six closely related values is connected to repetition of certain values observed within g-groups of randomizations. These repetitions are due to patterns in the projector of $\mathbf{X}_s$. See also Graph 4.7.1, (p. 331). The extreme values of

both Graphs 4.5.1 and 4.5.3 (below) occur in the g-group
which assigns to each treatment the covariate with the
same rank in each block. For example, treatment one is
assigned to the smallest covariate in each block.
Ignoring the blocking, this group of randomizations is
the least balanced on the covariate, as discussed in
Section 4.7.1, (p. 321).

Graph 4.5.2 shows the distribution of $\tau_1 - \hat{\tau}_1$ for the
case of experimental unit-treatment interaction error.
The case of normally distributed random error is similar,
and is omitted.

The Randomization Model with Blocks and a Covariate
Data set is sasglm.glm004

S-err trihata Midpoint		Freq	Cum. Freq	Percent	Cum. Percent
-38		0	0	0.00	0.00
-34		0	0	0.00	0.00
-30		1	1	0.17	0.17
-26		0	1	0.00	0.17
-22		1	2	0.17	0.35
-18	•	3	5	0.52	0.87
-14	•	2	7	0.35	1.22
-10	•••	9	16	1.56	2.78
-6	••••••••••••••••••	52	68	9.03	11.81
-2	••	186	254	32.29	44.10
2	•••	200	454	34.72	78.82
6	••••••••••••••••••••••••••••••••	90	544	15.63	94.44
10	••••••••	21	565	3.65	98.09
14	••	6	571	1.04	99.13
18	•	2	573	0.35	99.48
22		1	574	0.17	99.65
26		0	574	0.00	99.65
30	•	2	576	0.35	100.00
34		0	576	0.00	100.00
38		0	576	0.00	100.00

```
----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
    2    4    6    8   10   12   14   16   18   20   22   24   26   28   30   32   34
                              Percentage
```

Graph 4.5.2
The Distribution of $\tau_1 - \hat{\tau}_1$ for
Experimental Unit-Treatment Interaction Error

This distribution is also more peaked than the
normal distribution. Some 67 percent of the observations
are between -4 and +4 or (+-).77 standard deviation
units; a normal distribution would have only 56 percent
of its observations in this region. However, it is not
tail heavy. Beyond -12 and +12 or (+-)2.32 standard
deviation units, fall but 1.7 percent of the observa-
tions. This is close to the 2 percent found in this
region under the normal distribution curve.

Experimental unit-treatment interaction error
provides a distribution of $\hat{\tau}_1$ closer to the normal
distribution than does experimental unit error. With
experimental unit error, estimates of τ_k are either
better, closer to τ_k, or much worse, farther from τ_k,
than one would expect had the error been normally
distributed.

4.5.2 Distribution of $\beta - \hat{\beta}$

As per Theorem 3.7.9(4), (p. 202), when the only
error is experimental unit error, each estimate of β
repeats itself (t!) = 4! = 24 times. Graph 4.5.3
displays these clusters of estimates, each subtracted
from β=2.0.

As expected from Table 4.4.1, Graph 4.5.3 is centered
at 2.0 - 2.41 = -.41. The one outlying cluster is 4.2
percent of the observations and represents an estimate of
$\hat{\beta}$=11.3, an error of some magnitude (β=2.0). Randomiza-

tions in this cluster, or g-group, should be avoided. It is the group which assigns each treatment to the covariate with the same rank in each block. It is the least balanced of all g-groups. The value of the noncentrality parameter $\lambda_{\beta|\mu, b, \beta}$ associated with this g-group is 10, the minimum of all 24 values of $\lambda_{\beta|\mu, b, \tau}$. The maximum value of this noncentrality parameter is 990. Theorem 3.7.15, (p. 219), suggests selecting the randomization used to conduct the experiment from among those randomizations with the maximum value of $\lambda_{\beta|\mu, b, \tau}$; doing so would avoid estimates in the outlying cluster of Graph 4.5.3.

An alternate method of creating experimental unit errors yielded a set of errors with characteristics of both this and experimental unit-treatment interaction errors. The mix was some 90% the former and 10% the latter. The plot corresponding to Graph 4.5.3 based on these errors did not show an outlaying cluster as seen in Graph 4.5.3.

About 95 percent of the observations fall within one standard deviation (1.93) of the mean. For a normal distribution, one would expect only 68 percent of the observations to lie within this region. For this simulation, the distribution of the $\hat{\beta}$, using experimental unit errors, is not the normal distribution. Estimates are either better or far worse than one would expect had the distribution been normal.

The Randomization Model with Blocks and a Covariate
Data set is sasglm.glm002

S-err betahat Midpoint		Freq	Cum. Freq	Percent	Cum. Percent
-9.9		0	0	0.00	0.00
-9.3	••••••••	24	24	4.17	4.17
-8.7		0	24	0.00	4.17
-8.1		0	24	0.00	4.17
-7.5		0	24	0.00	4.17
-6.9		0	24	0.00	4.17
-6.3		0	24	0.00	4.17
-5.7		0	24	0.00	4.17
-5.1		0	24	0.00	4.17
-4.5		0	24	0.00	4.17
-3.9		0	24	0.00	4.17
-3.3		0	24	0.00	4.17
-2.7		0	24	0.00	4.17
-2.1		0	24	0.00	4.17
-1.5		0	24	0.00	4.17
-0.9	•••••••••••••••••••••	72	96	12.50	16.67
-0.3	•••	192	288	33.33	50.00
0.3	••	216	504	37.50	87.50
0.9	•••••••••••••••••••••	72	576	12.50	100.00
1.5		0	576	0.00	100.00

```
    2  4  6  8  10 12 14  16 18 20 22 24 26 28 30 32 34 36
```
Percentage

Graph 4.5.3

The Distribution of $\beta - \hat{\beta}$ for
Experimental Unit Error

Graph 4.5.4 shows the distribution of $\hat{\beta}$ using

experimental unit-treatment interaction errors. The

graph using normally distributed random errors is similar

and, thus, omitted. As per Table 4.4.1, this distribu-

tion is centered at $\beta-\hat{\beta}$ = 2.0-2.007 = .007. Although not

always indicated by an asterisk, there are a few

observations (.5%) five standard deviation units from the

mean. Within (+-).75 (one standard deviation) lie 78% of

the observations. One would expect 66% of the observa-

tions in this region if the distribution were the normal

distribution. Within (+-)1.50 (two standard deviations)

lie 95% percent of the observations, about what one would

expect under the normal distribution. In this simulation,

both experimental unit-treatment interaction error and

normally distributed random error yielded a near normal

distribution for the $\hat{\beta}$'s.

```
                        The Randomization Model with Blocks and a Covariate
                                 Data set is  sasglm.glm004

  S-err betahat                                                 Cum.              Cum.
  Midpoint                                              Freq    Freq  Percent   Percent

    -4.5   |                                              1       1     0.17     0.17
    -4.0   |                                              0       1     0.00     0.17
    -3.5   |•                                             2       3     0.35     0.52
    -3.0   |                                              0       3     0.00     0.52
    -2.5   |•                                             2       5     0.35     0.87
    -2.0   |•••                                           9      14     1.56     2.43
    -1.5   |•••••                                        13      27     2.26     4.69
    -1.0   |•••••••••••                                  29      56     5.03     9.72
    -0.5   |•••••••••••••••••••••••••••••••••••••••••••• 135     191    23.44    33.16
     0.0   |••••••••••••••••••••••••••••••••••••••••••••••• 198  389    34.38    67.53
     0.5   |•••••••••••••••••••••••••••••••••••••••••    122     511    21.18    88.72
     1.0   |•••••••••••••••••                            43     554     7.47    96.18
     1.5   |••••                                         12     566     2.08    98.26
     2.0   |••                                            5     571     0.87    99.13
     2.5   |•                                             2     573     0.35    99.48
     3.0   |                                              0     573     0.00    99.48
     3.5   |•                                             1     574     0.17    99.65
     4.0   |                                              0     574     0.00    99.65
     4.5   |•                                             1     575     0.17    99.83
     5.0   |•                                             1     576     0.17   100.00
           ----+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+--
              2.  4   6   8  10  12  14  16  18  20  22  24  26  28  30  32  34
                                       Percentage
```

Graph 4.5.4

The Distribution of $\beta - \hat{\beta}$ for

Experimental Unit-Treatment Interaction Error

4.5.3 Distribution of the OSL for $H_o: \tau = 0$ –

Experimental Unit Error

As proved in Part 5 of Theorem 3.7.9, (p. 203), when the null hypothesis $\tau = 0$ is true, the F-ratios, and hence the OSL's, for the treatments as adjusted for the mean, blocks, and covariate(s) are clustered. When the error is only experimental unit error, each cluster collapses to a single point. Graph 4.5.5 displays the 24 collapsed clusters of this discrete distribution, each containing 24 identical values. The automatic scaling has combined clusters with similar values; note that the values of the frequency column (Freq) are all multiples of 24.

By the discrete analogue of the probability integral transformation [Mood, Graybill, Boas, 1974, p. 202], if the F-ratio has the F-distribution, then the tail area or OSL as computed using the F-distribution for these F-ratios has the uniform distribution over [0,1]. Graph 4.5.5 shows excess probability in the tails, near zero and near one, relative to what one would observe for the uniform distribution. Thus, these F-ratios do not follow the usual F-distribution. Graph 4.6.1, (p. 295), will show in a more direct manner that these F-ratios do not follow the F-distribution with the usual degrees of freedom.

The Randomization Model with Blocks and a Covariate
Data set is sasglm.glm002

OSL T_ M,B,Z Midpoint		Freq	Cum. Freq	Percent	Cum. Percent
0.00	●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●	48	48	8.33	8.33
0.05	●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●	48	96	8.33	16.67
0.10	●●●	72	168	12.50	29.17
0.15	●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●	48	216	8.33	37.50
0.20	●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●	48	264	8.33	45.83
0.25		0	264	0.00	45.83
0.30	●●●●●●●●●●●●●●●●●●	24	288	4.17	50.00
0.35	●●●●●●●●●●●●●●●●●●	24	312	4.17	54.17
0.40		0	312	0.00	54.17
0.45		0	312	0.00	54.17
0.50	●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●	48	360	8.33	62.50
0.55		0	360	0.00	62.50
0.60		0	360	0.00	62.50
0.65		0	360	0.00	62.50
0.70		0	360	0.00	62.50
0.75		0	360	0.00	62.50
0.80	●●●●●●●●●●●●●●●●●●	24	384	4.17	66.67
0.85	●●●	72	456	12.50	79.17
0.90	●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●	48	504	8.33	87.50
0.95	●●●	72	576	12.50	100.00

```
----+----+----+----+----+----+----+----+----+----+----+----+---
    1    2    3    4    5    6    7    8    9   10   11   12
                        Percentage
```

Graph 4.5.5

The Distribution of OSL'S for $F(\tau=0|\tau=0)$
Experimental Unit Error

One consequence of the heavy tails of Graph 4.5.5 is that the Type I error rate exceeds α. Since $\tau=0$, a Type I error is committed when the OSL, computed under the assumption of a central F-distribution, is less than α, say, $\alpha=.075$. Graph 4.5.5 shows 16.7 percent ("Cum Percent" column) of the randomizations will yield an OSL less than or equal to .075 (the category boundry on the vertical, lefthand, axis between the midpoints of .05 and .10). The actual Type I error rate is .167, not .075.

Graph 4.5.6 presents the OSL's when the treatments
take the nonzero values of this simulation.

The Randomization Model with Blocks and a Covariate
Data set is saaglm.glm102

OSL T_ M.B.Z Midpoint		Freq	Cum. Freq	Percent	Cum. Percent
0.025	••	115	115	19.97	19.97
0.075	••	84	199	14.58	34.55
0.125	•••••••••••••••••••••••	41	240	7.12	41.67
0.175	••••••••••••••••	29	269	5.03	46.70
0.225	••••••••••••••	25	294	4.34	51.04
0.275	••••••••	11	305	1.91	52.95
0.325	•••••••••••	18	323	3.13	56.08
0.375	••••••••••••	22	345	3.82	59.90
0.425	••••••••••••••••••	32	377	5.56	65.45
0.475	••••••••••••••••••••	37	414	6.42	71.88
0.525	••••••••••••••••	28	442	4.86	76.74
0.575	•••••••••••	20	462	3.47	80.21
0.625	••••••••••	18	480	3.13	83.33
0.675	••••••••••	18	498	3.13	86.46
0.725	•••••••••••••	23	521	3.99	90.45
0.775	•••••••••	17	538	2.95	93.40
0.825	•••••••	11	549	1.91	95.31
0.875	••••••••	15	564	2.60	97.92
0.925	•••••	9	573	1.56	99.48
0.975	••	3	576	0.52	100.00

```
----+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
    1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
                              Percentage
```

Graph 4.5.6
The Distribution of OSL'S for $F(\tau=\mathbb{O}|\tau\neq\mathbb{O})$
Experimental Unit Error

Since $\tau\neq\mathbb{O}$, Part 5 of Theorem 3.7.9, (p. 203), does
not apply. Whatever clustering is provided by the
denominator of this F-ratio is dispersed by the numera-
tor, more specifically, by the changes in the projector
for $\Sigma = [\Sigma_\mu|\Sigma_b|\mathbb{Z}]$ within each g-group.

The visible difference between Graphs 4.5.5 ($\tau=0$) and 4.5.6 ($\tau\neq0$) is a shift to more frequent lower valued OSL's (higher valued F-ratios). This suggests that the F-ratio is distinguishing between the two cases of $\tau=0$ and $\tau\neq0$.

By the probability integral transformation, this distribution would be the uniform, if the proper distribution were used to compute the tail area. These OSL's are from the output of SAS's PROC GLM and are computed under the assumption of a true null hypothesis and a full-rank variance-covariance matrix for the errors. Thus, the usual central F-distribution provided these OSL's. However, since the null hypothesis is false, it is the non-central F-distribution (perhaps with adjusted degrees of freedom) which would provide a uniform distribution if the F-ratios approximate an F-distribution. Section 4.6.4, (p. 315), discusses adjusting the degrees of freedom. One would not expect this graph to resemble the uniform distribution.

For this case, one wants to reject the null hypothesis that $\tau=0$. One commits a Type II error, with probability β_{II}, by accepting the erroneous null hypothesis that $\tau=0$. This occurs when the observed F-ratio is less than that tabled F-ratio which provides an OSL equal to α, say, .05, for the central F-distribution. Graph 4.5.6 shows that some 80 percent of the randomizations will result in a Type II error for nonzero τ with a

near-zero noncentrality parameter λ_s. The actual power

$(1-\beta_{II})$ for the nonzero τ of this simulation requires

using the λ_s corresponding to randomization s.   These λ_s

change with each randomization, as displayed in Graph

3.6.1, (p. 178).   This adjustment appears to give power

greater than 1-.8=.2, but the power appears to be low

enough to suggest altering one's strategy, (p.148).

4.5.4   Distribution of the OSL for $H_o: \tau = 0$ - Experimental

Unit-Treatment Interaction Error

The experimental unit-treatment interaction errors

change with each randomization, dispersing the clustering

effect of projectors identical within each of the

g-groups.   (See Theorem 3.7.9(5), p. 203.)   The heavy

tails seen with experimental unit error are not visible

in Graph 4.5.7. This distribution appears uniform, but

has not been subjected to a statistical test.   As it was

computed using the central F-distribution, the logic of

the probability integral transformation suggests that

these F-ratios take the central F-distribution.   The

graph for normally distributed random error is similar

and omitted.   Graph 4.6.1 below, (p. 295), also

suggests a close approximation to the F-distribution.

The Randomization Model with Blocks and a Covariate
Data set is sasglm.glm004

OSL_T_M,B,Z Midpoint		Freq	Cum. Freq	Percent	Cum. Percent
0.025	••	28	28	4.86	4.86
0.075	•••••••••••••••••••••••••••••••••••••••	27	55	4.69	9.55
0.125	••	30	85	5.21	14.76
0.175	••••••••••••••••••••••••••••••••••••	22	107	3.82	18.58
0.225	••	28	135	4.86	23.44
0.275	••	30	165	5.21	28.65
0.325	•••••••••••••••••••••	14	179	2.43	31.08
0.375	•••	36	215	6.25	37.33
0.425	•••••••••••••••••••••	14	229	2.43	39.76
0.475	••	34	263	5.90	45.66
0.525	••	31	294	5.38	51.04
0.575	•••••••••••••••••••••••••••••••••••••••	29	323	5.03	56.08
0.625	••	31	354	5.38	61.46
0.675	••••••••••••••••••••••••••••••••••••••	27	381	4.69	66.15
0.725	•••••••••••••••••••••••••••••••••••	24	405	4.17	70.31
0.775	•••	33	438	5.73	76.04
0.825	••	42	480	7.29	83.33
0.875	••	31	511	5.38	88.72
0.925	•••••••••••••••••••••••••••••••••••••••	30	541	5.21	93.92
0.975	•••	35	576	6.08	100.00

```
         1       2       3       4       5       6       7
                        Percentage
```

Graph 4.5.7

The Distribution of OSL'S for $F(\tau=0|\tau=0)$
Experimental Unit-Treatment Interaction Error

One consequence of the near uniformity of Graph 4.5.7 is that the actual (randomization) Type I error rate is close to α. Since $\tau=0$, a Type I error is committed when the OSL, computed under the assumption of a central F-distribution, is less than α, say, $\alpha=.125$. Graph 4.5.7 shows 9.55+(5.21/2) = 12.2 percent of the randomizations will yield an OSL less than or equal to

$\alpha = .125$. The actual Type I error rate is quite close to such stated values of α, for $\alpha > .05$.

Graph 4.5.8 displays the distribution of the OSL's when the treatments take the nonzero treatment τ values of this simulation. As in Graph 4.5.7, the errors are experimental unit-treatment interaction errors. As with Graph 4.5.6, these OSL's are computed assuming a central F distribution. Thus, the distribution is not uniform. The distribution using normally distributed random error is similar to Graph 4.5.8 and is omitted.

The discussion of power following Graph 4.5.6, applies here, except that adjustments to the degrees of freedom appear necessary for the normally distributed random error case, but not the experimental unit-treatment interaction error case. For a near-zero noncentrality parameter and $\alpha = .125$, Graph 4.5.8 shows some $100 - 32.81 \approx 67$ percent of the randomizations will result in a Type II error.

The Randomization Model with Blocks and a Covariate
Data set is sasglm.glm104

OSL T_ M.B.Z Midpoint		Freq	Cum. Freq	Percent	Cum. Percent
0.00	••••••••••••••••••••••••••	41	41	7.12	7.12
0.05	••	76	117	13.19	20.31
0.10	••	72	189	12.50	32.81
0.15	•••	62	251	10.76	43.58
0.20	••••••••••••••••••••••••••••••••••••••	58	309	10.07	53.65
0.25	•••••••••••••••••••••••••	38	347	6.60	60.24
0.30	•••••••••••••••••••••••••••	41	388	7.12	67.36
0.35	•••••••••••••••••••••	32	420	5.56	72.92
0.40	•••••••••••••••••••••	32	452	5.56	78.47
0.45	••••••••••••••••••	28	480	4.86	83.33
0.50	••••••••••••••	21	501	3.65	86.98
0.55	•••••••••	14	515	2.43	89.41
0.60	•••••••••••	17	532	2.95	92.36
0.65	••••••••	13	545	2.26	94.62
0.70	••••••••	13	558	2.26	96.88
0.75	••••	5	563	0.87	97.74
0.80	•••	4	567	0.69	98.44
0.85	•••••	6	573	1.04	99.48
0.90	•	1	574	0.17	99.65
0.95	••	2	576	0.35	100.00

```
----+----+----+----+----+----+----+----+----+----+----+----+----+
    1    2    3    4    5    6    7    8    9   10   11   12   13
                         Percentage
```

Graph 4.5.8

The Distribution of OSL'S for $F(\tau=0|\tau\neq0)$
Experimental Unit–Treatment Interaction Error

4.5.5 Distribution of the OSL for $H_0: \beta=0$

Graph 4.5.9 displays the distribution of the OSL's

corresponding to $F(\beta|\mu,b,\tau)$, the test of the covariate as

adjusted for the mean, blocks, and treatments, for the

case of experimental unit error. As per Part 4 of Theorem

3.7.9, (p. 202), these OSL's are clustered. Each cluster

contains $N(\xi)t!=1(24)$ randomizations. The automatic

scaling has grouped clusters with similar values.

Graph 4.5.10 displays the same OSL's for the case of experimental unit-treatment interaction error. Note the change in both scales from Graph 4.5.9. For these errors, the clusters remain disperse as per Theorem 3.7.9(4), (p. 202). Both sets of OSL's are derived under the null hypotheses of $\beta=0$, that is, under the central F distribution. As $\beta=2.0$, the proper distribution is the non-central F distribution, assuming that the experimental unit errors approximate the normal distribution. Thus, neither graph is expected to display the uniform distribution.

In the following cases, Graph 4.5.9 and 4.5.10, one wants to reject the null hypothesis that $\beta=0$ and does so when the OSL is larger than, say, $\alpha=.125$. One commits a Type II error, accepting the erroneous null hypothesis, should the OSL be greater than α, say greater than .125. For a near-zero noncentrality parameter and $\alpha=.125$, Graph 4.5.9 (experimental unit error) shows that some

$$100-[62.5+(16.67/2)] = 100-70.83 = 30$$

percent of the randomizations will result in a Type II error. For experimental unit treatment interaction error, Graph 4.5.10 shows that some

$$100-[73.1+(7.64/2)] = 100-76.9 = 23$$

percent of the randomizations will result in a Type II error. The graph for normally distributed random error resembles Graph 4.5.10 and is omitted.

The discussion of power following Graph 4.5.6

applies here, except that adjustments to the degrees of

freedom appear necessary for the normally distributed

random error case, but not the experimental unit-

treatment interaction error case.

The Randomization Model with Blocks and a Covariate
Data set is sasglm.glm002

OSL Z_ M.B.T Midpoint		Freq	Cum. Freq	Percent	Cum. Percent
-0.025		0	0	0.00	0.00
0.000	●●	96	96	16.67	16.67
0.025	●●	120	216	20.83	37.50
0.050	●●●●●●●●●●●●●●	24	240	4.17	41.67
0.075	●●	96	336	16.67	58.33
0.100	●●●●●●●●●●●●●●	24	360	4.17	62.50
0.125	●●	96	456	16.67	79.17
0.150	●●●●●●●●●●●●●●	24	480	4.17	83.33
0.175		0	480	0.00	83.33
0.200	●●●●●●●●●●●●●●	24	504	4.17	87.50
0.225	●●●●●●●●●●●●●●●●●●●●●●●●●●	48	552	8.33	95.83
0.250		0	552	0.00	95.83
0.275		0	552	0.00	95.83
0.300		0	552	0.00	95.83
0.325		0	552	0.00	95.83
0.350		0	552	0.00	95.83
0.375		0	552	0.00	95.83
0.400		0	552	0.00	95.83
0.425	●●●●●●●●●●●●●●	24	576	4.17	100.00
0.450		0	576	0.00	100.00

```
  1  2  3  4  5  6  7  8  9  10 11 12 13 14 15 16 17 18 19 20
                          Percentage
```
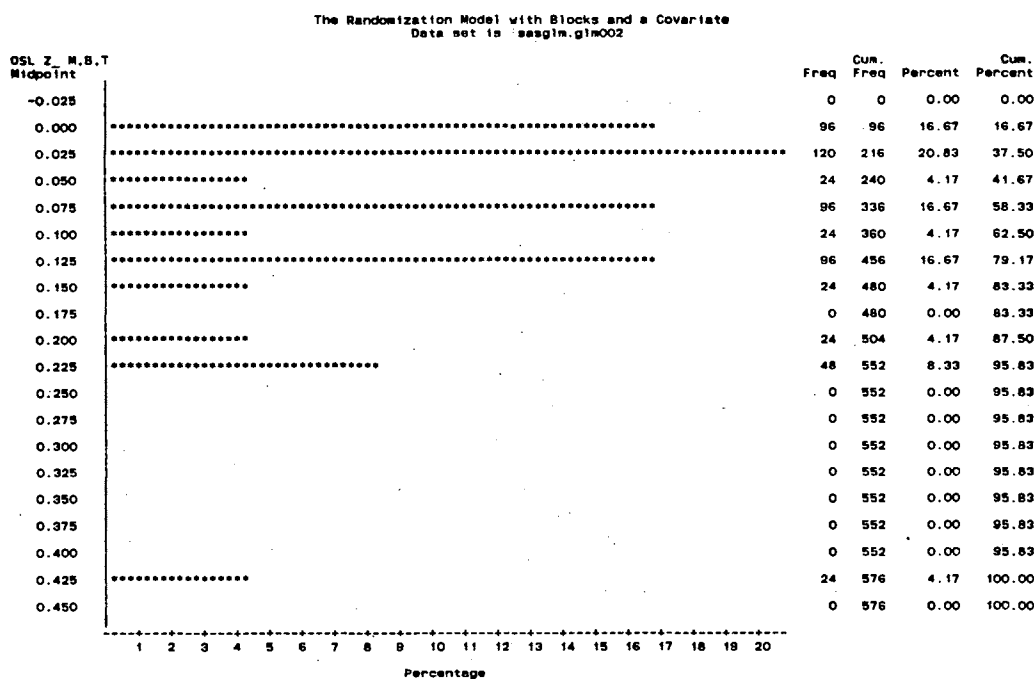
Graph 4.5.9
The Distribution of OSL'S for F(β=0 | β≠0)
Experimental Unit Error

The Randomization Model with Blocks and a Covariate
Data set is sasglm.glm004

OSL Z_ M.B.T Midpoint		Freq	Cum. Freq	Percent	Cum. Percent
0.025	••	328	328	56.94	56.94
0.075	•••••••••••••••••	93	421	16.15	73.09
0.125	•••••••••	44	465	7.64	80.73
0.175	•••••	27	492	4.69	85.42
0.225	••	13	505	2.26	87.67
0.275	••	10	515	1.74	89.41
0.325	••	9	524	1.56	90.97
0.375	••	9	533	1.56	92.53
0.425	•	5	538	0.87	93.40
0.475	•	3	541	0.52	93.92
0.525	•	3	544	0.52	94.44
0.575	•	5	549	0.87	95.31
0.625	•	7	556	1.22	96.53
0.675	•	3	559	0.52	97.05
0.725	•	4	563	0.69	97.74
0.775	•	3	566	0.52	98.26
0.825		1	567	0.17	98.44
0.875		1	568	0.17	98.61
0.925	•	3	571	0.52	99.13
0.975	•	5	576	0.87	100.00

```
     ----+----+----+----+----+----+----+----+----+----+----+--
         5   10   15   20   25   30   35   40   45   50   55
                            Percentage
```

Graph 4.5.10

The Distribution of OSL'S for $F(\beta=0 | \beta \neq 0)$
Experimental Unit-Treatment Interaction Error

The graphs for experimental unit-treatment interaction and normally distributed random errors, Graphs 4.5.8 and 4.5.10, resemble, respectively, two and one parameter gamma functions. The latter is the exponential function. Such a relationship has not been investigated.

We next derive the OSL's from the central and noncentral F-distributions and display them in probability plots.

Section 4.6

F-Distribution Probability Plots

Section 4.6.1 describes the F-distribution paper probability plots. Section 4.6.2 presents plots for the three types of errors with equal within-block error variances. Section 4.6.3 presents plots with unequal within-block error variances. Section 4.6.4 discusses potential problems with these plots.

Major findings are that the F distribution with the usual degrees of freedom is a poor approximation in the case of experimental unit error, but is a reasonable approximation for the other two error types. Also, the case of experimental unit error is sensitive to the equality or nonequality of the within-block error variances, while the case of experimental unit-treatment interaction error is not and the case of normally distributed random error is slightly sensitive.

4.6.1 Display of Results using Probability Paper

Normal probability paper provides a visual check on the normality of a set of data. If the observations are from a normal distribution, proper plotting on such special paper will yield a straight line. Snedecor and Cochran [1980, Ed. 7, p. 59-63] discuss the construction and interpretation of such plots. On probability paper for the normal distribution, the vertical axis is scaled in units of the lower tail probabilities from the normal

distribution. The horizontal axis is scaled in standard
distribution units.

The graphs presented below are an F-distribution
analog of normal probability paper, with axis switched
and rescaled. If the observations are from an F
distribution, proper plotting on such graphs will yield a
straight line from lower left to upper right along the
diagonal. For the upper half of the diagonal line,
points below the diagonal indicate a distribution heavy
in its upper tail. Points above the diagonal indicate a
distribution light in its upper tail. For the lower half
of the diagonal line, points *above* the line indicate a
distribution heavy in its lower tail. Points *below* the
diagonal indicate a distribution light in its lower tail.
The remainder of this section details the reading of
this type of graph.

The horizontal, bottom, axis is the upper tail
probability, the OSL, that is, the area under the
F-distribution curve to the right of the calculated
F-ratio. Its values are calculated by assuming that the
F-ratio follows the F distribution, with known numerator
and denominator degrees of freedom and known non-
centrality parameter. To present the upper tail area
value in its usual righthand side position, the
horizontal axis is reversed, the maximum OSL value, 1.0,
is to the left and the minimum OSL value, 0.0, is to the
right. The horizontal, bottom, axis of this probability

paper corresponds to the vertical axis of the usual
normal probability paper.

The vertical, lefthand, axis of this probability
paper is the rank of the OSL, more properly, the rank
divided by the number of observations, ranked from the
smallest (0.0) to the largest (1.0). The smallest OSL
has a rank of 0, the median OSL has a rank of .5 and the
largest OSL has a rank of 1. The vertical, lefthand,
scale is also reversed, with 1.0 at the bottom and 0.0 at
the top. This brings the desired diagonal to the
positive 45 degree position.

One way to read the graph is to hold a ruler
vertically, up and down, along the right-hand side of the
graph, at the 0.0 mark on the lower, horizontal, axis.
Move the ruler to the left to, say, the .2 value on the
horizontal axis. You have covered 20 percent of the
upper tail area, the OSL, of an F distribution.

If the plotted values are based upon statistics
which have the F distribution, then the line connecting
them should be located in such a position that 20 percent
of the plotted points, as read on the lefthand, vertical
axis, are above the 45 degree line. The lefthand scale
is read to the left of the point where the line crosses
the ruler. If the line of plotted points is below the 45
degree line, then there are excessive observations with
OSL's less than or equal to .2. There are too many large
calculated F ratios for their distribution to be the F

distribution.   The actual distribution is tail heavy

relative to the F distribution.   In such a case, in some

other region, the line of plotted points will be above

the 45 degree line.   This region is usually near the

other tail, near the left, the 1.0 end of the horizontal

axis.

One reason for plotting the OSL's instead of the

F-ratios is to maintain a common scale for all sets of

randomizations.   The F-ratios vary widely with values

from near zero to the tens of thousands.   Such large

ranges force most values into an uninformative clump at

one corner of a graph of the F-ratios.   A second reason

is that the OSL incorporates the degrees of freedom and

the noncentrality parameter, permitting OSL's computed

from central and noncentral F distributions, perhaps with

different degrees of freedom, to be displayed on the same

graph.   Finally, one is usually interested in the OSL, or

p-value, not the actual value of the statistic.

When $\tau=0$, the central F-distribution provides the

OSL's or upper tail probabilities.   For $\tau \neq 0$ and $\beta \neq 0$, the

noncentral F-distribution provides the OSL's.   The

noncentrality parameter changes with each randomization,

as per Theorems 3.7.10, (p. 211) and 3.7.11, (p. 213).

The variance term, σ^2, of the noncentrality parameters is

the variance of the errors for the randomization.   It is

constant for the experimental unit and normally

distributed random errors.   As discussed in Section 4.2,

(p. 243-253), the experimental unit errors are constant

for all randomizations; the normally distributed random

errors are standardized within each randomization to the

same variance.   The variance changes with each randomi-

zation for the experimental unit-treatment interaction

errors.   This may allow the F-distribution to adjust for

differences in the variances of the errors from one

randomization to the next, thus, in effect, standardizing

the variances.   The effect of changes in σ^2 appears to be

much less than that of the randomization itself, but

further examination is needed.   All OSL's are calculated

using the usual integer degrees of freedom.   Section

4.6.4 discusses the degrees of freedom as derived from

the singular-normal distribution.   Computer program probp

of Appendix F.10 details the construction of the

following graphs.

4.6.2 F Distribution Plots for Three Error Types

Graphs 4.6.1 through 4.6.6 come in pairs.   The

three pairs present the F-distribution plots for the

cases of $\tau = 0$, $\tau \neq 0$ and $\beta \neq 0$.   The first of each pair

shows the entire range of the OSL, (1.0 to 0.0).   The

second zooms in to the 20 percent area of the upper tail,

with OSL's from 0.2 to 0.0.

The plots of Graphs 4.6.1 and 4.6.2 are derived from

the hypothesis $H_{o; \mu, b, \beta} \tau = 0$.   When $\tau = 0$, they suggest that

the F-ratios, which produced the plotted OSL's, follow

the central F distribution, for experimental unit-
treatment interaction error (i) and for normally

distributed random error (n), but not for experimental

unit error (u).   Points for the experimental unit error

(u) represent clusters of equal-valued g-groups, as per

Theorem 3.7.9(5), (p. 203).   Their off-diagonal positions

indicate excess probability in both tails.   Thus, the

actual Type I error probability is larger than the stated

α.   Note that the points for normally distributed random

error (n) are not on the diagonal.   This is more clearly

seen in Graph 4.6.2 below.   Section 4.6.4 discusses this.
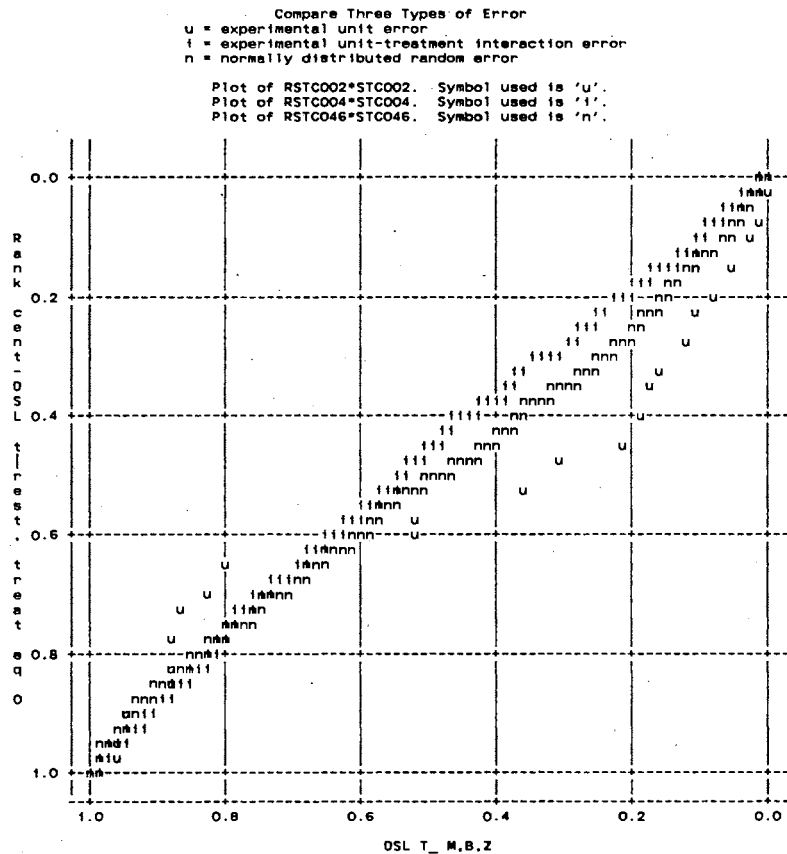
Graph 4.6.1
F Distribution Plot: $\tau = 0$
Entire Range

```
                      Compare Three Types of Error
                u = experimental unit error
                i = experimental unit-treatment interaction error
                n = normally distributed random error

                  Plot of RSTC002*STC002.  Symbol used is 'u'.
                  Plot of RSTC004*STC004.  Symbol used is 'i'.
                  Plot of RSTC046*STC046.  Symbol used is 'n'.

      0.00 +-+------------------+----------------+--------------------+in----
           | |                  |                |                 nn|
           | |                  |                |                 in|
   R       | |                  |                |               iinn |
   a       | |                  |                |               ii n |
   n       | |                  |                |              i nn   u
   k       | |                  |                |            ii inn
   c       | |                  |                |          ii i  n
   e 0.05 +-+------------------+----------------+--------ii--i--nn-----+
   n       | |                  |                |       ii     n  n
   t       | |                  |                |      ii    n  n      u
   -       | |                  |                |     ii   nn
   O       | |                  |                |  ii i    n
   S       | |                  |                | iii     nn|
   L       | |                  |              iii       nn|
           | |                  |             i         n  n
   t 0.10 +-+------------------+----------ii--+-------nn---+-----u-----+
   r       | |                  |          i         nnn
   e       | |                  |         i  i       nn
   s       | |                  |        i  i       nn
   t       | |                  |       ii         n
   .       | |                  |      ii        nn
           | |                  |    ii ii      nn
   t       | |                  |    ii        nn
   r       | |                  |   ii       n  n
   e       | |                  |  i        nn  i            u
   a 0.15 +-+------------i-i+--------n-+--i--+------------------+
   t       | |            ii       nnn
   e       | |           i        nn
   q       | |          i i      nnn
   o       | |         ii       nn
           | |        i        n
           | |       i i      nn
           | |      i i     nn
           | |    ii      n  n            u
     0,20 +-+--i-------nn-+----------------+--------------------+
           | |          nn|
           | i          n-+
           +-+-------------+----------------+--------------------+
            +-+------------+----------------+--------------------+
           0.20      0.15         0.10          0.05          0.00

                        OSL T_ M.B.Z
```
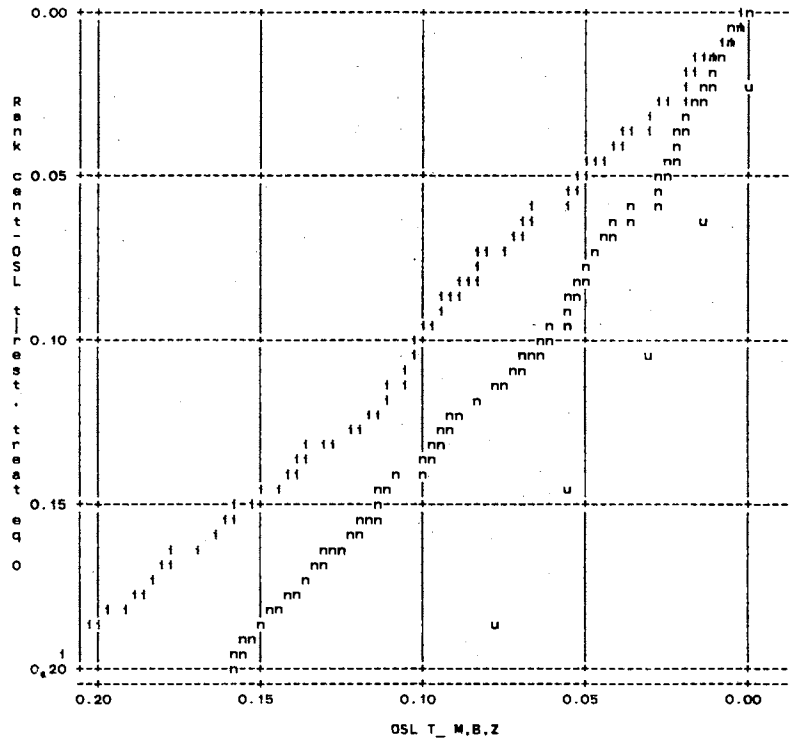
Graph 4.6.2

F Distribution Plot: τ=0
Upper Tail 20 Percent

The detailed plot, Graph 4.6.2, shows that at the

point where the F-distribution has 10 percent (.10 on the

lower, horizontal, axis) of the F-ratios more extreme

than the point, the actual percentage (on the lefthand,

vertical, axis) of randomizations with more extreme

F-ratios is 10 percent in the case of experimental unit-

treatment interaction error (i), about 14 percent in the

case of normally distributed random error (n) and about

20 percent in the case of experimental unit error (u).

The next pair of plots, Graphs 4.6.3 and 4.6.4, uses

the non-central F distribution to obtain OSL's for the

nonzero τ of this simulation. These plots the usual

degrees of freedom; that is, the variance-covariance

matrix of each error type is assumed to be nonsingular.

```
                        Compare Three Types of Error
              u = experimental unit error
              i = experimental unit-treatment interaction error
              n = normally distributed random error

              Plot of RSTN102*STN102.  Symbol used is 'u'.
              Plot of RSTN104*STN104.  Symbol used is 'i'.
              Plot of RSTN146*STN146.  Symbol used is 'n'.

       0.0 +-------------+-------------+-------------+-------------+---------nin---
           | |                                                           inu
  R        | |                                                        inina|
  a        | |                                                       inin  uu
  n        | |                                                      inn  uu
  k        | |                                                    iinn uu
     0.2 +-------------+-------------+-------------+--------------iinnuuuu
  n        | |                                                  iiinn uu
  o        | |                                                |innn uu
  n        | |                                               iinnn   u
  c        | |                                            iiiinnn   uu
  o        | |                                           innn   uu
  s        | |                                         ninn   |uuu
  L     0.4 +-------------+-------------+------------+nnni---uuu|
           | |                                      nnnii  uuuu
  t        | |                                     nnn|iiuuuu
  i        | |                                    nnn iiduuu
  r        | |                                   nn udduu
  e        | |                                  uoonuddi
  s        | |                                uuuonn 111
  t        | |                               uuuonnn 111
     . 0.6 +-------------+-------------+----uuunnn---11------
           | |                            uuunnn   111
  t        | |                           uuunnn  | 111
  r        | |                          uuonnn  |11
  e        | |                         uunn   111
  a        | |                        uon    11
  t        | |                       uoo   111
     n 0.8 +-------------+----oon----11-------
  e        | |              uoon|  1111
           | |             oon  | 111
  o        | |            uoo   11
           | |           oon   111
           | |          uoo  1111
           | |         uoon 111
           | |        oonii1
     1.0 +-oon111----------+-------------+-------------+---
           | |
           | |
         --+-------------+-------------+-------------+-------------+-------------+--
           1.0           0.8           0.6           0.4           0.2           0.0

                                    STN102
```

Graph 4.6.3

F Distribution Plot: $\tau \neq 0$

Entire Range

```
                    Compare Three Types of Error
          u = experimental unit error
          i = experimental unit-treatment interaction error
          n = normally distributed random error

             Plot of RSTN102*STN102.  Symbol used is 'u'.
             Plot of RSTN104*STN104.  Symbol used is 'i'.
             Plot of RSTN146*STN146.  Symbol used is 'n'.

     0.00 +-+-----------+-----------+-----------+-----------+--n-------+-+
                                                              nu
R                                                           nniu
a                                                           n   u
n                                                          ni  u
k                                                        ninh  u
                                                        iinn   u
n  0.05 +---------------------------------------iiinn---u-------+
o                                              |i i nn      u   |
n                                              iinn         u
c                                              nn         u
.                                          i ni           u
O                                         ii  n           uu
S                                         ii  nn          u
L                                       i   nn          uu
t                                    ii    n           u
i                                    ii    n           u
r  0.10 +------------------------ii-----n------u----------------+
e                                    ii   nnn          uu
s                                    ii    n           uu
t                                    i  .n n          uu
.                                    ii    nn         u
                                     ii    nn        u
t                                 ii    nn          uu
r                              iii  nnn           u
e                           ii'    n             u
a                         ii      nn            u
t                        i      nnn           uu
   0.15 +--------------ii-----nn--+-------u-------+
n                      ii    n            uu
e                       i     n           u
                          i    nn          uu
o                        ii   nnn         u
                        ii  n n          u
                      iii  . n          uu
                    iii    nn          uuu
                    i      n           u
                     ii    nn         uu
                     ii    nn        u
   0.20 +----------ii-----nn-------------u--------+
         +--+-----------+-----------+-----------+-----------+--+
         0.20        0.15        0.10        0.05        0.00

                            STN102
```

Graph 4.6.4

F Distribution Plot: T≠0

Upper Tail 20 Percent

Graphs 4.6.3 and 4.6.4 suggest that none of the types of errors yields an F-ratio which follows the non-central F distribution. In the detailed plot, Graph 4.6.4, all three lines are below the diagonal in the upper half of the plot. This indicates excessive probability in the upper tail. The excess is slight for experimental unit-treatment interaction error (i), but larger for normally distributed random error (n) and experimental unit error (u). Also there is a deficiency

of probability in the lower tail of the distribution.
This is indicated by the line's being below the diagonal
in the lower half of the plot.

The last two error plots, Graphs 4.6.5 and 4.6.6,
display the OSL's derived from the F-ratio testing the
hull hypothesis of $\beta=0$, when the value of $\beta=2.0$. They
use the non-central F distribution and the usual degrees
of freedom by assuming that the variance-covariance
matrix of each error type is nonsingular.

```
                    Compare Three Types of Error
            u = experimental unit error
            i = experimental unit-treatment interaction error
            n = normally distributed random error

            Plot of RSZN002*SZN002.  Symbol used is 'u'.
            Plot of RSZN004*SZN004.  Symbol used is 'i'.
            Plot of RSZN046*SZN046.  Symbol used is 'n'.

      0.0 +-------------+-----------+-----------+-----------+------------nn----
          |             |           |           |           |            nnu
          |             |           |           |           |           inn
      R   |             |           |           |           |          iinnu
      a   |             |           |           |           |         innn u
      n   |             |           |           |           |         inn
      k   |             |           |           |           |        iinn u
      0.2 +-------------+-----------+-----------+-----------+------iinn----u---+---
      n   |             |           |           |           |     iinn
      o   |             |           |           |           |    innn
      n   |             |           |           |           |   inno
      c   |             |           |           |           |  nnn
      -   |             |           |           |         u inn
      0   |             |           |           |        unnn
      S   |             |           |           |       nnni
      L   0.4 +---------+-----------+-----------+------unni----
          |             |           |           |     nnni
      z   |             |           |           |    unni
      i   |             |           |           |   unnn
      r   |             |           |           |   nnni
      e   |             |           |           |  onii
      s   |             |           |           | nnnii
      t   |             |           |           | nn uii
      0.6 +-------------+-----------+-----------+--on---ii----+---
      c   |             |           |         nnn  ii
      o   |             |           |        nno   ii
      v   |             |           |       nnn  iii
      a   |             |           |  u    nn   ii
      r   |             |           | u   nnn  |ii
          |             |           |   nn   |ii
      n   0.8 +---------+-----------+-u nnn----ii----+---
      e   |             |           | nnn   ii
      o   |             |           | non   ii
          |             |           | nnu   ii
          |             |          nnn   iiii
          |             |         nnn  oii
          |             |        nn | iii
          |             |       nnniiii  u
      1.0 +--n--nni-----+---------+--+u---
          | |           |           |           |           |           |
          +-+-----------+-----------+-----------+-----------+-----------+---
          1.0          0.8         0.6         0.4         0.2         0.0

                              SZN002
```

Graph 4.6.5
F Distribution Plot: $\beta\neq0$
Entire Range

Compare Three Types of Error
u = experimental unit error
i = experimental unit-treatment interaction error
n = normally distributed random error

Plot of RSZN002*SZN002. Symbol used is 'u'.
Plot of RSZN004*SZN004. Symbol used is 'i'.
Plot of RSZN046*SZN046. Symbol used is 'n'.

Graph 4.6.6
F Distribution Plot: $\beta \neq 0$
Upper Tail 20 Percent
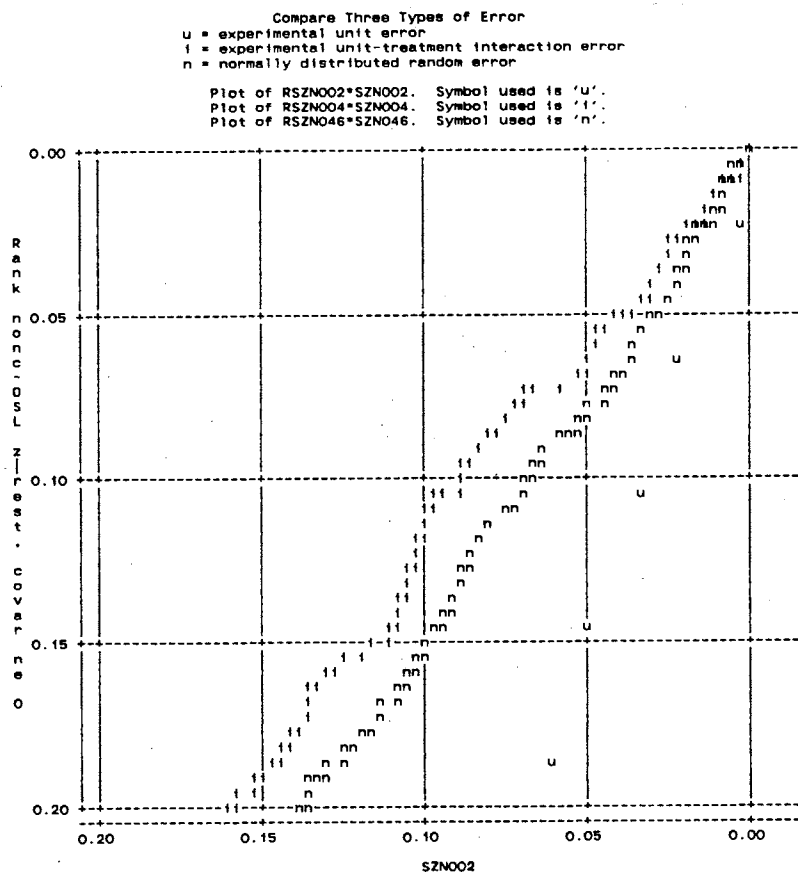
The points in Graphs 4.6.5 and 4.6.6 for

experimental unit error (u) show the clustering

discussed in Theorem 3.7.9(3), (p. 202). The conclusions

are the same as those for Graphs 4.6.3 and 4.6.4. The

tails are heavy in the case of experimental unit error,

and slightly so in the other two cases.

In summary, the three pairs of graphs cast doubt on

the assumption of an F-distribution, central or non-

central, with the usual degrees of freedom in the case of

experimental unit errors (u). The other two error types

yield close approximations to the F-distribution.

4.6.3 Equal and Unequal Within-Block Error Variances

The next seven pairs of F-probability plots compare (a) the OSL's from cases with the same (s) error variances in both blocks to (b) the OSL's from cases with different (d) error variances in the two blocks. The ratio of error variances is 1 to 4, block one vs block two. The two unit errors each have three pairs of plots, one pair each for the cases of $\tau = 0$, $\tau \neq 0$ and $\beta = 0$. The normally distributed random error has one pair of graphs ($\tau = 0$), as its other two ($\tau \neq 0$ and $\beta \neq 0$) resemble the one for $\tau = 0$. All graphs assume the usual degrees of freedom.

When the errors are experimental unit errors and the treatments are zero, as in Graphs 4.6.7 and 4.6.8, the F-ratios and accompanying OSL's display the clusters discussed in Theorem 3.7.9(5), (p. 203). The off-diagonality of the u's of Graph 4.6.1, (p. 295), is more clearly seen here. Graph 4.6.8 shows that the equal within-block variance case (s) has a heavy upper tail, while the unequal case (d) has a light upper tail, both for $\alpha \leq .075$. For $\alpha \leq .075$ the actual Type I error level exceeds α in the equal within-block variance case (s), but is less than α in the unequal within-block variance case (d). For $\alpha > .075$ the actual Type I error exceeds α for both cases.

Effect of Unequal Within-block Variances
s = same block variance   d = different block variances

Plot of RSTC002*STC002.   Symbol used is 's'.
Plot of RSTC003*STC003.   Symbol used is 'd'.
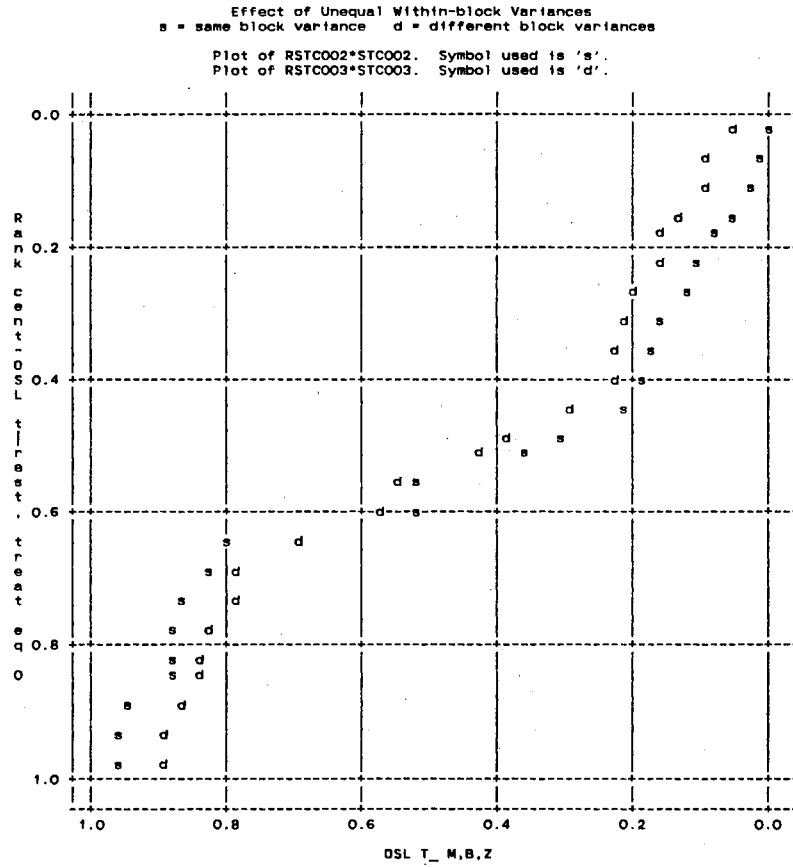
Graph 4.6.7
Equal vs Unequal Within-Block Error Variances $\tau = 0$
Experimental Unit Error -- Entire Range

Effect of Unequal Within-block Variances
s = same block variance   d = different block variances

Plot of RSTC002*STC002.   Symbol used is 's'.
Plot of RSTC003*STC003.   Symbol used is 'd'.

```
            | |                |            |            |            |
   0.00 +---+-----------------+------------+------------+------------+
            | |                |            |            |            |
            | |                |            |        d|              s
R           | |                |            |            |            |
a           | |                |            |            |            |
n           | |                |            |            |            |
k           | |                |            |            |            |
c  0.05 +---+-----------------+------------+------------+------------+
e           | |                |            |            |            |
n           | |                |            |            |            |
t           | |                |        d  |            |        s   |
-           | |                |            |            |            |
O           | |                |            |            |            |
S           | |                |            |            |            |
L           | |                |            |            |            |
t  0.10 +---+-----------------+------------+------------+------------+
i           | |                |        d  |            s            |
r           | |                |            |            |            |
e           | |                |            |            |            |
s           | |                |            |            |            |
t           | |                |            |            |            |
.           | |                |            |            |            |
t           | |                |            |            |            |
r           | |                |   d        |        s   |            |
e  0.15 +---+-----------------+------------+------------+------------+
a           | |                |            |            |            |
t           | |                |            |            |            |
e           | |                |            |            |            |
q           | |                |            |            |            |
            | |                |            |            |            |
o           | |                |            |            |            |
            | |   d            |        s   |            |            |
   0.20 +---+-----------------+------------+------------+------------+
            | |                |            |            |            |
          0.20            0.15         0.10         0.05         0.00
```

OSL T_ M.8.Z
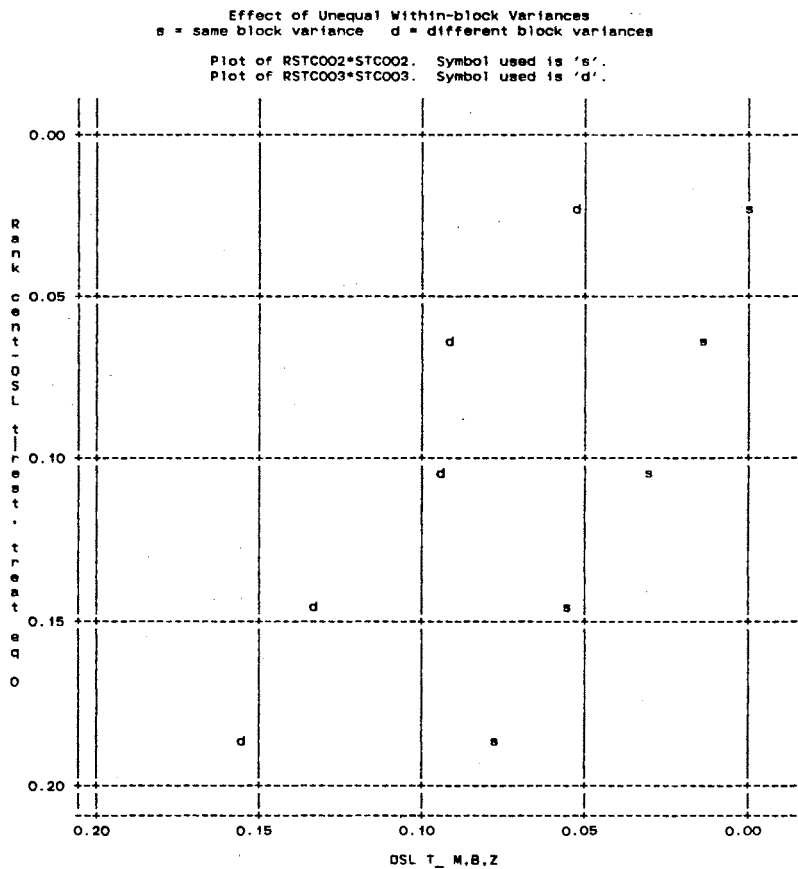
Graph 4.6.8

Equal vs Unequal Within-Block Error Variances τ=0
Experimental Unit Error - Upper Tail 20 Percent

For experimental unit error and nonzero treatments, as in Graphs 4.6.9 and 4.6.10, the treatment effects disperse the clustering seen in Graphs 4.6.7 and 4.6.8. For equal within-block error variances (s), the distribution is tail-heavy. Where the non-central F-distribution would have covered .2 of the F-ratios (.2 on the lower, horizontal, axis), the randomization distribution has covered some .35 (on the lefthand, vertical, axis) of the F-ratios. There are more larger F-ratios than one would expect from the usual non-central F-distribution. For

unequal within-block error variances (d), the distribution is tail heavy for OSL's > .10 and tail light for OSL's < .05.

The asymmetry about the center of Graph 4.6.9, (.5,.5), is due to the nonzero value of the non-centrality parameter's having a larger effect in the upper tail of the F-distribution than in the lower tail.

```
                    Effect of Unequal Within-block Variances
                  s = same block variance    d = different block variances

                      Plot of RSTN102*STN102.  Symbol used is 's'.
                      Plot of RSTN103*STN103.  Symbol used is 'd'.

        0.0 +--------------------------------------------------dd--s---
              | |                                               dd   s
              | |                                               d   ss
              | |                                              dd  ss
              | |                                             ddd  ss
              | |                                             dd   ss
          R   | |                                            dd   ss
          a   | |                                            dd   ss
          n 0.2 +---------------------------------------------d--ss-------+--
          k     | |                                          |ddd  ss
              | |                                           ddd   s
          n   | |                                          dd|   ss
          o   | |                                          ddd    ss
          n   | |                                         dd     ss
          c   | |                                        ddd    |ss
          -   | |                                       dd    sss|
          O   | |                                      ddd    sss|
          S 0.4 +--------------------------------ddd------s---+------------+--
          L   | |                                |dd    ssss
              | |                              ddd    ssss
          t   | |                             ddd  | sss
          i   | |                            dd   sss
          r   | |                           dddsssssss
          e   | |                           dddsss
          s   | |                          |sddds
          t   | |                          sddd
          . 0.6 +-------------------ssdd-----------+------------+------------+--
              | |                 ssddd
          t   | |                ssdd|
          r   | |               sssdd
          e   | |              sssddd
          a   | |              ssddd
          t   | |             ssdd
              | |             ssdd
          n   | |            sssdd
          e 0.8 +----------ssddd---------+------------+------------+------------+--
          o   | |          ss dd
              | |          sssdd|
              | |          ss dd
              | |         ssddd
              | |         sddd
              | |        sddd
              | |        sdd
              | |       sdd
              | |       ddd
        1.0 +-ddd------------------------------------------------------
              | |
              +-----------------------------------------------------------
               1.0        0.8        0.6        0.4        0.2        0.0

                                      STN102
```

Graph 4.6.9

Equal vs Unequal Within-Block Error Variances τ≠0
Experimental Unit Error - Entire Range

Effect of Unequal Within-block Variances
s = same block variance   d = different block variances

Plot of RSTN102*STN102. Symbol used is 's'.
Plot of RSTN103*STN103. Symbol used is 'd'.
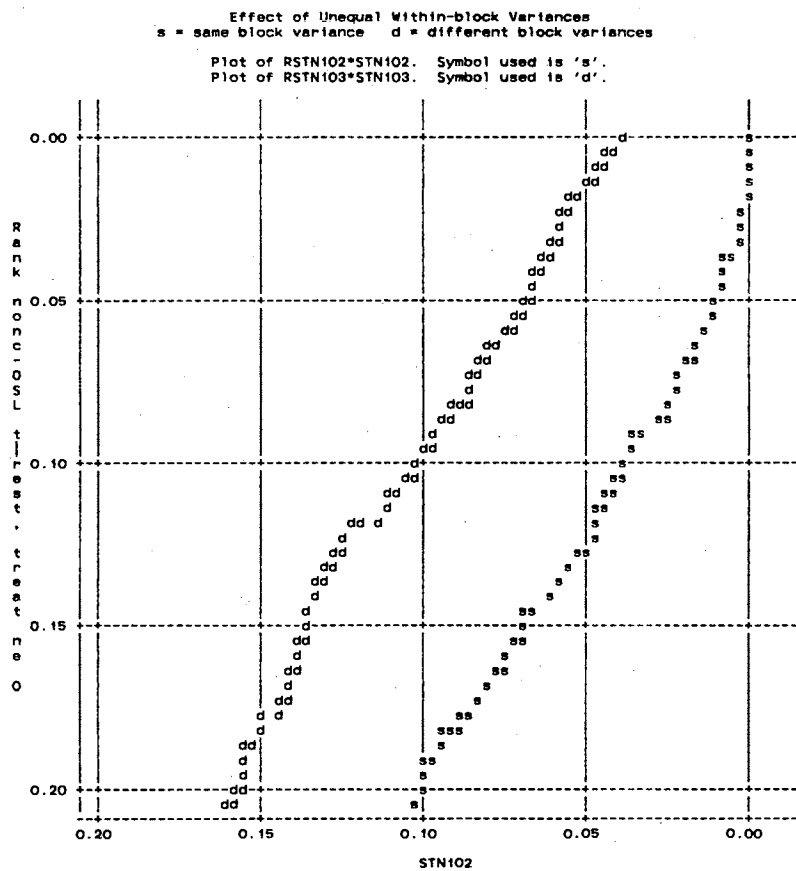
Graph 4.6.10

Equal vs Unequal Within-Block Error Variances τ≠0
Experimental Unit Error - Upper Tail 20 Percent

The F-ratios, and hence the OSL's, for the cases

with β=2.0 display the clustering proved in Theorem

3.7.9(3), (p. 202). Graphs 4.6.11 and 4.6.12 display

these OSL's for the case of experimental unit error.

While the points are near the diagonal for OSL's less

than .7, the lower tail is light for both equal and

unequal within-block error variances. In this region,

1.0 to .7, the F-ratios do not follow this noncentral

F-distribution; there are fewer very small F-ratios than

one would expect. When the errors have the same

within-block error variances (s), Graph 4.6.12 shows that
the randomization distribution of the F-ratios is tail
heavy for OSL's ≤.2.   For unequal within-block error
variances (d), the near diagonal pattern for small OSL's
suggests a good approximation to this noncentral
F-distribution within this region.

```
                    Effect of Unequal Within-block Variances
              s = same block variance    d = different block variances

                  Plot of RSZN002*SZN002.  Symbol used is 's'.
                  Plot of RSZN003*SZN003.  Symbol used is 'd'.

          | |        |         |         |         |          |
   0.0  +-+---------+---------+---------+---------+----------+---+---
        | |                                                   |  ds
        | |                                                 d   s
        | |                                                d     s
   R    | |                                            d        s
   a    | |                                          d         s
   n 0.2 +-+---------+---------+---------+---------+ d  s ----+---+---
   k    | |                                        d  s
        | |                                      d
   n    | |                                    ds
   o    | |                                   d
   n    | |
   c    | |
   - 0.4 +-+---------+---------+-------d-+-------s-+----------+---+---
   O    | |                          d         s
   S    | |                        d   d  s
   L    | |                        d    s
        | |                      d      s
   z    | |
   i    | |
   r 0.6 +-+---------+---------d-+---s-----+---------+----------+---+---
   e    | |                  d   s
   s    | |                s   d
   t    | |              s    d
   .    | |              s    d
   c 0.8 +-+---------+--s---d--+---------+---------+----------+---+---
   o    | |           s    d
   v    | |           sd
   a    | |           sd
   r    | |          d s
   n    | |          ds
   1.0  +-+---------+---------+---------+---------+----------+---+---
        | |        |         |         |         |          |
        +-+--------+---------+---------+---------+----------+---+
         1.0      0.8       0.6       0.4       0.2        0.0

                                SZN002
```
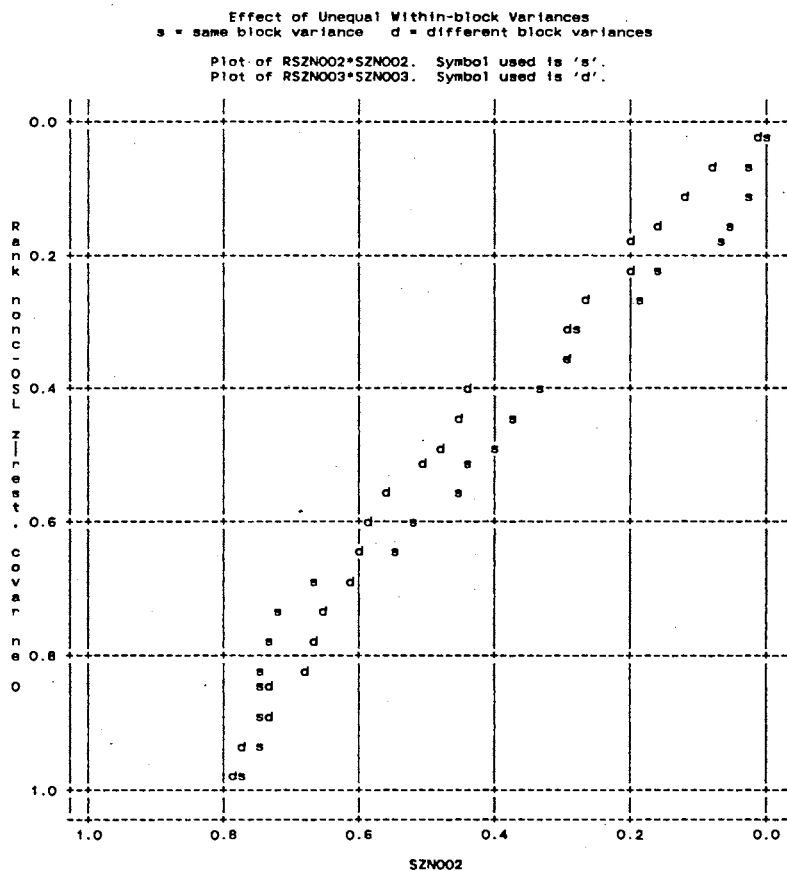
Graph 4.6.11
Equal vs Unequal Within-Block Error Variances ρ≠0
Experimental Unit Error - Entire Range

Effect of Unequal Within-block Variances
s = same block variance   d = different block variances

Plot of RSZN002*SZN002.   Symbol used is 's'.
Plot of RSZN003*SZN003.   Symbol used is 'd'.

```
R
a
n
k

n 0.05 +----------------------------------------------------------------
o                                                    d              s
n
c
-
O
S
L

z
i
r 0.10 +----------------------------------------------------------------
e                          d                      s
s
t
.

c
o
v
a
r 0.15 +----------------------------------------------------------------
n          d                                      s
e

o

              d                              s

0.20 +
```

        0.20   0.15   0.10   0.05   0.00

                            SZN002
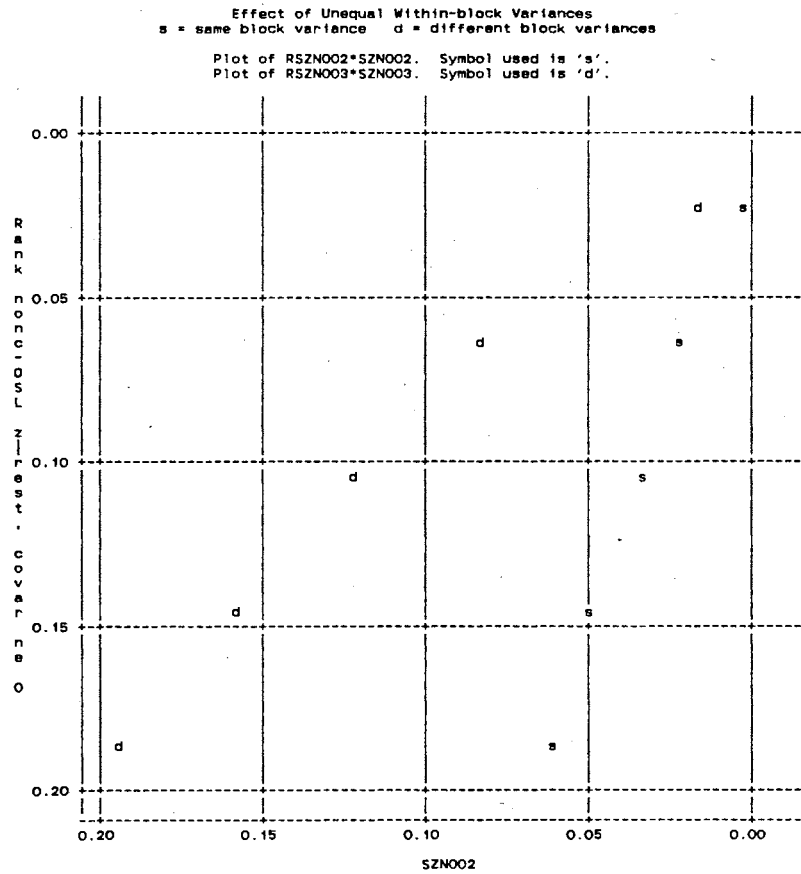
            Graph 4.6.12

Equal vs Unequal Within-Block Error Variances $\rho \neq 0$
Experimental Unit Error - Upper Tail 20 Percent

    The next six graphs, Graphs 4.6.13 through 4.6.18,

display the case of experimental unit-treatment interac-

tion errors.   The equality or nonequality of within-block

variances has little affect on the OSL's for this type of

error.   Points for all three pairs are on the diagonal in

the important upper tail area.   The graphs displaying the

entire range show all points near the diagonal, suggest-

ing the F-distribution is a good approximation to the

randomization distribution for this error type.   The

detailed graph for the case of $\tau = 0$, Graph 4.6.14, shows

the equal and unequal cases differing slightly for OSL's
between .2 and .05. The actual Type I error rate is
lower than α in this region for the case of unequal
within-block error variances (d) and higher than α for
the case of equal within-block error variances (s).
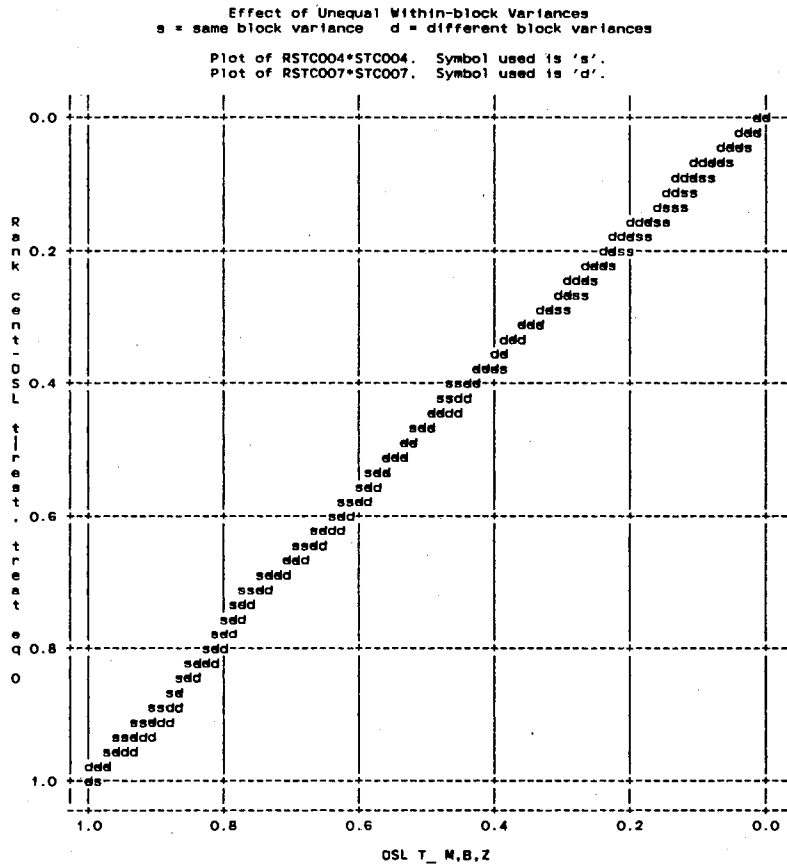
Graph 4.6.13
Equal vs Unequal Within-Block Error Variances $\tau=0$
Experimental Unit-Treatment Interaction Error
Entire Range

Effect of Unequal Within-block Variances
s = same block variance   d = different block variances

Plot of RSTC004*STC004. Symbol used is 's'.
Plot of RSTC007*STC007. Symbol used is 'd'.
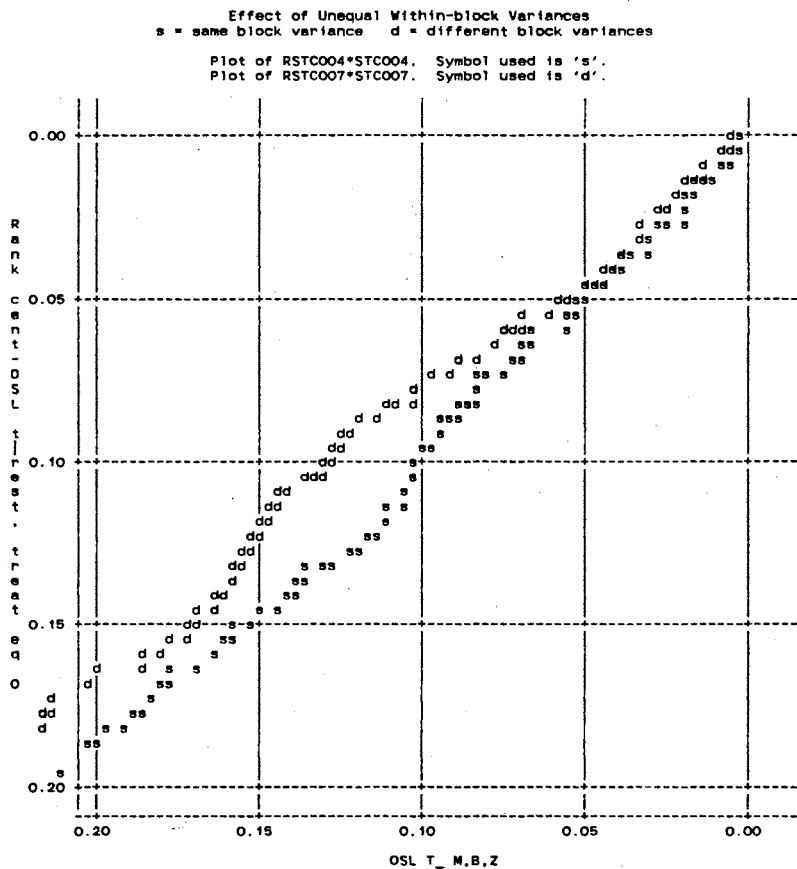
Graph 4.6.14
Equal vs Unequal Within-Block Error Variances τ=0
Experimental Unit-Treatment Interaction Error
Upper Tail 20 Percent

Graphs 4.6.15 and 4.6.16 are for the experimental unit-treatment interaction error cases with $\tau \neq 0$. Deviation from the diagonal throughout the range 1.0-.85 suggests this non-central F distribution differs from the randomization distribution. The detailed graph, Graph 4.6.16, suggests a good approximation in the upper tail, say, from .15 to zero.
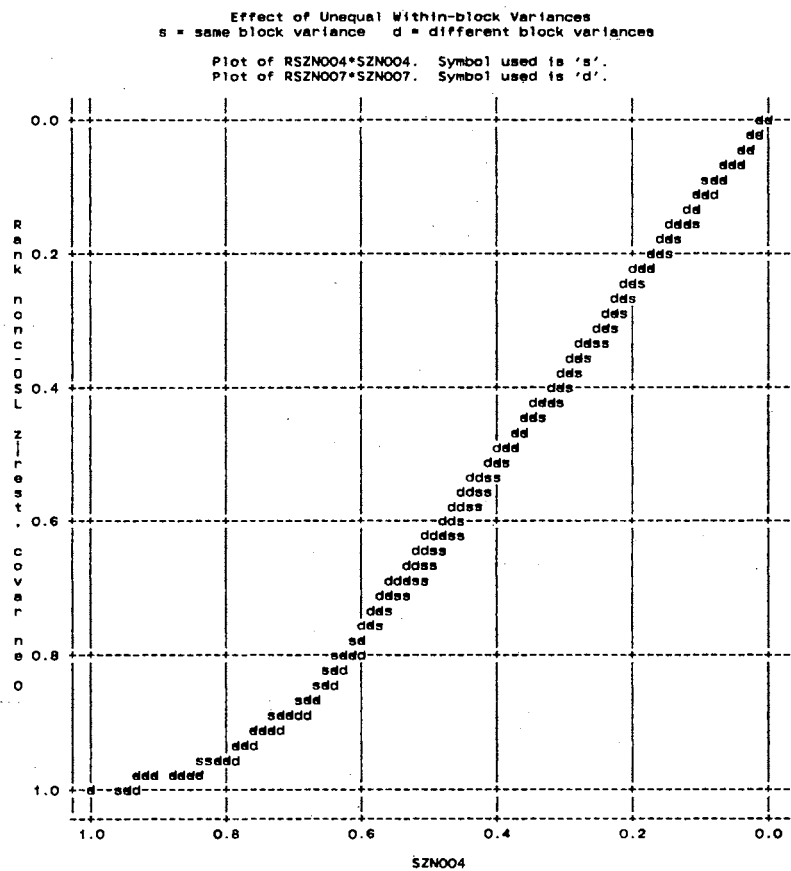
Graph 4.6.15

Equal vs Unequal Within-Block Error Variances $\tau \neq 0$
Experimental Unit-Treatment Interaction Error
Entire Range

```
                    Effect of Unequal Within-block Variances
                 s = same block variance   d = different block variances

                    Plot of RSTN104*STN104.   Symbol used is 's'.
                    Plot of RSTN107*STN107.   Symbol used is 'd'.

       0.00 +-+-----------+------------+------------+------------ds------+
                                                                      dd |
            |                                                         ds
    R                                                                ddd
    a                                                                ds
    n                                                              d  ss
    k                                                            ddds  s
                                                                  dss
       n 0.05 +-+------------+------------+------------+----------ddds-------+
       o                                                          ddd s
       n                                                        -ddss-------
       c                                                        |d  s
       -                                                         sd
       O                                                       dd  s
       S                                                     dd s s
       L                                                      dss
         t                                                    dss
         |                                                   dds
       r                                                    ssdd
       e                                                    ssd
       s                                                    ssd
       t 0.10 +-+------------+------------+----------ssd----------+------------+
       .                                            |sdd
                                                    dds
       t                                            dds
       r                                           ddss
       e                                          d  s
       a                                        ddd ss
       t                                       d dss|
       n                                        ddss
       e                                        dds
       o                                         ds
       0.15 +-+-----------+----------d--dss----+------------+------------+
                                     d  ss
                                    d    s
                                   ddd  s
                                   dd   ss
                                 d  d|  ss
                                 dd   sss
                                 d sss
                                dds
                                ddss |
                              d dss
       0.20 +-+----dd--ss----+------------+------------+------------+----
            | |    d  sss     |            |            |            |
            +-+-----------------------------------------------------------
             0.20         0.15         0.10         0.05         0.00

                                      STN104
```
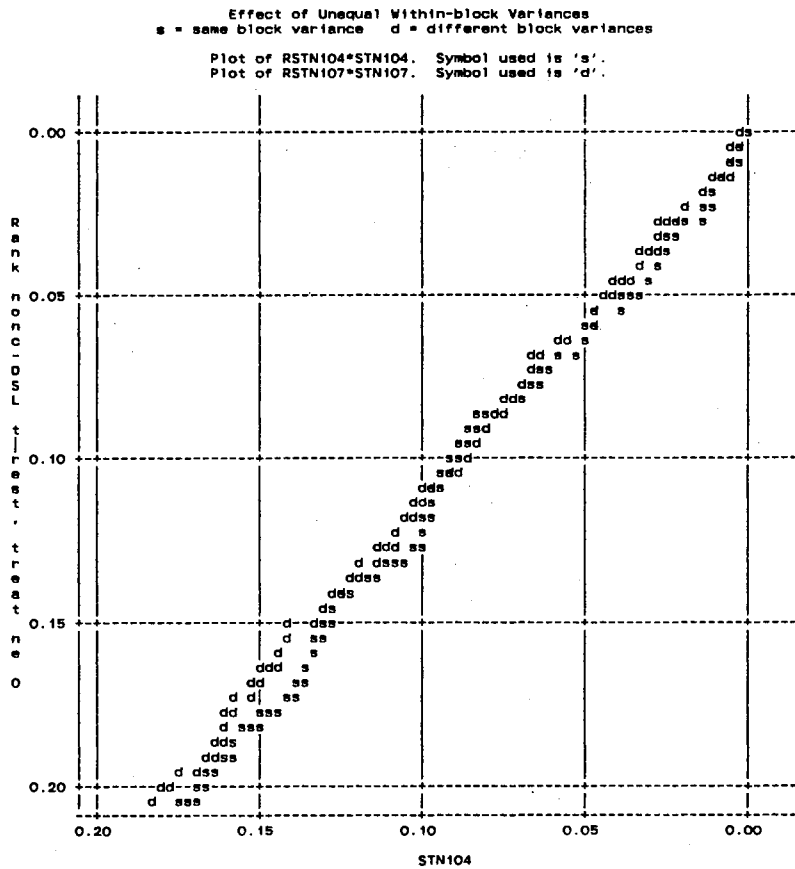
Graph 4.6.16
Equal vs Unequal Within-Block Error Variances $\tau \neq 0$
Experimental Unit-Treatment Interaction Error
Upper Tail 20 Percent

For $\beta \neq 0$, Graph 4.6.17, shows a deviation from the diagonal suggesting that the randomization distribution and the F-distribution differ for this case. The detailed graph, Graph 4.6.18, suggests a good approximation in the upper tail, say from .075 to zero. Both graphs are for the experimental unit-treatment interaction error case.

Effect of Unequal Within-block Variances
s = same block variance   d = different block variances

Plot of RSZN004*SZN004.   Symbol used is 's'.
Plot of RSZN007*SZN007.   Symbol used is 'd'.
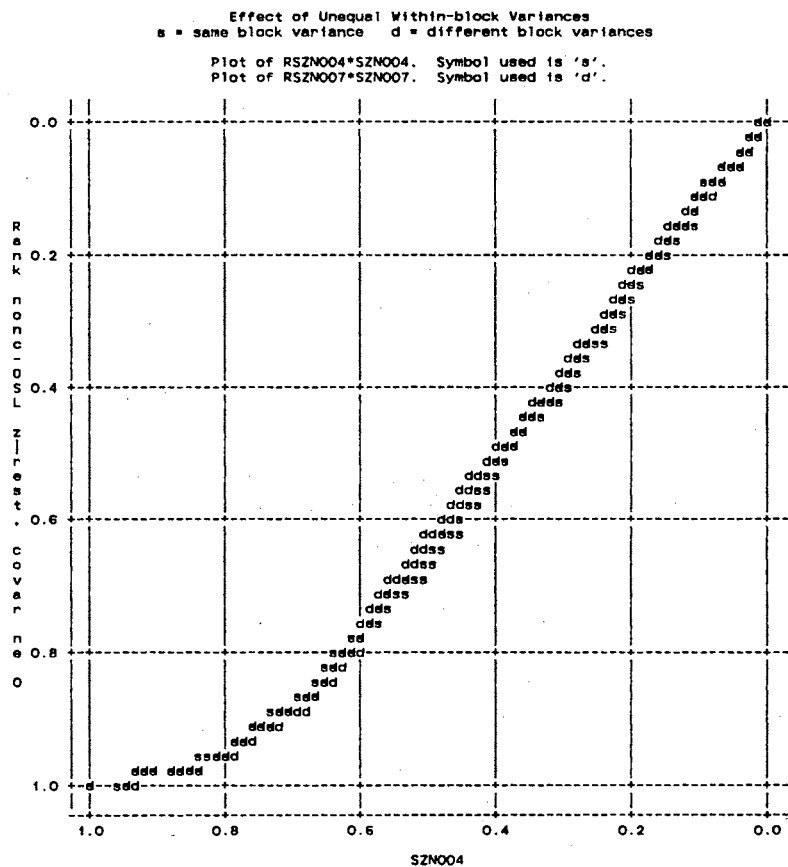
Graph 4.6.17
Equal vs Unequal Within-Block Error Variances $\beta \neq 0$
Experimental Unit-Treatment Interaction Error
Entire Range
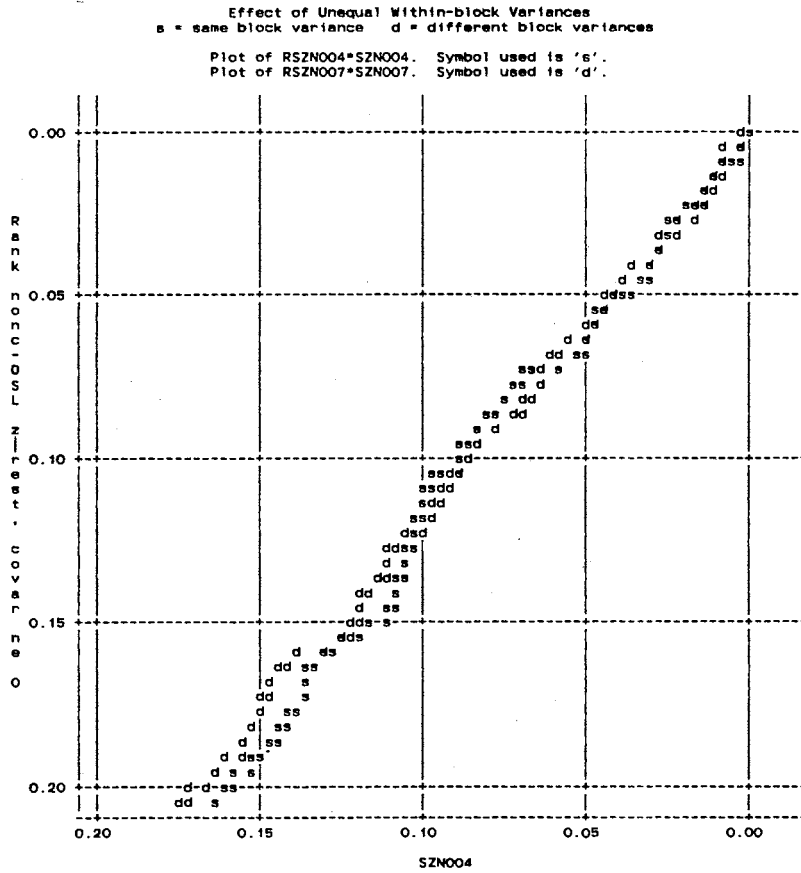
Graph 4.6.18

Equal vs Unequal Within-Block Error Variances ρ≠0
Experimental Unit-Treatment Interaction Error
Upper Tail 20 Percent

For the case of normally distributed random errors, the three pairs of graphs for $\tau=0$, $\tau\neq0$ and $\beta\neq0$, are remarkably alike. Only the pair for $\tau=0$ is shown. These are Graphs 4.6.19 and 4.6.20. Equality of within-block error variances (s) leads to a distribution slightly tail-heavy in both tails. At the $\alpha=.1$ level (.10 on the lower, horizontal axis) the probability of a Type I error is about .14 when the error variances are equal in both blocks (s) and is about .075 when the error variances differ (d) in the two blocks (.14 and .075 on the left-hand, vertical, axis).

```
                    Effect of Unequal Within-block Variances
               s = same block variance    d = different block variances

                    Plot of RSTCO46*STCO46.  Symbol used is 's'.
                    Plot of RSTCO55*STCO55.  Symbol used is 'd'.

          |  |         |         |         |         |        |
     0.0  +--+---------+---------+---------+---------+--------dd---
          |  |                                              ddd|
          |  |                                             ddds
          |  |                                           ddddsss
          |  |                                         dddd  ss
   R      |  |                                         ddd   sss
   a      |  |                                        |dd    ss
   n 0.2  +--+---------+---------+---------+---------ddd+--ss------+---
   k      |  |                                       dd |sss
          |  |                                      ddd  ss
   c      |  |                                      dd  sss
   e      |  |                                     ddd   ss
   n      |  |                                     ddd  ss
   t      |  |                                     dd  sss
   -      |  |                                    dd  ss
   0      |  |                                   dd sss
   S 0.4  +--+---------+---------+---------+----dd-ss----------+---
   L      |  |                                 ddd|ss
   t      |  |                                dddsss
   i      |  |                               ddsss
   r      |  |                              ddsss
   e      |  |                              ddds
   s      |  |                             ddd
   t      |  |                            ddds
   . 0.6  +--+---------+---------+-------dds--------+---------+---
   t      |  |                          dds
   r      |  |                         ddd
   e      |  |                        sdd
   a      |  |                       dddd
   t      |  |                      ddd
   e      |  |                     sdd
   q      |  |                    ssdd
     0.8  +--+--------ssddd-------+---------+---------+--------+---
   o      |  |        ssddd|
          |  |        sddd
          |  |        sdd
          |  |       ssdd
          |  |      sddd
          |  |     ssdd
          |  |    ssdd
     1.0  +--+--dd-+---------+---------+---------+---------+--------+---
          |  |
          |  |
             1.0       0.8       0.6       0.4       0.2      0.0

                    OSL T_ M,B,Z
```

Graph 4.6.19
Equal vs Unequal Within-Block Error Variances $\tau=0$
Normally Distributed Random Error – Entire Range
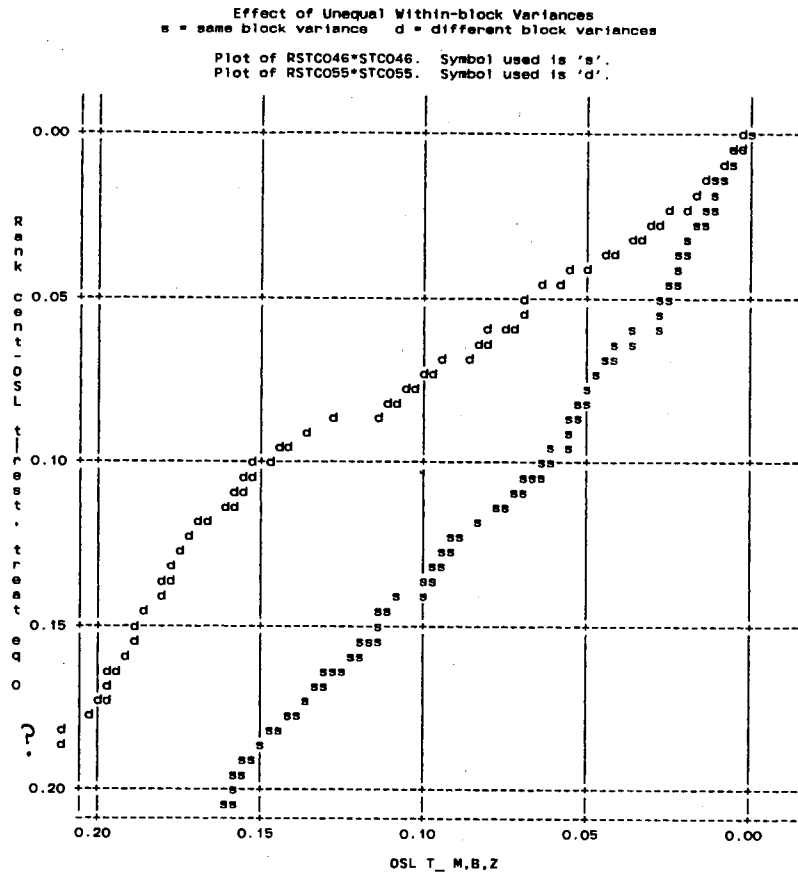
Graph 4.6.20

Equal vs Unequal Within-Block Error Variances $\tau=0$

Normally Distributed Random Error

Upper Tail 20 Percent

4.6.4 A Possible Problem with the Graphs

One possible problem with these graphs is that the
line of plotted points for those sets of randomizations
using normally distributed random errors does not fall
on the 45 degree line, as it should.

To make the normally distributed random errors more like the experimental unit errors, the normally distributed random errors are forced to sum to zero within each randomization. Tests indicate that forcing normally distributed random errors to have a mean of (to sum to) zero within each randomization costs one degree of freedom. That is, the plot of the resulting F-ratios is brought to the 45 degree line by reducing the appropriate (denominator) degree of freedom by one. Computer program testf of Appendix F.8 details this test.

For those randomizations using normally distributed random error, such a reduction did not bring the plotted line of F-ratios to the 45 degree line. A fractional reduction, from 2 to 1.75 helped, but no degree of freedom, constant for all randomizations, was able to bring the entire line of plotted values to the 45 degree line.

One possible reason is that standardizing to zero within each randomization gives the errors the singular normal distribution. See the discussion following (2.3.3), (p. 17). Following Corollary 2s.2 of Searle [1971, p. 69], we computed the Chi-square degrees of freedom for the denominator quadratic form, $x'Ax$ in Searle's notation. The 8 by 1 vector of normally distributed random errors provided the x; the denominator term within braces of (3.7.21), (p. 205), times σ^2 provided the matrix A. The errors and center

matrix A both change with each randomization. The

variance-covariance matrix, $V=LL'(I\sigma^2)$, in Searle's

notation, for the singular normal is, in this case, block

diagonal with identical diagonal blocks. The diagonal

blocks are $\{I_{4\times4} - (1/4)J_{4\times4}\}\sigma^2$; the off-diagonal blocks

are $O_{4\times4}$. The standardization transformation subtracts

the block-wise mean from each original error within each

block, hence $L=I-(1/t)J$. The variance, $I\sigma^2$, cancels in

the product AV.

The resulting degrees of freedom, trace(AV), differ-

ed for each randomization, ranging from 1.0 to 1.96 with

a mean of 1.4.

Unfortunately, the condition on Searle's Corollary

2s.2, that

$$VAVAV - VAV = O, \qquad (4.6.1)$$

appears not to hold for all randomizations. The maximum

element of the right-hand side of (4.6.1), over all ran-

domizations, is not zero, but 1.14, with the mean of the

maximum elements equalling .39. Furthermore, applying

these degrees of freedom as the divisor of the residual

sum of squares and as the denominator degrees of freedom

in the calculation of the OSL's brought the line of

plotted points farther from, not closer to, the diagonal.

Unable to justify altering the denominator degrees

of freedom, we use the one provided by the analysis of

variance table, 2, for all randomizations using normally

distributed random error. Thus, the line of plotted
points remains off of the diagonal line. As the F-ratio
for each randomization should be an F statistic and the
ensemble on the diagonal, further investigation is
suggested.

One could argue that the experimental unit errors
have the singular normal distribution. This is because
our method of construction forces them to be the
residuals from a least squares procedure. When the
errors of the original least squares model are normally
distributed, the residuals are known to have the singular
normal distribution. In this case $\Psi = \mathbb{L}\mathbb{L}' = \mathbb{L}$ of the
discussion following (2.3.3), (p. 22) and $\mathbb{A}$ is as above.
The condition on Corollary 2s.2 (4.6.1) is more nearly
satisfied. The maximum element of the right-hand side of
(4.6.1), over all randomizations, is .09 and the mean of
the maximum elements is .02. However, use of trace($\mathbb{A}\Psi$)
for the degrees of freedom moved the line of plotted
points farther from the diagonal.

Uncertainty of this approach led us to again use the
original, constant, degrees of freedom (2) for those
randomizations using experimental unit error.

Zyskind, Kempthorne, et $al.$, [1968, p. 11], point
out that forcing the experimental unit errors to sum to
zero also forces the covariance matrix of the Y_{ijk}'s (and
the errors) to be singular.

The experimental unit-treatment interaction errors for each treatment are also residuals from a least squares procedure. However, as each treatment appears but once per block, the entirety of any one set of residuals never appears in any one randomization. We see no argument that they have a singular normal distribution.

The SAS program snormal in Appendix F.9 details the tests of these singular normal distributions.

We now turn from equal selection probability for all randomizations to examine one proposed method which has all selection probabilities greater than zero, but unequal.

Section 4.7

Weighted Selection Probabilities

Section 4.7.1 reviews Cox's suggested weights for
the selection probabilities, $p(s)$, for the analysis of
covariance. Section 4.7.2 presents means and standard
deviations obtained via his weights for this simulation.
They should be compared with those in Tables 4.4.1-4.4.4,
(p. 259-268).   Section 4.7.3 displays graphs suggesting
that the ratio may be used to avoid randomizations with
large errors in estimation.

The major findings are that, for this simulation,
weighting the selection probabilities with Cox's third
ratio modestly improves the estimates of the model
equation parameters and/or their standard deviations.
Weighting has little effect on the expectation and
standard deviation of the mean squares and F-ratios. A
second finding is that using the weights to assign zero
selection probabilities to some randomizations excludes
from consideration those randomizations with highly
biased estimates for the adjusted τ and β.

4.7.1 Review of Cox's Suggested Weights

Cox [1956, 1957, and 1982] presents several ratios
useful in the analysis of covariance. Our understanding
of three of these ratios follows.

The first,

$$[SST_{zz} - SSM_{zz}]/ [(SSR_{zz}-SSM_{zz})/(t-1)], \qquad (4.7.1)$$

is discussed by Finney [1946, p. 54, Eq. 3], Lucas
[1950], Greenberg [1953, p. 698], Cox [1956, p. 1146, Eq.
7 (part)], Cox [1957, p. 153, Eq. 6b (part)], and Cox
[1982, p. 198, Eq. 3 (part)]. See (2.5.1), p. 35, for
the definition of SSM_{zz}, SST_{zz} and SSR_{zz}. The term SST_{zz}
first sums, over all blocks, those covariate values whose
experimental unit receives treatment 1, then sums those
receiving treatment 2, and so on for all k treatments.
If the randomization is such that all these sums (over
treatments) are equal, then the numerator of (4.7.1) is
zero; otherwise it is positive. Equation (4.7.1) follows
a negative sign in the expectation of the numerator of
the F-ratio for the treatments [Greenberg, 1953, p. 698,
Eq. (A)]. This equation follows under the assumption
that the true mean of the covariate is identical in each
of the treatment groups; that is, the design is balanced
on the covariate. Under this condition, minimizing
(4.7.1) maximizes the numerator of this F-ratio when the
treatment effects are nonzero [Greenberg, 1953, p. 698].
Equation 3 of Cox [1982, p. 198] extends (4.7.1) to the
multiple covariate case with the trace function reducing
the dimensionality to one.

The second ratio is

$$R(\mu,b,\tau) \; / \; E_R\{SSR_{ZZ}/(\text{d.f. of } SSR_{ZZ})\}. \qquad (4.7.2)$$

The $R(\cdot)$ notation is detailed in Appendix B, p. 385.
This reduction in sums of squares, $R(\mu,b,\tau)$, is for Z and
has the form $Z'AZ$, where A is the projector matrix for
$[\Sigma_\mu \mid \Sigma_b \mid \Sigma_\tau]$. The residual sum of squares, SSR_{ZZ}, is
the matrix version as in Table 2.5.2, (p. 38), for the
model equation $Z=\mu+b_i+\tau_k+\text{error}$. The expectation is over
all randomizations. This ratio appears in Cox [1982, p.
200, the final paragraph]. Cox denotes the denominator
as Ω_z, which he defines on the bottom of his page 198.

The third ratio is

$$SSR_{ZZ} \; / \; E_R\{SSR_{ZZ}/(\text{d.f. of } SSR_{ZZ})\}. \qquad (4.7.3)$$

The denominator is as in (4.7.2); the numerator, SSR_{ZZ},
differs from the numerator of (4.7.2) by terms constant
over all randomizations. The difference is such that
ratios two and three are linearly and inversely related.
The third ratio is based upon Cox [1956, p. 1148, Eq.
14]. Cox [1956, p. 1150, iii] states,

> Arrangements with a large value of SSR_{ZZ}
>
> will have a small value for SST_{ZZ} [better
>
> balanced on the covariate], and conversely.
>
> Hence the weighting proportional to SSR_{ZZ}
>
> attaches greater chance of selection to those
>
> arrangements in which the treatment groups are
>
> balanced with respect to the mean value of Z.

Ratio one (4.7.1) did not yield the informative
graphs of ratio two (equivalently ratio three) and has
not been examined as closely as the latter two ratios. A
graph of ratio one vs ratio three resembles the probabil-
ity density function for the exponential distribution.
The SAS program glm of Appendix F.6 details the
computation of these three ratios.

4.7.2 Cox's Third Ratio Applied to Simulations

A step in the SAS program glm uses the third ratio
as a weighting variable to compute the values of
Tables 4.7.1 - 4.7.4. These tables should be compared
with Tables 4.4.1 - 4.4.4, (p. 259 - 268), which display
the same means and/or standard deviations for parameter
estimates and analysis of variance terms, but without the
weights of ratio three.

For the case of experimental unit error, Table 4.7.1
below, weighting with the third ratio improves the expec-
tation of the mean, the block effects, and the covariate
coefficient. All are closer to their true values than in
the unweighted case. The expectations of the treatment
estimators remain equal to their true values because the
weights are equal within each g-group and, within each
g-group, the treatment estimators are unbiased. Table
4.7.2 below shows that the standard deviations for all
parameter estimates are reduced. Weighting affects these
terms in a like manner in the unequal within-block
error variance case; the corresponding table is not shown.

The effect upon the expectation of the mean squares and F-ratios, Table 4.7.3, is modest and mixed; changes are desirable in some cases and undesirable in others. The standard deviations are slightly smaller than in the unweighted case. This also holds for the unequal within-block error variance cases for the experimental unit errors, Table 4.7.4.

For the case of experimental unit-treatment interaction error, weighting with the third ratio again improves all estimates of model equation parameters and reduces the standard deviation of the estimates, as seen in Tables 4.7.1 and 4.7.2. The estimates for the blocks and covariate parameters are unbiased or almost so.

The effect of weighting upon the expectation and standard deviation of the residual and adjusted treatment mean squares and their F-ratios is negligible, for both equal and unequal within-block error variance cases, as in Tables 4.7.3 and 4.7.4. The weighting increases the expectation of the mean square and F-ratio for the adjusted covariate, but has little effect upon their standard deviations. This increase occurs in both equal and unequal within-block error variance cases for the experimental unit-treatment interaction errors.

For the case of normally distributed random error, the weighting has little effect upon the expectations of the estimated equation parameters, but does reduce the standard deviation of these estimates, Tables 4.7.1-4.7.2.

Weighting has little, if any, effect upon the expectation and standard deviation of the mean squares for the residual and adjusted treatment effects, as seen in Tables 4.7.3 and 4.7.4. Weighting increases the expectation of the mean square for the adjusted covariate, but leaves its standard deviation nearly unchanged. The expectation and standard deviation of all F-ratios are modestly increased in the equal within-block error variance case. The effect is mixed in the unequal within-block error case for the normally distributed random errors.

The expected mean square residual differs from the expected mean square for $\tau = 0 | \mu, b, \beta$ about as much in the weighted as in the unweighted case. See Table 4.7.3. Weighting reduces the difference between the two mean squares in the case of experimental unit-treatment interaction error and experimental unit error, and has little effect for the normally distributed random error.

TABLE 4.7.1

WEIGHTED EXPECTED VALUE OF PARAMETER ESTIMATES
EQUAL WITHIN-BLOCK ERROR VARIANCES

| Parameter: | | Three Types of Error Experimental: | | Normally |
Name	Value	Unit	Unit-Treat. Interaction	Distributed
μ	10.0	8.90	9.98	9.46
b_1	-1.5	-1.20	-1.50	-1.35
b_2	1.5	1.20	1.50	1.35
τ_1	0/-6.5	0/-6.5	-.38/-6.88	.24/-6.26
τ_2	0/-3.5	0/-3.5	.44/-3.06	.20/-3.30
τ_3	0/ 2.5	0/ 2.5	-.17/ 2.23	.08/ 2.42
τ_4	0/ 7.5	0/ 7.5	.10/ 7.60	-.35/ 7.15
β	2.0	2.11	2.002	2.056

TABLE 4.7.2

WEIGHTED STANDARD DEVIATION OF PARAMETER ESTIMATES
EQUAL WITHIN-BLOCK ERROR VARIANCES

| Parameter: | | Three Types of Error | | Normally |
Name	Value	Experimental: Unit	Unit-Treat. Interaction	Distributed
μ	10.0	10.8	5.42	7.27
b_1	-1.5	2.93	2.86	1.98
b_2	1.5	above	above	above
τ_1	0/-6.5	7.56	4.13	5.53
τ_2	0/-3.5	above	4.17	5.38
τ_3	0/ 2.5	above	4.04	5.13
τ_4	0/ 7.5	above	3.70	5.21
β	2.0	1.11	.58	.76

TABLE 4.7.3

WEIGHTED EXPECTATION & STANDARD DEVIATION OF ANOVA TERMS
EQUAL WITHIN-BLOCK ERROR VARIANCES

Anova Term	Three Types of Error Experimental:		Normally
Expectation:	Unit	Unit-Treat. Interaction	Distributed
MS Residual	48.61	25.72	40.58
MS $\tau=\mathbb{O}\|\mu,b,\beta$	47.59	25.48	39.31
MS $\tau\neq\mathbb{O}\|\mu,b,\beta$	115.72	94.52	103.23
MS $\beta\neq0\|\mu,b,\tau$	639.90	646.60	680.60
F $\tau=\mathbb{O}\|\mu,b,\beta$	35.89	7.62	60.28
F $\tau\neq\mathbb{O}\|\mu,b,\beta$	65.14	26.15	129.00
F $\beta\neq0\|\mu,b,\tau$	131.84	126.03	293.15
Standard Deviation:			
MS Residual	38.51	21.44	32.60
MS $\tau=\mathbb{O}\|\mu,b,\beta$	25.67	18.79	23.06
MS $\tau\neq\mathbb{O}\|\mu,b,\beta$	70.33	52.41	65.92
MS $\beta\neq0\|\mu,b,\tau$	316.5	380.1	395.75
F $\tau=\mathbb{O}\|\mu,b,\beta$	146.92	41.40	1162.8
F $\tau\neq\mathbb{O}\|\mu,b,\beta$	311.87	153.65	2464.9
F $\beta\neq0\|\mu,b,\tau$	447.89	487.87	4486.5

TABLE 4.7.4

WEIGHTED MEAN AND STANDARD DEVIATION OF ANOVA TERMS
UNEQUAL WITHIN-BLOCK ERROR VARIANCES

Anova Term Expectation:	Three Types of Error Experimental:		Normally Distributed
	Unit	Unit-Treat. Interaction	
F $\tau=0\|\mu,b,\beta$	2.47	4.29	5.45
F $\tau\neq0\|\mu,b,\beta$	5.02	18.72	13.95
F $\beta\neq0\|\mu,b,\tau$	21.10	111.89	66.79
Standard Deviation:			
F $\tau=0\|\mu,b,\beta$	3.05	14.80	31.53
F $\tau\neq0\|\mu,b,\beta$	6.53	66.45	97.95
F $\beta\neq0\|\mu,b,\tau$	18.95	409.10	303.65

4.7.3 Distribution of Parameter Estimates vs Cox's

Second Ratio

Graphs 4.7.1 - 4.7.4 show the error in estimation,
$\tau_1-\hat{\tau}_1$ or $\beta-\hat{\beta}$, vs the value of Cox's second ratio (4.7.2),
(p. 322), for the two experimental unit error cases. The
ratio is on the lower, horizontal, axis; the error in
estimation is on the lefthand, vertical, axis. The graphs
for the normally distributed random error case are nearly
identical to those for the experimental unit-treatment
interaction error case, and are omitted. Graphs for all
levels, k, and all values ($\tau=0$ and $\tau\neq0$) of $\tau.-\hat{\tau}.$ are

identical; only those for k=1 are shown. Graphs for

unequal within-block error variances are omitted as they

are nearly identical to those shown, which all have equal

within-block error variances. Recall that the second

ratio is linearly and inversely related to the third ratio

used to weight the randomizations of Tables 4.7.1-4.7.4.

Each graph indicates that, for this simulation,

selecting a randomization with a small value of ratio two

(or equivalently, a large value of ratio three) avoids

randomizations with large errors in estimating τ_1 and β.

Graphs 4.7.1 and 4.7.2 for $\tau_1 - \hat{\tau}_1$ display, respec-

tively, the cases of experimental unit errors and

experimental unit-treatment interaction errors for the

case of equal within-block error variances. For

experimental unit errors, Graph 4.7.1, identical values

of $\tau_1 - \hat{\tau}_1$ repeat six times (F=6); adjacent groups of six

are combined and are represented by (L=12). The cause of

this is believed to be related to repetitions of certain

values within g-groups of randomizations. In both

graphs, the distribution appears centered about zero and

appears so for all treatments. Table 4.4.1, (p. 259),

shows the expectation to be zero for experimental unit

error, but to be slightly biased for the other two

errors. Thus, in the latter case, the observations are

not centered, in the sense of the mean, about zero.

```
                The Randomization Model with Blocks and a Covariate
                        Data set is  sasglm.glm002

          Plot of STR1*RATIO2.  Legend: A = 1 obs, B = 2 obs, etc.

       80 +
S
-
e
                                                         F
       60 +
e
r
       40 +                                              F
r
S  20 +
-                                             F
e
r                                    . F   F      F       F
r                                       F    F            FF
          LLL        F  L  FRF   FFFLL       L         FF
t    0 +--------L-------L-F-FRF--FLL---F-F-------FF--------
r          FFX        F     LF R LRFFL F              L
i          FF                    FFF  F  F            L
h                        F  FF      F                 F
a
t
a -20 +


      -40 +                    ,                        F


      -60 +                                             F


      -80 +
          -+-----------+-----------+-----------+-----------+-
           8           9          10          11          12

                  R2-for Z: Adj. Treat. ss/E(residual ms)
```
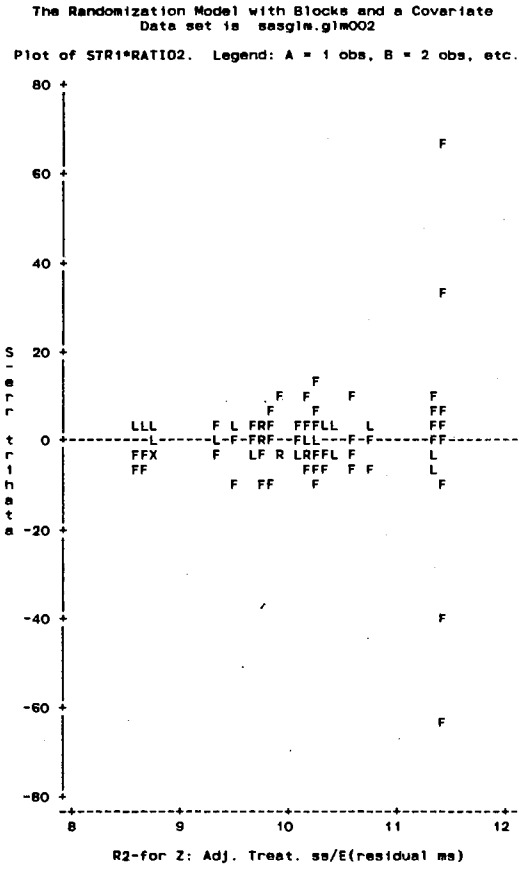
Graph 4.7.1
The Distribution of $\tau_1 - \hat{\tau}_1$ vs Ratio Two
Experimental Unit Error

```
          The Randomization Model with Blocks and a Covariate
                    Data set is  sasglm.glm004

        Plot of STR1*RATIO2.  Legend: A = 1 obs, B = 2 obs, etc.
          |
       40 +
          |
          |
       30 +                                    A
          |                                    A
          |
          |                                  . A
       20 +
     S    |                                   AA
     -    |                                   AA
     e    |                                   AB
     r    |                                   DB
     r 10 +       A A      A                   E
          |              A B                   EB
     t    |  AA     A A AAA   ABCB  A A        DA
     r    |  ADC    B B DDCC BFBAC  D B        HD
     1    |  DBJ    A B BGC  BFKDE  C C        FC
     h    |  EDI    C B BIDB AJJEE  D I        DC
     a  0 +--EDG------D-E-ELED-BHFCC-C-B------CC--------
     t    |  OBE     E D EHBF GBFDD  D B       CB
     a    |  EBI     G D CDDF EGCAB  A B       BG
          |  CC      B ABA  AEDBB             C
          |  B       A AA A AABB   B          C
          |  A          A       A A        A
      -10 +                             A        A
          |                   AA      B          A
          |                                      A
          |                                      A
          |                                      B
      -20 +                                      A
          |                   .
          |
          |
      -30 +                                      A
          |
        --+------------+------------+------------+------------+--
          8            9           10           11           12

              R2-for Z: Adj. Treat. ss/E(residual ms)
```

Graph 4.7.2
The Distribution of $\tau_1-\hat{\tau}_1$ vs Ratio Two
Experimental Unit-Treatment Error

Graphs 4.7.3 and 4.7.4 for $\beta-\hat{\beta}$ display, respectively, the cases of experimental unit errors and experimental unit-treatment interaction errors for the case of equal within-block error variances. For experimental unit errors, Graph 4.7.3, identical values repeat 24 times (X=24), as proved in Theorem 3.7.9(4), (p. 202). The highly erroneous estimate of $\hat{\beta}$ = -8.95 (β = 2.0) can be avoided by assigning a zero selection probability value

to $p(s)$ for those randomizations with large values of the
second ratio.   As seen along the horizontal line at zero
in Graphs 4.7.3 and 4.7.4, unbiased estimates of β are
obtained by some randomizations throughout the range of
the second ratio.   This suggests searching for an
alternative function which would better distinguish
randomizations yielding unbiased estimates.   Selecting
the randomization used to conduct the actual experiment
on the basis of the second ratio (4.7.2) (or third ratio)
discards many unbiased or low biased randomizations for
the gain of avoiding those with large bias.

The Randomization Model with Blocks and a Covariate
Data set is sasglm.glm002

Plot of SBETA*RATIO2.   Legend: A = 1 obs, B = 2 obs, etc.

R2-for Z: Adj. Treat. ss/E(residual ms)
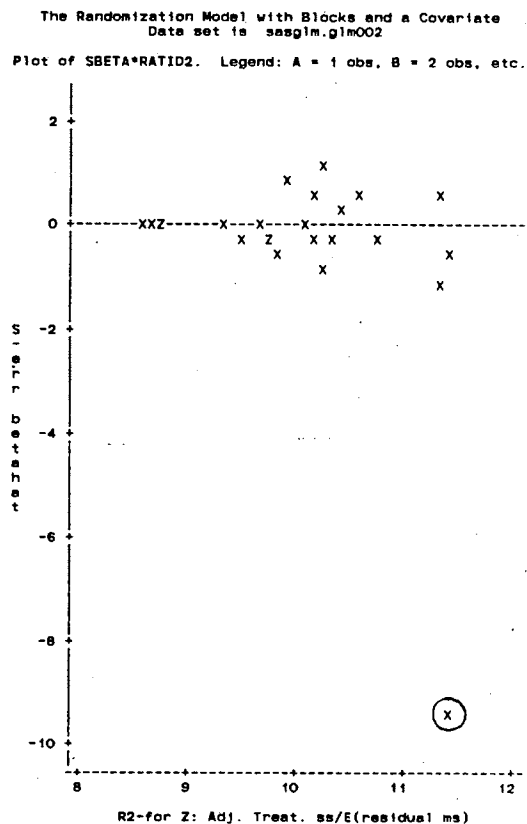
Graph 4.7.3
The Distribution of $\beta - \hat{\beta}$ vs Ratio Two
Experimental Unit Error

```
                  The Randomization Model with Blocks and a Covariate
                              Data set is  sasglm.glm004

               Plot of SBETA*RATIO2.   Legend: A = 1 obs. B = 2 obs. etc.

                    |
                 6  +
                    |
                    |
                    |
                    |
                    |                                        A
                    |
                 4  +                                        A
                    |
                    |                                        A
                    |
                    |                                        A
                    |                                        D
             S   2  +                                        A
             -      |                                        BB
             e      |                           B     A      DB
             r      |              A B  AB        B          CA
             r      |     B B     B  BE  BAC A A E           DB
                    |     BDE     B E CCBC CGHBC H B         ED
             b      |     DDI     F C CEED AHHGG E B         BC
             e   0  +----HHP------B-E-HPGA-EJKFC-C-D------CC--------
             t      |     CFI     G H GHEE MMHBC B C         CD
             a      |     DBE     E C AGDI  CDEC B A         CB
             h      |     A B        CA   CBBB A B           CD
             a      |                     BB   B             BB
             t      |                                        GA
                    |                            B           DA
                -2  +                                        BA
                    |                                        AB
                    |                                        B
                    |
                    |                                        B
                    |
                -4  +                                        A
                    |
                    |
                    |
                    |
                -6  +
                    |
                    --+---------+---------+---------+---------+--
                      8         9        10        11        12

                        R2-for Z: Adj. Treat. ss/E(residual ms)
```
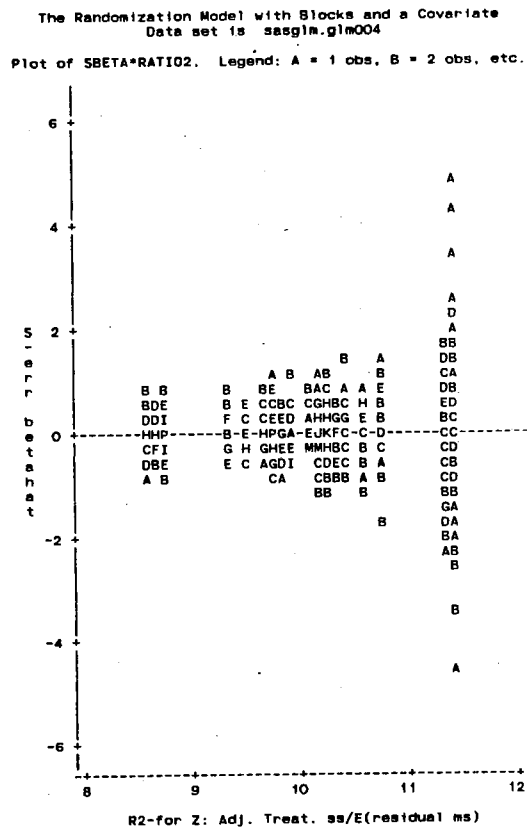
Graph 4.7.4
The Distribution of $\beta - \hat{\beta}$ vs Ratio Two
Experimental Unit-Treatment Error

The final section considers restricting randomization. Instead of assigning varing selection probabilities as in this section, $p(s)$ is either zero or a constant. The noncentrality parameter for the adjusted covariate is used to select those randomizations with nonzero selection probabilities.

Section 4.8

G-groups, Noncentrality Parameters, and OSL's

Section 4.8 discusses the use of the noncentrality

parameter for the covariate as adjusted for the mean,

blocks, and treatment, $\lambda_{s,\beta|\mu,b,\tau}$, to select those random-

izations better able to detect nonzero values of β and τ.

Section 4.8.1 contains examples of Theorem 3.7.15,

(p. 219), which plays a key role throughout Section 4.8.

Section 4.8.2 suggests that Cox's second (and third)

ratio (4.7.2), (p. 322), for a single covariate is a

linear function of this non-centrality parameter. The

graphs of Section 4.8.3 show the effect on $H_{o;\mu,b,\tau,\beta}\beta=0$

of using $\lambda_{s,\beta|\mu,b,\tau}$ to select the randomization used to

test the hypothesis. Section 4.8.4 displays together the

OSL's for the usual two hypotheses as one moves from the

lowest to the highest values of $\lambda_{s,\beta|\mu,b,\tau}$. Section

4.8.5 compairs the means and standard deviations of the

estimated model equation parameters as computed within

(a) the lowest and (b) the highest values of $\lambda_{s,\beta|\mu,b,\tau}$.

The major result is that for, this simulation,

selecting the randomization used to conduct the experi-

ment from among the g-group with the largest value of

$\lambda_{s,\beta|\mu,b,\beta}$ improves the probability of detecting nonzero

covariates for all error types, but offers only modest

help in detecting nonzero treatments in the cases of

experimental unit-treatment interaction and normally

distributed random errors and no help in the case of experimental unit error. If proper, use of the singular normal distribution may improve the situation for experimental unit errors. Limiting randomization to randomizations within the maximum-valued g-group substantially improves the expectations and standard deviations of the model parameter estimates for all three error types. These randomizations are the most balanced on the covariate.

4.8.1 A Comparison of the Two Noncentrality Parameters

Graphs 4.8.1 and 4.8.2 illustrate Theorem 3.7.15, in particular equation (3.7.34), (p. 219), for the three types of error. The lower, horizontal, axis is the noncentrality parameter for the covariate as adjusted for the mean, blocks, and treatments. The lefthand, vertical, axis is for the noncentrality parameter for the treatments as adjusted for the mean, blocks, and covariate.

The errors enter the noncentrality parameters only via the variance. The experimental unit and normally distributed random errors have constant and equal-valued variances, thus, their graphs are identical; both cases are represented by Graph 4.8.1. The vertical lines display groups with equal values of the noncentrality parameter for the covariates as adjusted for the mean, blocks, and treatments. These are the g-groups of Theorem 3.7.10, (p. 211). The letters on the graphs label the g-groups.

The variance for the experimental unit-treatment interaction errors changes with each randomization, as discussed in Section 4.2.3, (p. 252). As shown in Graph 4.8.2, this disperses the values of each g-group, blurring the columns seen in Graph 4.8.1.

Each graph shows that one can avoid low-valued noncentrality parameters for the adjusted treatments by restricting randomization to g-groups with large values of the noncentrality parameter for the adjusted covariate. The desirable randomizations are those in the top righthand corner of the graphs, or at least the right-most column. As discussed in (2.6.6a), (p. 41), one thus avoids randomizations having low power to detect nonzero model equation parameters. Such lower-powered randomizations are to the left and/or bottom of the graphs. In the case of a single covariate, the desired g-groups may be discovered without knowledge of the true value of the covariate slope parameter. The discovered g-group(s) will be favorable for all values of the treatment effects.

Consider the twenty-four points (X,Y) of Graph 4.8.1 to be: $X = \lambda_{g,\beta|\mu,b,\tau}$, on the horizontal axis and

$Y = \min\{\lambda_{s,\tau|\mu,b,\beta}\}$, on the vertical axis.

A few such points are circled on the graph. The Y value is the bottom point in each of the twenty-four columns of numbers. The least squares fit of a straight line through these points, as in (3.7.35), (p. 221),

estimates the slope parameter as .22.   This value is a

close approximation to the ratio $\{\lambda_{f,T} / \lambda_{f,F}\} = .235$.   This

ratio is part of (3.7.34), (p. 219).   As discussed below

(3.7.35), (p. 221), one increases the minimum value of

the noncentrality parameter for the adjusted treatments

by .22 for each unit increase in the noncentrality

parameter for the adjusted covariate.   Had the slope been

near zero, no g-group of randomizations would have been

advantageous in avoiding low-valued, adjusted-treatment,

noncentrality parameters.
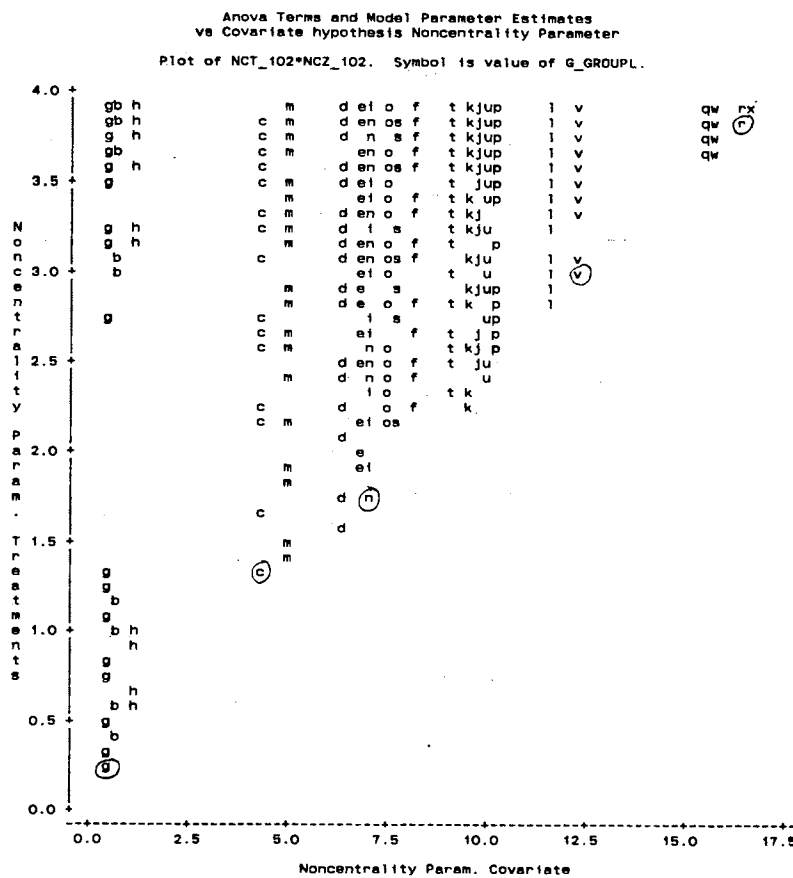
Graph 4.8.1
Adjusted Noncentrality Parameters
Experimental Unit and Normally Distributed Random Errors

Anova Terms and Model Parameter Estimates
vs Covariate hypothesis Noncentrality Parameter

Plot of NCT_104*NCZ_104. Symbol is value of G_GROUPL.
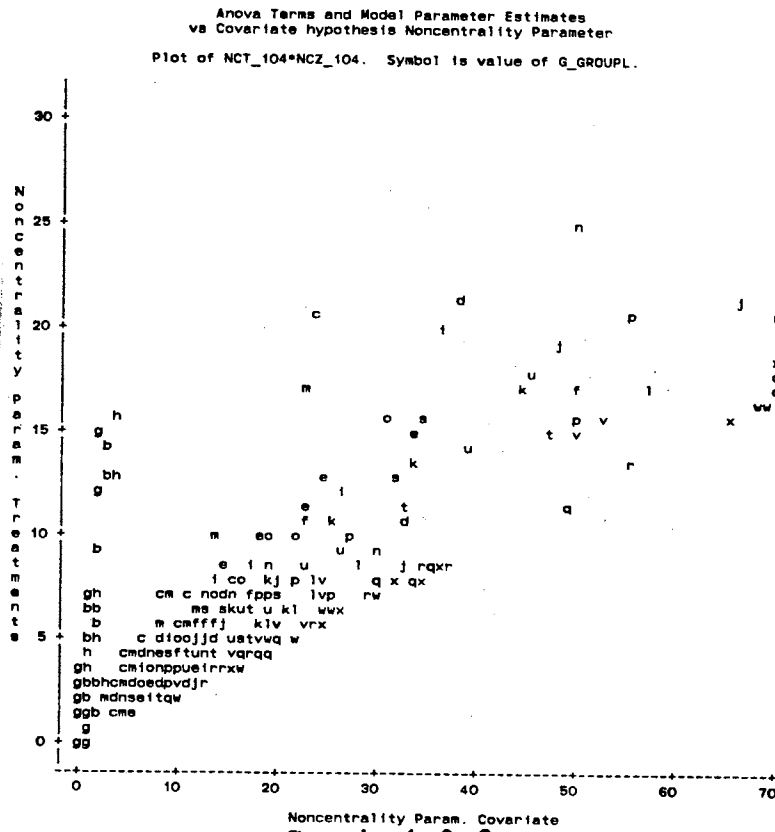
Graph 4.8.2

Adjusted Noncentrality Parameters
Experimental Unit-Treatment Interaction Errors

4.8.2 Cox's Second Ratio and $\lambda_{\beta|\mu,b,\tau}$

Graph 4.8.3 shows the apparent linear relationship between Cox's second ratio and the noncentrality parameter for the covariate as adjusted for the mean, blocks, and treatments. Values for Cox's second ratio are on the vertical axis. The horizontal axis shows the noncentrality parameter prior to multiplication by $.5\sigma^2$; the axis plots $.5\sigma^2\lambda_{s,\beta|\mu,b,\tau}$. This linear relationship suggests Theorem 3.7.5, (p. 190), and (2.6.6a),

(p. 41), as rationale for applying Cox's second or third
ratio for the randomization process. An algebraic proof
of the relationship of these ratios and the noncentrality
parameter is an extension.
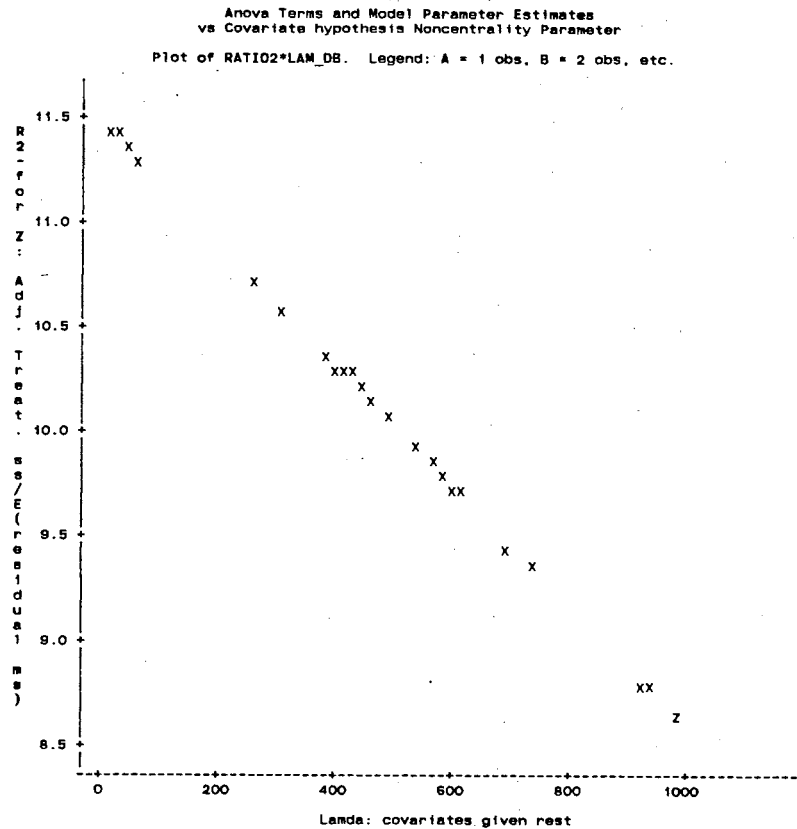
Graph 4.8.3

Cox's Second Ratio vs $.5\sigma^2\lambda_{s,\beta|\mu,b,\tau}$

4.8.3 Restricting Randomization via $\lambda_{s,\beta|\mu,b,\tau}$

According to (2.6.6a), (p. 41), the best randomiza-
tions for detecting a nonzero covariate slope coefficient
are the ones with the largest noncentrality parameter for
the adjusted covariate slope parameter. The graphs of
this section display the OSL for the hypothesis test

$H_{o;\mu,b,\tau,\beta}\beta=0$. Since $\beta=2.0$, this hypothesis should be rejected. Small values of the OSL, say, less than .10, do this; larger values result in accepting a false null hypothesis, a Type II error. Each of Graphs 4.8.4-4.8.6 shows that large values of $\lambda_{\beta|\mu,b,\tau}$, to the right on the lower, horizontal, axis, avoid large values of the OSL, that is, avoid a Type II error for this hypothesis test. These are the points below the horizontal dashed line. There is one graph for each of the three error types.

Theorem 3.7.10, (p. 211), proves the clustering of noncentrality parameter values seen along the lower, horizontal, axis of Graph 4.8.4, for experimental unit error, and of Graph 4.8.6, for normally distributed random error. When the variance is not constant for each randomization, as for the experimental unit-treatment interaction error of Graph 4.8.5, those clusters on the horizontal axis are dispersed.

Theorem 3.7.9(3), (p. 202), proves the clustering of OSL values along the lefthand, vertical, axis of Graph 4.8.4 for experimental unit error. This error is constant for all randomizations. In contrast, the values of the errors η_S and n_S change with each randomization. This disperses the clusters on the vertical axis, as discussed in Theorem 3.7.9, and as seen in Graph 4.6.5, for experimental unit-treatment interaction error η_S, and in Graph 4.6.6, for normally distributed random error n_S.

Anova Terms and Model Parameter Estimates
vs Covariate hypothesis Noncentrality Parameter

Plot of SZC002*NCZ_102.   Symbol is value of G_GROUPL.

```
O
S
L 0.6 +
Z
  0.5 +
M
B 0.4 +        b
T
  0.3 +
      g                          d
  0.2 +
       h
                                    f
           c          o        up
  0.1 +-----------------------------------------v------------------------------
              e          t
              n                                        qw
  0.0 +                           j        l
                m      i  s      k                      rx
```

```
   +----------+----------+----------+----------+----------+----------+----------+
  0.0        2.5        5.0        7.5       10.0       12.5       15.0       17.5
```

Noncentrality Param. Covariate
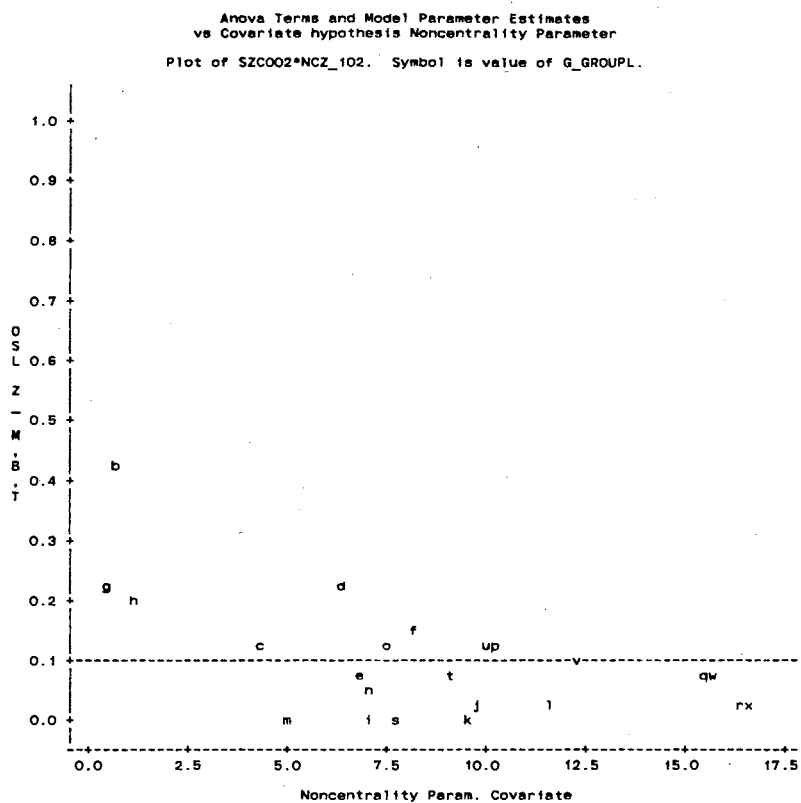
Graph 4.8.4

OSL's   vs   Noncentrality Parameters for β
Experimental Unit Errors

```
                    Anova Terms and Model Parameter Estimates
                    vs Covariate hypothesis Noncentrality Parameter

                 Plot of SZC004*NCZ_104.  Legend: A = 1 obs, B = 2 obs, etc.

       |
   1.0 +
       |  B
       |  BB
       |  A
   0.9 + B
       |
       |
       |  A
   0.8 + A  A
       |  A
       |
       |  C
   0.7 + AA
       |  B
  D    |  B  A
  S    |  A  A
  L 0.6 + C
       |  A
  Z    |  BB A
       |  A
  ‾ 0.5 +
  M    |    A A
  .    |  B
  B    |  A A
  .  0.4 +  BA   A
  T    |   AA   A
       |  ACA A   A
       |   B  A      A
   0.3 + AAA A
       |  ACB
       |   AA    AA
       |  AA  A  A  A   A
   0.2 +  CAA  A  A  C  AA
       | ADAAAAA    A A    A
       |  BA    ABDABABAA
       |  AC    BB A A BAAA A
   0.1 +-AA--BACBCFBDBAAA-A-B-----A------------------------------------------
       |   C   ACAEFECGCABAAAA A    A    A
       |  BA   CBEFJCHDDCGDBCDBFA  A B      A  BA  A
       |  B   BEHIEIFFDJGLFDDDDIDACBDC AAAAA AAADAABAA     A  A  C     A A        A  BAC
   0.0 +  A A CCBCDGFDFDFADDCCEBABB ACC  AC ABAA  A      AA  AAA  A  A         A   B
       |
        --+----------+----------+----------+----------+----------+----------+----------+-
          0         10         20         30         40         50         60         70
                             Noncentrality Param. Covariate
```

Graph 4.8.5

OSL's vs Noncentrality Parameters for β
Experimental Unit-Treatment Interaction Errors

```
              Anova Terms and Model Parameter Estimates
             vs Covariate hypothesis Noncentrality Parameter

          Plot of SZCO46*NCZ_146.  Legend: A = 1 obs, B = 2 obs, etc.

     1.0 +  A
        |   AA
        |   AA
        |   A   B
     0.9 +      A
        |
        |   A   A
     0.8 +  A A A
        |   AA
        |   A A
        |   A
     0.7 +  A              A
   O    |   BA A
   S    |   A
   L 0.6 +  A              A
        |   A A
   Z    |   A  A
        |   AAB
     0.5 +  B  A
   M    |   ABB
   .    |   A
   B    |   B B
   . 0.4 +  A          A
   T    |   AA         A      A
        |              B      A
     0.3 +  BA  A      A      A
        |   A   A      B      B A
        |   AAA A          A  A  A
        |   B   A      A       B  B        A
     0.2 +  B          C     B AD AC A     A
        |   A  A       B  A    BC BA C   A   BC
        |   A   A      B      B A DB A   C AB      C B
        |   BC  B      B  C   B  B AC C  C A AA      A              DA
     0.1 +----B-B------------B--A-----B-CE--B-C---D-BCAC-----B--A------------CE--BB----
        |   AA          B  C    B CC DC A   A DEEB   B  C          BB  CG
        |      B B       A  A    E DH ED C   D CEDE   C  C          DA  DA
        |    C C B      F  C    B GK CB D    D EEFH   I  I          DG  JE
     0.0 +  A   B       B  D    E CE CA E    D FOEB   E  E          GH  EI
        |
        +----------+---------+---------+---------+---------+---------+---------+-
           0.0      2.5       5.0       7.5      10.0      12.5      15.0      17.5

                         Noncentrality Param. Covariate
```

Graph 4.8.6

OSL's vs Noncentrality Parameters for β
Normally Distributed Random Errors

4.8.4 OSL's for the Best and Worst Cases

Displays 4.8.7 and 4.8.8 present the OSL's for the adjusted treatments vs those for the adjusted covariate. The randomizations are grouped by equal values of the noncentrality parameter for the adjusted covariate. This is equivalent to grouping by the g-groups. The value used is $.5\sigma^2 \lambda_{g,\beta|\mu,b,\tau}$, with 10 the minimum value (Display 4.8.7) and 990 the maximum (Display 4.8.8). The

multiplication by $.5\sigma^2$ avoids the changes in σ^2 from one
randomization to the next which occur in the experimental
unit-treatment interaction errors.

In the top row of three graphs, the OSL's for the
adjusted treatments (lefthand scale) use $\tau=0$. The
central F-distribution provides the OSL's for the top
row. In the bottom row of three graphs, the OSL's for the
adjusted treatments use the nonzero τ of this simulation.
The noncentral F-distribution provides the OSL's of the
bottom row. The noncentrality parameter, λ, is for the
appropriate randomization and variance. The covariate
slope coefficient, β, is always equal to 2.0. The noncen-
tral F-distribution provides the OSL's for the adjusted
covariate (the bottom scale of both rows of graphs). The
usual degrees of freedom are used in all cases.

The left-most column of two graphs, one above the
other, uses experimental unit errors. The center column
of two graphs uses experimental unit-treatment
interaction errors, and the right-most column of two
graphs uses normally distributed random errors.

Each graph has the OSL for the adjusted treatments
on the lefthand, vertical, axis and the OSL for the
adjusted covariates on the lower, horizontal, axis.

For experimental unit error, the left-most column of
two graphs of both displays, the covariate values are
clustered, as per Theorem 3.7.9(3), (p. 202). This
provides the columns of values seen in these graphs in

both displays.   When the treatment effects are zero, as
in the top lefthand graph of both displays, the
clustering discussed in Theorem 3.7.9(5), (p. 203), also
operates.   Together, the two sources of clustering result
in the single point of the top, left-most graph of both
displays.   The OSL for the adjusted treatments is the
same for each of the display's 24 randomizations, as is
the OSL for the adjusted covariates, hence the single
point.

In all plots, the value of β is nonzero.   One wants
to reject this null hypothesis, $H_{o:\ \beta,\mu,b,\tau}\beta=0$.   Thus, one
wants low valued OSL's, OSL's to the left of each graph.
Note that as one moves from minimum values of the non-
centrality parameter (Display 4.8.7) to the maximum value
(Display 4.8.8), the OSL's shift to the left, to lower
values, in the center and right-most columns of graphs.
These two columns represent experimental unit-treatment
interaction and normally distributed random errors.
While most randomizations of these two error types of
Display 4.8.7 make the Type II error of accepting the
erroneous null hypothesis $H_{o;\ \beta,\mu,b,\tau}\beta=0$, for $\alpha=.12$, none
of the randomizations of Display 4.8.8 commit a Type
II error.   By restricting randomization to those rando-
mizations with the maximum value of the noncentrality
parameter (Display 4.8.8), the researcher improves
his/her chance of selecting a randomization which will
lead to the correct hypothesis test conclusion when $\beta\neq0$.

In contrast to the experimental unit-treatment
interaction and normally distributed random errors, the
experimental unit errors appear to be unaffected by the
shift in the value of the adjusted covariate's non-
centrality parameter. The column of values in the
left-most column of graphs has the same OSL for the
adjusted covariate hypothesis in both displays, about
.05. Increasing the value of this noncentrality parameter
did not lower OSL's for this error type. The advice of
(2.6.6a), (p. 41), applys to distributions affected by a
noncentrality parameter or a function of it. The failure
to lower the OSL's suggests that the F-distribution is
not applicable for this error type, at least not in the
upper tail. Recall that the probability plot of Graph
4.6.12, (p. 219), also suggested a poor fit of the upper
tail of the F-distribution for experimental unit error.
Use of the singular normal distribution, if proper,
may result in the g-group affecting the OSL for this
error type.

For the top row of three graphs, the null hypothesis
of $H_{0:\tau|\mu,b,\beta}\tau = 0$ is true. One wants to accept this
hypothesis. Thus, one wants high-valued OSL's, ones
larger than .10 towards the top of each graph. However,
in this case, the noncentrality parameter is zero and
Theorem 3.7.15, (p. 219), does not apply. The OSL's for
experimental unit error, the top left-most graph, take a
higher value as one moves from minimum to maximum values

of the noncentrality parameter (Display 4.8.7 to 4.8.8).
Some OSL's for experimental unit-treatment interaction
error and normally distributed random error take lower
values, but the effect of the noncentrality parameter
for the adjusted covariates upon this OSL is small.

For the bottom row of three graphs, the null
hypothesis of $H_{0:\tau,\mu,b,\beta}\tau=0$ is false.   One wants to
reject this hypothesis. Thus, one wants low-valued
OSL's, ones near the bottom of the graphs, ideally below,
say $\alpha=.10$. The center and right-most graphs, for
experimental unit-treatment interaction and normally
distributed random errors, show a slight improvement at
the larger value of the adjusted covariate's noncentral-
ity parameter (Display 4.8.8). More randomizations have
adjusted treatment OSL's below .10 in Display 4.8.8 than
in Display 4.8.7. For experimental unit-treatment
interaction, the change is from none to four randomiza-
tions; for normally distributed random error, the change
is from six to ten. Again, by restricting randomization
to those randomizations with the maximum value of the
noncentrality parameter, the researcher improves his/her
chance of selecting a randomization which will lead to
the correct conclusion.

For experimental unit error, the lower, left-most
graph, the number of randomizations correctly rejecting
this null hypothesis decreases as one moves from the
lowest to the highest noncentrality parameter (Display

4.8.7 to 4.8.8). Again, this contradicts the advice of
(2.6.6a), (p. 41), suggesting that the F-distribution is
not applicable for this error type, at least in the upper
tail.

While the effect of the value of the noncentrality
parameter for the adjusted covariate appears large for
the hypothesis test of this parameter, the effect upon
the hypothesis test of the adjusted treatments appears to
be small, and limited to the cases of experimental
unit-treatment interaction and normally distributed
random errors. This may or may not hold in other
simulations. Possibly, this noncentrality parameter has
little effect on the tail probabilities for the degrees
of freedom used to compute the OSL's for the adjusted
treatments. Examination of such is an extension.

Observed Significance Levels for Treatments and Covariates
Treatments are Zero in Top Row, Nonzero in Bottom Row
Left to Right: experimental unit error, experimental-unit
treatment interaction error, normal random error

-- Lamda: covariates given rest=10 ---

A=1, B=2, etc. Plot of STC002*SZC002.

```
      -+-----+-----+-----+-----+-----+-
 1.00 +                                +
0     |                                |
S 0.75+-----+-----+-----+-----+-----+--+
L     |                                |
T 0.50+-----+-----+-----+-----+-----+--+
-     |                                |
M 0.25+-----+-----+-----+--~--+-----+--+
.
B
, 0.10+--X--+-----+-----+-----+-----+--+
Z     |                                |
 0.00 +                                +
      -+-----+-----+-----+-----+-----+-
     0.00  0.10  0.25  0.50  0.75 1.00
```

OSL Z_ M,B,T

A=1, B=2, etc. Plot of STC004*SZC004.

```
      -+-----+-----+-----+-----+-----+-
 1.00 +         |A       AA|     A +
0     |                A  A        |
S 0.75+-----+-----+--A+-----+--A---A--+
L     |            |ABA      A       |
      |            |A   A            |
T 0.50+-----+--A---+-A---+-----+-----A+
-     |      A          |   A        |
M 0.25+-----+-----+-----+-----+-----+--+
B
, 0.10+-----+-----+-----+-----+-----+--+
Z     |                                |
 0.00 +                                +
      -+-----+-----+-----+-----+-----+-
     0.00  0.10  0.25  0.50  0.75 1.00
```

OSL Z_ M,B,T

A=1, B=2, etc. Plot of STC046*SZC046.

```
      -+-----+-----+-----+-----+-----+-
 1.00 +          |      A A|   |A  A |
0     .                    A          |
S 0.75+-----+-----+--A----+-----+-----+
L     |               |A  |A      A  |
      |            A  |  A  |A  A  A |
T 0.50+-----+-----+--A--+-----+-----+
-     |        |A  |A     |         |
M 0.25+-----+-----+--A+-----+-----+
B     |     A |AA   |               |
, 0.10+-----+-----+-----+-----+-----+
Z     |   A                         |
 0.00 + B                           +
      -+-----+-----+-----+-----+-----+-
     0.00  0.10  0.25  0.50  0.75 1.00
```

OSL Z_ M,B,T

A=1, B=2, etc. Plot of STC102*SZC002.

```
      -+-----+-----+-----+-----+-----+-
 1.00 +                                +
0     |                                |
S 0.75+-----+-----+-----+-----+-----+--+
L     |                                |
T 0.50+-----+-----+-----+-----+-----+--+
-     |                                |
M 0.25+-----+-----+-----+-----+-----+--+
B     | A                             |
, 0.10+--D--+-----+-----+-----+-----+--+
Z     | S                             |
 0.00 +                                +
      -+-----+-----+-----+-----+-----+-
     0.00  0.10  0.25  0.50  0.75 1.00
```

OSL Z_ M,B,T

A=1, B=2, etc. Plot of STC104*SZC004.

```
      -+-----+-----+-----+-----+-----+-
 1.00 +           |A      A |
0     |         A|  |     |  |
S 0.75+-----+-----+-----+-----+
L     |           |AA|     A  |
T 0.50+-----+-----+--AA-----+
-     |           |   A     B |
M 0.25+-----+-----+--A--+-----+
B     |     A  |A          |
, 0.10+-----+--A--+-A---+-----A+
Z     |                         |
 0.00 +                         +
      -+-----+-----+-----+-----+-----+-
     0.00  0.10  0.25  0.50  0.75 1.00
```

OSL Z_ M,B,T

A=1, B=2, etc. Plot of STC146*SZC046.

```
      -+-----+-----+-----+-----+-----+-
 1.00 +              |  A|     |
0     |              |   |     |  |
S 0.75+-----+-----+-----A-----+---A-+
L     |                        A |
T 0.50+-----+--A-+B-----+-----+-A----+
-     |           |A            A  A|
M 0.25+-----+-----+--AA----+-A--+
B     |        |A  |     A| A      |
, 0.10+-----+--+A--+-----+-----+
Z     |     A |A A |             |
 0.00 + B                       +
      -+-----+-----+-----+-----+-----+-
     0.00  0.10  0.25  0.50  0.75 1.00
```

OSL Z_ M,B,T

Display 4.8.7
OSL's for β vs OSL' for τ
Minimum Noncentrality Parameter for β

Observed Significance Levels for Treatments and Covariates
Treatments are Zero in Top Row, Nonzero in Bottom Row
Left to Right: experimental unit error, experimental-unit
treatment interaction error, normal random error

-- Lamda: covariates given rest=990 --

A=1, B=2, etc. Plot of STC002*SZC002.   A=1, B=2, etc. Plot of STC004*SZC004.   A=1, B=2, etc. Plot of STC046*SZC046.

```
        -+----+----+----+----+----+-                    -+----+----+----+----+----+-                    -+----+----+----+----+----+-
   1.00 +    |    |    |    |    |    +              1.00 +  A A|    |    |    |    +              1.00 +    A |    |    |    |    +
 O      |    |    |    |    |    |    |            O      |  B  |    |    |    |    |            O      |  AAB|    |    |    |    |
 S 0.75 +----+----+----+----+----+----+            S 0.75 +--B--+----+----+----+----+            S 0.75 +----BB+----+----+----+----+
 L      |    |    |    |    |    |    |            L      |     |    |    |    |    |            L      |    |    |    |    |    |
 T 0.50 +----+----+----+----+----+----+            T 0.50 +-A--+----+----+----+----+            T 0.50 +----+----+----+----+----+
 _      |    |    |    |    |    |    |            _      | BA |    |    |    |    |            _      | AA |    |    |    |    |
 M 0.25 +  X +----+----+----+----+----+            M 0.25 +-A--+----+----+----+----+            M 0.25 +  A +----+----+----+----+
 .      |    |    |    |    |    |    |            .      | A  |    |    |    |    |            .      |    |    |    |    |    |
 B      |    |    |    |    |    |    |            B      | A  |    |    |    |    |            B      | BA |    |    |    |    |
 . 0.10 +----+----+----+----+----+----+            . 0.10 +----+----+----+----+----+            . 0.10 +-B--+----+----+----+----+
 Z      |    |    |    |    |    |    |            Z      |    |    |    |    |    |            Z      | AB |    |    |    |    |
 0.00 +    |    |    |    |    |    +              0.00 +    |    |    |    |    +              0.00 + B  |    |    |    |    +
        -+----+----+----+----+----+-                    -+----+----+----+----+----+-                    -+----+----+----+----+----+-
       0.00  0.10  0.25  0.50  0.75 1.00              0.00  0.10  0.25  0.50  0.75 1.00              0.00  0.10  0.25  0.50  0.75 1.00
```

              OSL Z_ M,B,T   OSL Z_ M,B,T   OSL Z_ M,B,T
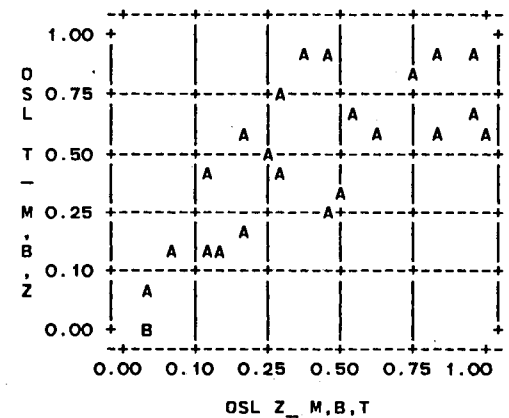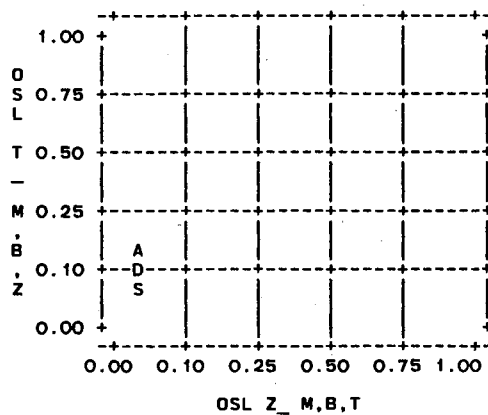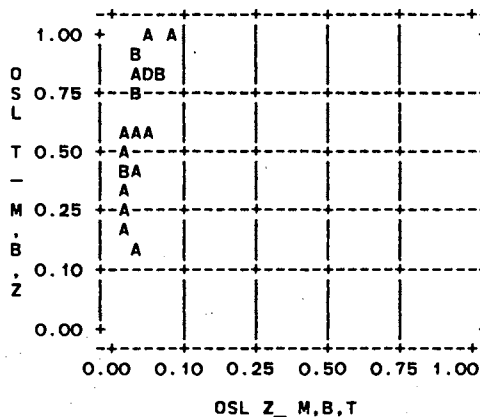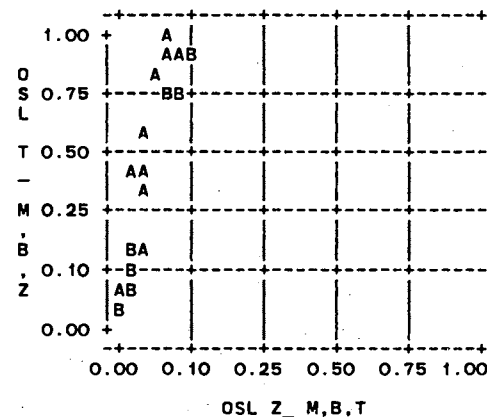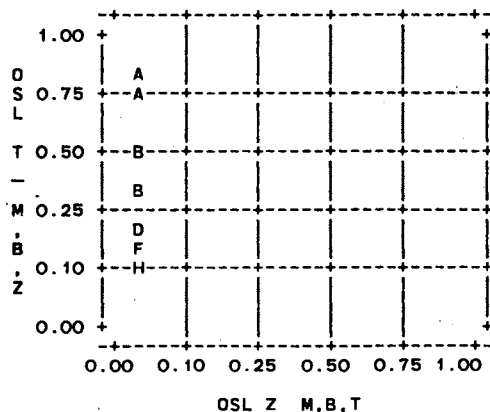
A=1, B=2, etc. Plot of STC102*SZC002.   A=1, B=2, etc. Plot of STC104*SZC004.   A=1, B=2, etc. Plot of STC146*SZC046.

```
        -+----+----+----+----+----+-                    -+----+----+----+----+----+-                    -+----+----+----+----+----+-
   1.00 +    |    |    |    |    |    +              1.00 +    |    |    |    |    +              1.00 +    |    |    |    |    +
 O      |  A |    |    |    |    |    |            O      |    |    |    |    |    |            O      |    A |    |    |    |    |
 S 0.75 +--A--+----+----+----+----+----+            S 0.75 +----+----+----+----+----+            S 0.75 +--A--+----+----+----+----+
 L      |    |    |    |    |    |    |            L      |    |    |    |    |    |            L      |  BA|    |    |    |    |
 T 0.50 +-B--+----+----+----+----+----+            T 0.50 +--A-AA+----+----+----+----+            T 0.50 +---AAA+----+----+----+----+
 _      |    |    |    |    |    |    |            _      | ABA |    |    |    |    |            _      |  A |    |    |    |    |
 M      | B  |    |    |    |    |    |            M      | BAA |    |    |    |    |            M      |    | B  |    |    |    |
 0.25 +----+----+----+----+----+----+            M 0.25 +--AC-+----+----+----+----+            M 0.25 +-A--+----+----+----+----+
 .      | D  |    |    |    |    |    |            .      | AA |    |    |    |    |            .      |    A|    |    |    |    |
 B      | F  |    |    |    |    |    |            B      | A A|    |    |    |    |            B      |    |    |    |    |    |
 . 0.10 +-H--+----+----+----+----+----+            . 0.10 +-A--+----+----+----+----+            . 0.10 +----+----+----+----+----+
 Z      |    |    |    |    |    |    |            Z      | AA |    |    |    |    |            Z      | AAA|    |    |    |    |
 0.00 +    |    |    |    |    |    +              0.00 + A  |    |    |    |    +              0.00 + BE |    |    |    |    +
        -+----+----+----+----+----+-                    -+----+----+----+----+----+-                    -+----+----+----+----+----+-
       0.00  0.10  0.25  0.50  0.75 1.00              0.00  0.10  0.25  0.50  0.75 1.00              0.00  0.10  0.25  0.50  0.75 1.00
```
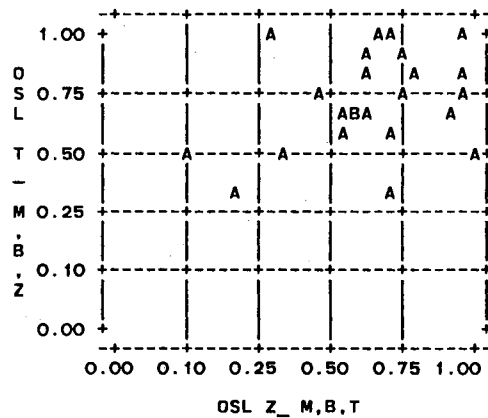
              OSL Z_ M,B,T   OSL Z_ M,B,T   OSL Z_ M,B,T
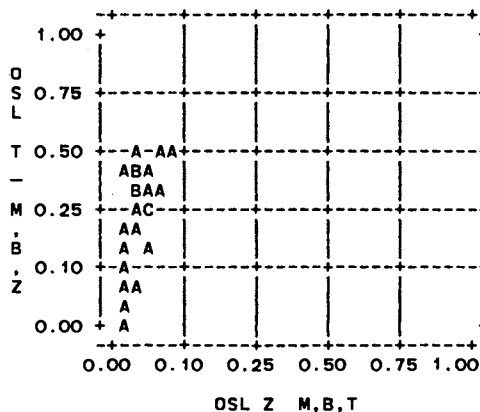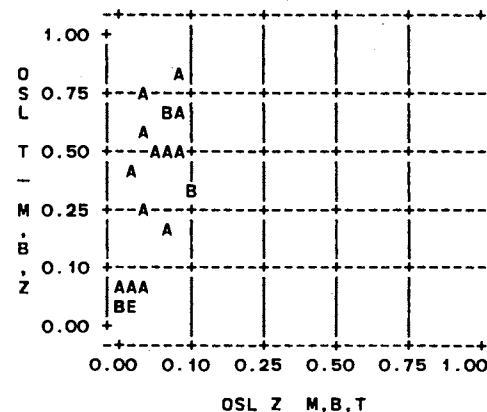
Display 4.8.8
OSL's for β vs OSL' for τ
Maximum Noncentrality Parameter for β

4.8.5 Hypothesis Tests for the Best and Worst Cases

One method to preserve the value of the stated Type I error level, α, is to first test the entire model equation, then test the two usual hypotheses of interest. When $\tau=0$ and $\beta\neq0$, one wants the model (M) and covariate (C) to be statistically significant (S), while the treatments (T) to be not statistically significant (N). Such an outcome is circled in the upper portion of Display 4.8.9. There one sees that for experimental unit error, 144 or 25% of the randomizations yield this correct outcome. For experimental unit-treatment interaction error, 47.4% yield the correct outcome; for normally distributed random error, 30.6% do so. All other randomizations err on one or more of the three hypotheses tested. All hypothesis tests use $\alpha=.10$ and the central F-distribution with the usual degrees of freedom.

Circled in the lower portion of Display 4.8.9 is the proper outcome when $\tau\neq0$ and $\beta\neq0$, this is statistical significance (S) for the model (M), the treatments (T) and covariate (C). Some 30.7% of the randomizations for the experimental unit error type yield the correct outcome. The other two error types yield the correct outcome in, respectively, 24.5% and 23.6% of their randomizations. For the error and parameter values of this simulation, one is more likely to select a randomization leading to erroneous results than one leading to the correct result, both for $\tau=0$ and $\tau\neq0$.

Counts and Percentages of Randomizations with Outcomes for
Hypotheses: Model (M), Adj. Treatments (T), Adj. Covariate (C)
of: Significant   (S) Reject Null Hypothesis and
Nonsignificant (N) Accept Null Hypothesis

TABLE 1 OF ERR_TYPE BY RESULT
CONTROLLING FOR TRT_VAL=Zero Treats

ERR_TYPE(Type of Error)   RESULT(Hypotheses Test)

Frequency Row Pct	M S T S / C S	M S T S / C N	M S T N / C S	M S T N / C N	M N T N / C S	M N T N / C N	Total
Exp. Unit	120 / 20.83	0 / 0.00	144 / 25.00	48 / 8.33	72 / 12.50	192 / 33.33	576
E.U. T.I.	54 / 9.38	1 / 0.17	273 / 47.40	55 / 9.55	94 / 16.32	99 / 17.19	576
N.D.R.E.	75 / 13.02	3 / 0.52	176 / 30.56	38 / 6.60	116 / 20.14	168 / 29.17	576
Total	249	4	593	141	282	459	1728

TABLE 2 OF ERR_TYPE BY RESULT
CONTROLLING FOR TRT_VAL=Nonzero Treats

ERR_TYPE(Type of Error)   RESULT(Hypotheses Test)

Frequency Row Pct	M S T S / C S	M S T S / C N	M S T N / C S	M S T N / C N	M N T N / C S	M N T N / C N	Total
Exp. Unit	177 / 30.73	22 / 3.82	92 / 15.97	31 / 5.38	67 / 11.63	187 / 32.47	576
E.U. T.I.	141 / 24.48	14 / 2.43	201 / 34.90	53 / 9.20	79 / 13.72	88 / 15.28	576
N.D.R.E.	136 / 23.61	10 / 1.74	121 / 21.01	36 / 6.25	110 / 19.10	163 / 28.30	576
Total	454	46	414	120	256	438	1728

Display 4.8.9
Tests of Model and Parameter Hypotheses
All Randomizations

Displays 4.8.10 and 4.8.11 detail the previous

display (Display 4.8.9) for, respectively, the smallest

value and the largest value of the noncentrality

parameter for the covariate as adjusted for the mean,

blocks, and treatments, $.5\sigma^2\lambda_{g,\beta|\mu,b,\tau}$. These two

noncentrality values are the same as in Displays 4.8.7

and 4.8.8; the conclusions are similar.

For experimental unit error and $\tau=0$, all 24

randomizations yield the correct conclusions for the

three hypotheses at both the lowest (Display 4.8.10) and

the highest (Display 4.8.11) values of the adjusted

covariate noncentrality parameter. For this error type,

and $\tau\neq0$, there are fewer correct conclusions at the

highest value than at the lowest value of the noncentral-

ity parameter, 22 vs 4 or 92% vs 17% (Displays 4.8.10 vs

4.8.11). As with Displays 4.8.7 and 4.8.8, this suggests

that the advice of (2.6.2a), (p. 41), does not apply to

the case of experimental unit error. That is, in their

upper tails, the distribution of the F-ratios over all

randomizations is not approximated by the F-distribution.

The effect of the singular normal method of computation

of the degrees of freedom, if appropriate, is an

extension.

For experimental unit-treatment interaction and

normally distributed random errors, more randomizations

yield the correct conclusion at the higher than at the

lower noncentrality parameter value, for both $\tau=0$ and

$\tau \neq 0$. For experimental unit-treatment interaction error,
no randomization yields the correct conclusions at the
lower value of the noncentrality parameter, both for $\tau =$
0 and $\tau \neq 0$. At the higher level of the noncentral- ity
parameter, 50% of the randomizations lead to correct
conclusions for $\tau = 0$ and 17% for $\tau \neq 0$. For normally
distributed random error, the improvement is from 1/24 to
9/24 for $\tau = 0$ and from 4/24 to 10/24 for $\tau \neq 0$. For
these two error types, restricting randomization to the
g-group with the largest value of the noncentrality
parameter for the adjusted covariate is more likely to
select a randomization which will lead to the correct
conclusions for these hypotheses.

Counts and Percentages of Randomizations with Outcomes for
Hypotheses: Model (M), Adj. Treatments (T), Adj. Covariate (C)
of: Significant   (S) Reject Null Hypothesis and
Nonsignificant (N) Accept Null Hypothesis

---------------- Lamda: covariates given rest=10 ----------------

TABLE 1 OF ERR_TYPE BY RESULT
CONTROLLING FOR TRT_VAL=Zero Treats

ERR_TYPE(Type of Error)   RESULT(Hypotheses Test)

Frequency Row Pct	M S T S C S	M S T S C N	M S T N C S	M S T N C N	M N T N C N	Total
Exp. Unit	0 0.00	0 0.00	24 100.00	0 0.00	0 0.00	24
E.U. T.I.	0 0.00	0 0.00	0 0.00	10 41.67	14 58.33	24
N.D.R.E.	3 12.50	0 0.00	1 4.17	7 29.17	13 54.17	24
Total	3	0	25	17	27	72

TABLE 2 OF ERR_TYPE BY RESULT
CONTROLLING FOR TRT_VAL=Nonzero Treats

ERR_TYPE(Type of Error)   RESULT(Hypotheses Test)

Frequency Row Pct	M S T S C S	M S T S C N	M S T N C S	M S T N C N	M N T N C N	Total
Exp. Unit	22 91.67	0 0.00	2 8.33	0 0.00	0 0.00	24
E.U. T.I.	0 0.00	1 4.17	0 0.00	12 50.00	11 45.83	24
N.D.R.E.	4 16.67	2 8.33	0 0.00	6 25.00	12 50.00	24
Total	26	3	2	18	23	72

Display 4.8.10
Tests of Model and Parameter Hypotheses
Minimum Noncentrality Parameter for β

Counts and Percentages of Randomizations with Outcomes for
Hypotheses: Model (M), Adj. Treatments (T), Adj. Covariate (C)
of: Significant   (S) Reject Null Hypothesis and
      Nonsignificant (N) Accept Null Hypothesis

------------------ Lamda: covariates given rest=990 ----------------

TABLE 1 OF ERR_TYPE BY RESULT
CONTROLLING FOR TRT_VAL=Zero Treats

ERR_TYPE(Type of Error)   RESULT(Hypotheses Test)

Frequency Row Pct	M S T S C S	M S T N C S	M N T N C S	M N T N C N	Total
Exp. Unit	0 0.00	24 100.00	0 0.00	0 0.00	24
E.U. T.I.	0 0.00	12 50.00	12 50.00	0 0.00	24
N.D.R.E.	5 20.83	9 37.50	8 33.33	2 8.33	24
Total	5	45	20	2	72

TABLE 2 OF ERR_TYPE BY RESULT
CONTROLLING FOR TRT_VAL=Nonzero Treats

ERR_TYPE(Type of Error)   RESULT(Hypotheses Test)

Frequency Row Pct	M S T S C S	M S T N C S	M N T N C S	M N T N C N	Total
Exp. Unit	4 16.67	20 83.33	0 0.00	0 0.00	24
E.U. T.I.	4 16.67	9 37.50	11 45.83	0 0.00	24
N.D.R.E.	10 41.67	4 16.67	8 33.33	2 8.33	24
Total	18	33	19	2	72

Display 4.8.11
Tests of Model and Parameter Hypotheses
Maximum Noncentrality Parameter for β

4.8.6   Parameter Estimates for the Best and Worst Cases

Tables 4.8.1 and 4.8.2 display, respectively, the means and standard deviations of the estimates of the model equation parameters for the g-groups with the lowest (l) and highest (h) values of the noncentrality parameter for the covariates as adjusted for the mean, blocks, and treatments, $\lambda_{s,\beta|\mu,b,\tau}$. The first column of numbers in each table shows the true value of the model equation parameter; the two treatment values are shown. For all error types and all estimates, the tables show less bias and smaller standard deviations for the g-group with the highest (h) value of this noncentrality parameter.

The second column of Table 4.8.1 is for experimental unit error. The expectation of the estimated treatment effects, $\hat{\tau}_k$, is unbiased within both g-groups. This holds for $\tau=0$ and $\tau\neq0$. Table 4.8.2 shows that the standard deviations of the estimates within the highest valued (h) g-group are much smaller than are those within the lowest valued (l) g-group, for example 4.96 vs 53.61 for $\hat{\tau}_1$. The standard deviation is equal for all estimated treatment effects, and is unaffected by the value of τ. Estimates for the other model equation parameters also show less bias at the highest valued g-group, for example $E(\hat{\beta})=11.34$ at the lowest valued (l) g-group, but equals 1.94 at the highest (h). The parameter value is $\beta=2.0$.

As per Theorem 3.7.9(4), (p. 202), $\hat{\beta}$ is constant within
the g-group for this error type; the standard deviation
shown in Table 4.8.2 is zero.

For experimental unit-treatment interaction error,
the third column of Table 4.8.1, all estimates are
biased, but much less so for the highest valued g-group.
This holds for $\tau=0$ and $\tau\neq0$, for example for $\tau_1=-6.5$, the
expectations are -9.33 and -6.54. The estimate at the
higher valued group (h) is closer to the true value.
Standard deviations again are much less at the highest
valued g-group than at the lowest valued g-group, for
example, 2.53 vs 13.6 for $\hat{\tau}_1$. As discussed in Theorem
3.7.9(4) the η_s disperse the clusters of estimates for β,
hence the within g-group standard deviation for $\hat{\beta}$ is not
zero in Table 4.8.2 for this error type. For the same
reason, neither is the standard deviation for $\hat{\beta}$ for the
normally distributed random error case (column four).

The standard deviation for $\hat{\beta}$ is due only to the
dispersion effects of η_s and n_s. The standard deviation
for $\hat{\tau}$ is due to these effects plus that of the change in
randomization. Should the experimental units be
reuseable, one might estimate the variance of the η_s
and/or the n_s by conducting all experiments within the
same g-group of randomizations.

The case of normally distributed random error is similar to the case of experimental unit-treatment interaction error. One exception is that the biases for $\hat{\tau}_2$ and $\hat{\tau}_3$ are larger than for other levels of $\hat{\tau}$.

The cases with unequal within-block error variances resemble the data of Tables 4.8.1 - 4.8.2 and are omitted. Their biases and standard deviations are slightly larger.

TABLE 4.8.1

RANDOMIZATION EXPECTED VALUE OF PARAMETER ESTIMATES
FOR G-GROUPS WITH EXTREME VALUES OF $\lambda_{s,\beta|\mu,b,\tau}$
EQUAL WITHIN-BLOCK ERROR VARIANCES

Parameter: Name Value (l=min., h=max.)		Three Types of Error: Experimental Unit	Unit-Treat. Interaction	Normally Distributed
μ	10.0	l -80.13	12.42	3.87
		h 10.54	10.06	9.68
b_1	-1.5	l 23.08	-2.16	0.17
		h -1.65	-1.52	-1.41
b_2	1.5	l -23.08	2.16	-0.17
		h 1.65	1.52	1.41
τ_1	0/-6.5	l 0/-6.5	-2.83/-9.33	-0.93/-7.43
		h 0/-6.5	-0.04/-6.54	0.01/-6.40
τ_2	0/-3.5	l 0/-3.5	4.04/ 0.55	-0.44/-3.94
		h 0/-3.5	-0.04/-3.54	0.63/-2.87
τ_3	0/ 2.5	l 0/ 2.5	-7.19/-4.69	6.51/ 9.01
		h 0/ 2.5	0.04/ 2.46	-0.67/ 1.83
τ_4	0/ 7.5	l 0/ 7.5	5.98/13.48	-5.14/ 2.36
		h 0/ 7.5	0.12/ 7.62	0.06/ 7.44
β	2.0	l 11.34	1.75	2.64
		h 1.94	2.00	2.03

TABLE 4.8.2

RANDOMIZATION STANDARD DEVIATION OF PARAMETER ESTIMATES
FOR G-GROUPS WITH EXTREME VALUES OF $\lambda_{s,\beta|\mu,b,\tau}$
EQUAL WITHIN-BLOCK ERROR VARIANCES

| | | | Three Types of Error: | | |
Parameter: Name Value (l=min., h=max.)			Experimental Unit	Unit-Treat. Interaction	Normally Distributed
μ	10.0	l	0.0	21.5	35.60
		h	0.0	3.9	4.17
b_1	-1.5	l	0.0	6.96	9.71
		h	0.0	2.90	1.14
b_2	1.5	l	as in	as in	as in
		h	b_1	b_1	b_1
τ_1	0/-6.5	l	53.61	13.6	23.21
		h	4.96	2.53	3.72
τ_2	0/-3.5	l	as in	13.6	22.29
		h	τ_1	3.49	3.81
τ_3	0/ 2.5	l	as in	10.1	18.48
		h	τ_1	1.87	3.53
τ_4	0/ 7.5	l	as in	11.2	19.48
		h	τ_1	2.16	4.64
β	2.0	l	0.0	2.30	3.40
		h	0.0	0.45	0.43

Tables 4.8.1 and 4.8.2 should be compared with

Tables 4.4.1 and 4.4.2, (p. 295-262), which average over

all randomizations with equal selection probability and

with Tables 4.7.1 and 4.7.2, (p. 326-327), which also

average over all randomizations, but with weighted

selection probabilities. All of these tables display the

equal within-block error cases. The comparison shows an
advantage by selecting the randomization used to conduct
the actual experiment from among those with the largest
value of the noncentrality parameter for the adjusted
covariate. For one covariate, these randomizations are
the most balanced on the covariate. Model equation
parameter estimates are improved for all three error
types. Conclusions in hypotheses tests are improved for
cases with experimental unit-treatment interaction or
normally distributed random errors

The reader interested in further developing these
topics will find that Appendix E, (p. 389), lists
suggested refinements of this work and extensions.
Affixed to the inside back cover is a 3.5" hard diskette
containing the SAS programs listed in Appendix F. The
files are in ASCII and contain the JCL statements used to
submit them. An IBM mainframe cartridge containing SAS
transportable versions of the data sets as well as the
SAS programs of Appendix F is available from the author.

END OF TEXT

BIBLIOGRAPHY and AUTHOR INDEX

Literature on the following topics is cited at the indicated pages:

Below each of the following bibliographical entries are pairs of page numbers:
/ page(s) of cited work : page(s) of citation /   .
For example, a passage on page 221 of Anderson and Bancroft's article will be found cited on page 118 of this work. That a general citation to the entire work is found on, say, page 25 herein, is indicated as / --:25 /. Multiple citations of a single work are ordered by that author's page numbers. Citations marked A.1, B, etc. are in the Appendices. Some titles without citations are included; they may be of interest.

Adrian .(1808)._____. The Analyst, or Mathematical
    Museum, Vol. 1, p. 83-87, 93-109.
      / --:25 /

Anderson, R. L., and T. A. Bancroft. (1952). Statitical
    Theory in Research. New York: McGraw-Hill. p.220-221.
      / 221:118 /

Atiqullah, M. (1962A). "The Estimation of Residual
    Variance in Quadratically Balanced Least-Squares
    Problems and the Robustness of the F-test".
    Biometrika, Vol. 49, p. 83-91.
      / --:226, 231 /

Atiqullah, M. (1962B). "On the Effect of Non-normality
    on the Estimation of Components of Variance".
    Journal of the Royal Statistical Society, Series B,
    Vol. 24, p. 140-147.
      / --:228 /

Atiqullah, M. (1963). "On the Randomization Distribution and Power of the Variance Ratio Test". Journal of the Royal Statistical Society, Series B, Vol. 25, p. 334-347.
/ --:224, 225, 231 / 338:226 /

Atiqullah, M. (1964). "The Robustness of the Covariance Analysis of a One-way Classification". Biometrika, Vol. 51, p. 365-372.
/ --:12 / 365, 368:228 /

Bailey, R. A. (1982). "Randomization". In Recent Trends in Statistics, S. Heiler, Ed. Gottingen: Vandenhoeck and Ruprecht. p. 9-26.
/ --:118 /

Bailey, R. A. (1985). "Restricted Randomization Versus Blocking". International Statistical Review, Vol. 53, p. 171-183.
/ --:119 /

Bailey, R. A. (1987). "Restricted Randomization: A Practical Example". Journal of the American Statistical Association, Vol. 82, p. 712-719.
/ --:114, 120 / 712:114, 240 /

Bartlett, M. S. (1935). "The Effect of Non-Normality on the t distribtuion". Proceedings of the Cambridge Philosophical Society, Vol. 31, p. 223-231.
/ --:125, 224 /

Basu, D. (1980). "Randomization Analysis of Experimental Data: The Fisher Randomization Test". The Journal of the American Statistical Association, Vol. 75, p. 575-581, with comment on p. 582-595.
/ --:118, 120 /

Ben-Israel, Adi and Thomas N. E. Greville. (1980). Generalized Inverses: Theory and Application. New York: Krieger, Huntington. Chapters 1 - 3.
/ 7-22:83 / Ch.1 - Ch.3:90 /

Bingham, Christopher and Stephen E. Fienberg. (1982). "Textbook Analysis of Covariance -- Is It Correct?". Biometrics, Vol. 38, p. 747-753.
/ --:21, E.IV.1.b /

Boullion, Thomas L. and Patrick L. Odell.   (1971).
    Generalized Inverse Matrices.   New York:
    Wiley-Interscience.
        / --:90 /

Boullion, Thomas L., John W. Seaman, Jr. and Dean M.
    Young. (1990). "Moments of Discrete Probability
    Distributions Derived Using a Differential Operator".
    Presented at the American Statistical Association's
    Annual Meeting in New Orleans, 1990.
        / --:232 /

Burton, David M. (1971). _Abstract and Linear Algebra_.
    Reading, Mass.: Addison-Wesley.
        / 48-53, 69-74:188 /

Campbell, S. L. (Ed.).   (1982).   _Recent Applications of
    Generalized Inverses_.   London: Pitman Advanced
    Publishing Program.
        / --:91 /

Cassel, Claes-Magnus, Carl-Erik Särndal and Jan Håkan
    Wretman. (1977).   _Foundations of Inference in Survey
    Sampling_. New York: John Wiley & Sons. p. 1-85.
        / --:64 / 1-31:141 / 10:143 / 11-12, 68:111 /
        / 72,85:165 /

Center, Bruce Aruthor. (1982). "An Investigation of the
    Behavior of Selected Variance Ratio Statistics under
    Randomization for the Analysis of Covariance".
    unpublished Ph.D. dissertation, University of
    Minnesota. University Microfilms Order No. DA8318049.
        / --:231 /

Chow, Yuan Shih and Henry Teicher. (1978). _Probability
    Theory, Independence, Interchangeability,
    Martingales_. New York: Springer-Verlag.
        / 33, 223-226, 309-311:165 /

Cifarelli, D. M. (1979).   "Impostazione bayesiana di un
    problema di analisi della varianza con approccio non
    parametrico". _Ist. Mat. Finanz._ Univ. Torino. III 17.
        / --:165 /

Cifarelli, D. M. and E. Regazzini. (1981A). "Some
    Considerations about Mathematical Statistics
    Teaching Methodology Suggested by the Concept of
    Exchangeability". In _Exchangeability in Probability
    and Statistics_, G. Koch and F. Spizzichino, Eds.,
    New York: North-Holland Publishing Co. p. 185-206.
        / --:165 /

Cifarelli, D. M., P. Muliere, and M. Scarsini. (1981B).
"IL modello lineare nell'approccio bayesiano non
parametrico". Ist. Mat. "G. Castelnuovo". Univ. Roma.
/ --:165 /

Cline, Randall E. (1979), Elements of the Theory of
Generalized Inverses for Matrices. Newton,
Mass.: EDC/UMAP.
/ --:90 /

Cochran, William G. (1957). "Analysis of Covariance: Its
Nature and Uses". Biometrics, Vol. 13, p. 261-281.
/ --:31 / 262-267:10,11 / 272:207 /

Cochran, William G. and Gertrude M. Cox. (1957).
Experimental Designs. New York: John Wiley & Sons.
p. 6-9.
/ 6-9:121 /

Cook, M. B. (Miss). (1951A). "Bi-Variate k-Statistics and
Cumulants of their Joint Sampling Distribution".
Biometrika, Vol. 38, p. 179-195. See Cook (1951B).

Cook, M. B. (Miss). (1951B). "Two Applications of
Bivariate k-Statistics". Biometrika, Vol. 38,
p. 368-376. See Cook (1951A).
/ --:E.V.1, G /

Cox, David Roxbee. (1956). "A Note on Weighted Randomiza-
tion". The Annals of Mathematical Statistics,
Vol. 27, p. 1144-1151.
/ --:320 / 1146:321 / 1147:227 / 1148:241, 322/
/ 1150:322 /

Cox, David Roxbee. (1957). "The Use of a Concomitant
Variable in Selecting an Experimental Design".
Biometrika, Vol. 44, p. 150-158.
/ --:320 / Methods III, IV, and V:116 / 153:321 /

Cox, David Roxbee. (1958). Planning of Experiments.
New York: Wiley and Sons. p. 14-17, 70-90.
/ 14-17:134 / 74-85:118 / 77-81:119 / 85-89:119 /

Cox, David Roxbee. (1982). "Randomization and
     Concomitant Variables in the Design of Experiments".
     In Kallianpur, G., P. R. Krishnaiah, J. K. Ghosh,
     Eds., <u>Statistics</u> <u>and</u> <u>Probability: Essays</u> <u>in</u> <u>Honor</u> <u>of</u>
     <u>C. R. Rao</u>. New York: North-Holland Publishing
     Company. p. 197-202.
       / --:320 / 198:321 / 200:180, 322, 328 /

Cox, David Roxbee and D. V. Hinkley. (1974). <u>Theoretical</u>
     <u>Statistics</u>. London: Chapman and Hall.
       / 180 and 196:168 /

Cox, David Roxbee and Peter McCullagh. (1982). "Some
     Aspects of Analysis of Covariance". <u>Biometrics</u>.
     Vol. 38, p. 541-561.
       / --:13 / 541-542:7, 10 / 543:15 / 543-544:27 /

de Finetti, Bruno. (1980). "Probability: Beware
     Falsifications!". In <u>Studies</u> <u>in</u> <u>Subjective</u>
     <u>Probability</u>, Henry E. Kyburg Jr. and Howard E.
     Smolker, Eds. Huntington, N. Y.: Robert E. Drieger
     Publishing Co., Huntington, N.Y. p. 193-224.
     Original article published in 1977.
       / 211:165 /

de Finetti, Bruno. (1974-1975). <u>Theory</u> <u>of</u> <u>Probability</u>,
     <u>Volumn</u> <u>Two</u>. New York: Wiley and Sons. Translated
     from *Teroia delle Probabilità*, Einaudi, Torino, 1970.
       / Vol.2 215:165 /

de Finetti, Bruno. (1980). "Foresight: Its Logical Laws,
     Its Subjective Sources". In <u>Studies</u> <u>in</u> <u>Subjective</u>
     <u>Probability</u>, Henry E. Kyburg Jr. and Howard E.
     Smolker, Eds. Huntington, N.Y.: Robert E. Drieger
     Publishing Co. p. 53-118, also p. 15-18. A
     translation of "La prévision: Ses lois logiques, ses
     sources subjectives" from the <u>Annales</u> <u>de</u> <u>i'Institut</u>
     <u>Henri</u> <u>Poincaré</u>, Vol. 7 (1937).
       / 15-17, 81-83, 99-101:165 /

De Lury, D. B. (1948). "The Analysis of Covariance".
     <u>Biometrics</u>, Vol. 4, p. 153-170.
       / --:11 /
Dolby, G. R. (1982). "The Role of Statistics in the
     Methodology of the Life Sciences". <u>Biometrics</u>, Vol.
     38. p. 1069-1083.
       / 1083:vii /

Duncan, Acheson J. (1957). "Charts of the 10% and 50% Points of the Operating Characteristic Curves for Fixed Effects Analysis of Variance F Tests, $\alpha=0.01$ and 0.05". The Journal of the American Statistical Society, Vol. 52, p. 345-349.
    / --:A.3 /

Eden, T. and F. Yates. (1933). "On the Validity of Fisher's Z-test when Applied to an Actual Sample of Non-normal Data". The Journal of Agricultural Science (Cambridge), Vol. 23, p. 6-16.
    / --:231 /

Edgington, Eugene S. (1987), Randomization Tests, 2nd Edition. New York: Marcel Dekker, Inc. (Volumn 77 in the series, Statistics: Textbooks and Monographs.)
    / --:166 /

Fedorov, V. V. (1972), Theory of Optimal Experiments. New York: Academic Press. Translated into English; originally published in 1969.
    / --:120 /

Feldt, Leonard S. and Moharram W. Manhmoud. (1958). "Power Function Charts for Specifiying Numbers of Observations in Analyses of Variance of Fixed Effects". Annals of Mathematical Statistics, Vol. 29, p. 871-887.
    / --:A.3 /

Finney, D. J. (1946). "Standard Errors of Yields Adjusted for Regression on an Independent Measurement". Biometrics, Vol. 2, p. 53-55.
    / --:116 / 54:321 /

Finney, D. J. (1957). "Stratification, Balance, and Covariance". Biometrics, Vol. 13, p. 373-386.
    / --:116 /

Fisher, Ronald Aylmer. (1928). The Design of Experiments. Edinburgh: Oliver and Boyd.
    / 671:A.4 /

Folks, J. Leroy. (1984). "Use of Randomization in Experimental Research", in Experimental Design, Statistical Models, and Genetic Statistics, Klaus Hinkelmann, Ed. New York: Marcel Dekker, Inc. p. 17-32.
    / --:118 /

Fox, M. (1956). "Charts of the Power of the F-Test".
    Annals of Mathematical Statistics, Vol. 27, p.
    484-497.
     / --:A.3 /

Gabriel, K. R., and W. J. Hall. (1983). "Rerandomization
    Inference on Regression and Shift Effects:
    Computationally Feasible Methods". Journal of the
    American Statistical Society, Vol. 78, p. 827-836.
     / --:168, 174 /

Galilei, Galileo (1953) 1953), Dialogue Concerning the
    Two Chief World Systems, Stillman Drake, translator
    of 1632 original. Berkely: University of California
    Press. p. 281-316 (The Third Day) and p. 483-487.
     / the above:92-96 /

Gauss, Karl Friedrich (1857) Theory of the Motion of the
    Heavenly Bodies Moving about the Sun in Conic
    Sections, (Theoria Motus), Charles Henry Davis,
    translator of 1809 original. New York: Dover.
    p. 249-260.
     / 249-260:92-98 /

Gayen, A. K. (1950). "The Distribution of the Variance
    Ratio in Random Samples of any Size Drawn from
    Non-Normal Universes". Biometrika, Vol. 37, p.
    236-251.
     / --:231 /

Good, Phillip I. (1979). "Detection of a Treatment
    Effect When Not All Experimental Subjects Will
    Respond to Treatment". Biometrics, Vol.35, p.483-489.
    See Zhou, et al. (1989).

Graybill, Franklin A. (1976). Theory and Application of
    the Linear Model. Boston: Duxbury Press.
     / 6:26 / 22:26 / 283-297:12-13 / 493-498:76 /

Graybill, Franklin A. (1983). Matrices with Applications
    in Statistics, 2nd Edition. Belmont, CA.: Wadsworth.
     / --:13 / 87,89:89 / 134: 75:78 / 135:84 / Ch. 6:90/

Greenberg, Bernard G. (1953). "The Use of Analysis of
    Covariance and Balancing in Analytical Surveys."
    American Journal of Public Health, Vol.43, p.692-699.
     / --:116 / 698:321 /

Greville, Thomas N. E. (1966). "Note on the Generalized
    Inverse of a Matrix Product". SIAM Review,
    Vol. 8, p. 518-521. Errata, Vol. 9, p. 249.
      / --:89 /

Godambe, V. P. (1966). "A New Approach to Sampling from
    Finite Populations I, II". Journal of the Royal
    Statistical Socitey, Series B, Vol. 28, p. 310-328.
      / --:120 /

Godambe, V.P. (1975). "A Reply to My Critics". Sankhyā C,
    Vol. 37, p. 53-76.
      / --:120 /

Godambe, V. P. and M. E. Thompson (1971), "Bayes,
    Fiducial and Frequency Aspects of Statistical
    Inference in Regression Analysis in Survey
    Sampling". Journal of the Royal Statistical Society,
    Series B, Vol. 33, p. 361-390.
      / --:120 /

Hartley, H. O. and J. N. K. Rao. (1971). "Foundations of
    Survey Sampling (a Don Quixote Tragedy)". The
    American Statistician, Vol 25, p. 21-7.
      / --:120 /

Harvey, James R. (1972). "On Expressing Moments in Terms
    of Cumulants and Vice Versa". The American
    Statistician, Vol. 26, p. 38-39.
      / --:G /

Henderson, Charles R., Jr. (1982). "Analysis of Covar-
    iance in the Mixed Model: Higher-Level, Nonhomogen-
    eous, and Random Regressions." Biometrics, Vol. 38,
    p. 623-640.
      / 623-624:13, 15, 19 /

Hoaglin, D. C. and R. E. Welsch. (1978). "The Hat Matrix
    in Regression and ANOVA". The American Statistician,
    Vol. 32, p. 17-22.
      / --:13 /

Hocking, R. R. (1982). "Discussion on a Paper by D. R.
    Cox and P. McCullagh". Biometrics,Vol.38, p. 559-561.
      / --:13 / 560:7 /

Holland, Paul W. (1986). "Statistics and Causal
    Inference". The Journal of the American Statistical
    Association, Vol. 81, p. 945-960, with discussion.
      / 945:8 /

Hooke, Robert. (1958). "Use of Randomization in the
    Investigation of Unknown Functions". The Journal of
    the American Statistical Association, Vol. 53,
    p. 176-186.
      / --:120 /

Hooper, Peter M. (1989). "Experimental Randomization and
    the Validity of Normal-Theory Inference". Journal of
    the American Statistical Association, Vol.84,
    p. 576-586.
      / --:230, 232 / 578-579:21 /

Hsu, P. L.   (1938). "On the Best Unbiased Quadratic
    Estimate of the Variance". Statistical Research
    Memoires, Vol. 2, p. 91-104.
      / --:12, 231 /

Hsu, P. L. (1941). "Analysis of Variance from the Power
    Function Standpoint". Biometrika, Vol. 32, p. 62-69.
      / --:A.3 /

Irving, J. O. and Maurice George Kendall. (1943-1945).
    "Sampling Moments of Moments for a Finite
    Population". Annals of Eugenics, Vol. 12, p. 138-142.
      / --:232 /

Jefferson, Craig. (1990). Personal communication.
      / --:50, (used on 53, 55), C /

Jonrup, H. and B. Rennermalm. (1976)."Regression Analysis
    in Samples from Finite Populations". Scandanavian
    Journal of Statistics, Vol. 3, p. 33-37.
      / --:157 /

Kempthorne, Oscar. (1952). The Design and Analysis of
    Experiments. New York: John Wiley & Sons. Chapter 8.
      / 97-103:31 / 98-103:11 / 128-129:181,182 /
      /130-131:171 / 132-134:139 / 137:126 /
      / 145-146:124, 125, 126, 128 / 159:138 /230-231:171/
      /232:107/

Kempthorne, Oscar. (1955). "The Randomization Theory of
    Experimental Inference". Journal of the American
    Statistical Society, Vol. 50, p. 946-967.
    Corrections in Vol. 51 (1956), p. 651.
     / 949:120 / 952:133, 134 / 956:135 /

Kempthorne, Oscar. (1975). "Inference from Experiments
    and Randomization". In A Survey of Statistical
    Design and Linear Models, Jagdish N. Srivastava,
    Ed. New York:North-Holland. p. 303-332.
     / 324:127 / 324-325:139 / 326:121 /

Kempthorne, Oscar, and T. E. Doerfler. (1969). "The
    Behaviour of Some Significance Tests Under
    Experimental Randomization". Biometrika, Vol. 56,
    p. 231-248.
     / 232:107, 110 / 231-235:164 / 234-235:168 /

Kempthorne, Oscar, G. Zyskind, S. Addelman, T. N.
    Throckmorton, and R. F. White. (1967).
    "Analysis of Variance Proceedures". Aerospace
    Research Laboratories, Technical Report 149.
    Wright-Patterson Air Force Base, Ohio.
    See also Zuskind, Kempthorne, et al. (1960, 1964,
    1968).
     / --:232 /

Kiefer, J. (1959). "On the Nonrandomized Optiomality and
    Randomized Nonoptimality of Symmetrical Designs".
    The Annals of Mathematical Statistics, Vol. 29,
    p. 675-699.
     / --:120 /

Kotlarsky, Ignacy. (1990). Personal communication.
     / --:142 /
Kotz, Samuel and Norman L. Johnson. (1985-1986).
    Encyclopedia of Statistical Sciences, Volumns 6 and
    7. New York: Wiley and Sons.
    Volumn 6 (1985):/ 277:A.2 /
    Volumn 7 (1986):/ 524-530:120,167-168 /

Lane, David A. (1980). "Comment" (on Basu 1980). The
    Journal of the American Statistical Society,
    Vol. 75, p. 587-589.
     / 589:119 /

Legendre, Adrien-Marie. (1806). New Methods for the
    Determination of the Orbits of Comets, (Nouvelles
    méthodes pour la détermination des orbites des
    comètes), Paris.
    / --:25 /

Lindley, D. V. and Melvin R. Novick. (1981). "The Role of
    Exchangeability in Inference". The Annals of
    Statistics, Vol. 9, p. 45-58.
    / --:165 /

Lucas, Henry L. (1950). "Design and Analysis of Feeding
    Experiments with Milking Dairy Cattle". Raleigh, N.C.:
    Institute of Statistics Mimeograph Series No. 18.
    / --:116, 321 /

Maistrov, L. E. 1974. Probability Theory A Historical
    Sketch, Samuel Kotz, translator. New York: Academic
    Press. p. 30-34 and 148-157.
    / 30-34, 148-157:93 / 152, 155:94 /

Mandel, J. (1961). "Non-additivity in Two-Way Analysis
    of Variance". The Journal of the American
    Statistical Association, Vol. 56, p. 878-888.
    / --:127 /

Mann, H. B. (1949). Analysis and Design of Experiments.
    New York: Dover Publications, Inc. p. 74-75.
    / 74-74:A.3 /

McCullagh, Peter. (1987). Tensor Methods in Statistics.
    London: Chapman and Hall.
    / --:232 /

Mood, Alexander McFarland, Franklin A. Graybill, and
    Duane C. Boas. (1974). Introduction to the Theory of
    Statistics, 3rd Edition. New York: McGraw-Hill.
    / 202:279 / 534:A.2 /

Morrison, Donald F. (1983), Applied Linear Statistical
    Methods. Englewood Cliffs, N.J.: Prentice-Hall, Inc.
    p. 120-131.
    / 122-124:45 / 122-131:13 / 127-131:40, A.3 /

Nashed, M. Zuhair, Editor. (1976). Generalized Inverses
     and Applications. New York: Academic Press.
      / --:91 /

Neter, John, William Wasserman, and Michael H. Kutner.
     (1990). Applied Linear Statitical Models, 3rd
     Edition. Homewood, Illinois: Irwin.
      / 861-906:11 /

Neyman, J. with the cooperation of K. Iwaszkiewicy and
     St. Kotodziejczyk. (1935). "Statistical Problems in
     Agricultural Experimentation". Journal of the Royal
     Statistical Society, Series B, Vol. 2, p. 107-154.
      / 110:124 / 111:127, 139 / 114:124 / 145:124, 127 /
      / 146:131 / 172-177:139 /

Ostle, Bernard (1954), Statistics in Research. Ames,
     Iowa: The Iowa State College Press. p. 288-293, 385,
     392-397, and 410-465.
      / 410-465:11 / 394:33 /

Ott, Leyman. (1988). An Introduction to Statistical
     Methods and Data Analysis, Third Edition. Boston:
     PWS-Kent Publishing Co.
      / 676-681:11 / (Only in First Edition,1977) 524:81 /

Pearson, Egon Sharpe and H. O. Hartley. (1976).
     Biometrika Tables for Statisticians, Vol. II.
     Biometrika Trust. p. 66-77.
      / 64, 66-77:A.3 /

Pitman, E. J. G. (1937). "Significance Tests Which May
     Be Applied to Samples from Any Populations, III The
     Analysis of Variance Test". Biometrika, Vol. 29,
     p. 322-335.
      / --:125, 224, 227 / 322-323:125 / 323:206 /
      / 331-335:225 /

Plackett, R. L. (1960). Principles of Regression
     Analysis. Oxford: Clarendon Press. p. 40-41 and
     152-159.
      / 40-41:231 / 152-157:125 / 155-156:224 /

Pope, P. T. and J. T. Webster. (1972). "The Use of an
     F-Statistic in Stepwise Regression Procedures".
     Technometrics, Vol. 14, p. 327-340.
      / 330-331:45 /

Pringle R. M. and A. A. Rayner. (1971). <u>Generalized</u>
<u>Inverse</u> <u>Matrices</u> <u>with</u> <u>Applications</u> <u>to</u> <u>Statistics</u>.
New York: Hafner Publishing Co.
/ --:90 / 88-90:76 /

Rao, Calyampudi Radhakrishna. (1952). "Some
Theorems on Minimum Variance Estimation". <u>Sankhya</u>,
Vol. 12, p. 27-42.
/ --:231 /

Rao, Calyampudi Radhakrishna. (1973). <u>Linear</u> <u>Statistical</u>
<u>Inference</u> <u>and</u> <u>Its</u> <u>Applications</u>, <u>2nd</u> <u>Edition</u>. New
York: Wiley and Sons. p. 54.
/ 54:217 / 161:100 / 122:229 / 288-294:13 /

Rao, Calyampudi Radhakrishna. (1975). "Theory of
Estimation of Parameters in the General
Gauss-Markoff Model". In <u>A</u> <u>Survey</u> <u>of</u> <u>Statistical</u>
<u>Designs</u> <u>and</u> <u>Linear</u> <u>Models</u>, J. N. Srivastava, Ed.,
New York: North-Holland Co. p. 475-488.
/ 475-487:13 /

Rao, Calyampudi Radhakrishna and Sujit Kumar Mitra.
(1971). <u>Generalized</u> <u>Inverse</u> <u>of</u> <u>Matrices</u> <u>and</u> <u>its</u>
<u>Applications</u>. New York: John Wiley and Sons, Inc.
/ --:91 /

Reisch, Joan S., and J. T. Webster. (1969). "The Power of
a Test in Covariance Analysis". <u>Biometrics</u>, Vol. 25,
p. 701-714.
/ 705:116 /

Richards, Winston A. (1980). "The Randomization Analysis
of Covariance and Change-Over Designs". Unpublished
dissertation. Ames, Iowa: Iowa State Univeristy.
See Stewart, Margaret Hill (1980).
/ --:238 /

Robinson, J. (1973A). "The Large-Sample Power of
Permutation Tests for Randomization Models". <u>The</u>
<u>Annals</u> <u>of</u> <u>Statistics</u>, Vol. 1, p. 291-296.
/ --:226 /

Robinson, J. (1973B). "The Analysis of Covariance under a
Randomization Model". <u>Journal</u> <u>of</u> <u>the</u> <u>Royal</u>
<u>Statistical</u> <u>Society</u>, <u>Series</u> <u>B</u>, Vol. 35, p. 368-376.
/ 368:227, 229 / 372:230 /

Roux, C. Z. (1982). "The Analysis of Covariance in
     Randomization Theory". South African Statistical
     Journal, Vol. 16, p. 1-23.
        / --:125, E.VI.8 / 4:137 /

Roux, C. Z. (1984). "Treatment x Unit Interactions in the
     Completely Randomized and Randomized Block Designs".
     In Experimental Design, Statistical Models, and
     Genetic Statistics, Essays in Honor of Oscar
     Kempthorne, Klaus Hinkelmann, Ed. Marcel Dekker,
     Inc. p. 141-153.
        / --:128, 132, E.VI.8 / 144:133 /

Royall, R. M. and J Herson. (1973). "Robust Estimation in
     Finite Populations I". The Journal of the American
     Statistical Society, Vol. 68, p. 880-889.
        / --:120 /

Royden, H. L. (1968). Real Analysis, 2nd Edition. New
     York: MacMillan Publishing Co., Inc. p. 270.
        / 270:158 /

Rubin, Donald B. (1978). "Bayesian Inference for Causal
     Effects: The Role of Randomization". Annals of
     Statistics, Vol. 6, p. 34-58.
        / --:118 /

Sarkar, Sahudeb. (1990). Personal communication.
        / --:142 /

SAS Institute (1990), SAS Language: Reference, Version 6,
     First Edition. Cary, N.C.: SAS Institute, Inc. SAS
     Function PROBF, p. 579.
        / 578:A.3 /

Scheffe, Henry. (1959). The Analysis of Variance.
     Toronto: Wiley and Sons. p. 42-45, 104-106, 291-329
     (Ch.9), 446-455.
        / --:12 / 42, 446-454:A.3 /

Searle, Shayle R. (1971). Linear Models. New York: John
     Wiley & Sons. p. 340-361.
        / --:12 / 20:17, 47 / 27:46 / 36:40, 43, 44 /
        / 37:44 / 43:A.1 / 54-64:40 / 58:264 / 66-72:18 /
        / 69:316-318, E.I.1 / 80, 169:17 / 182:26 /
        / 209-220:20 / 218:75 / 221-224:13 / 240-361:12 /
        / 246-249:B / 342-343:27 / 343-344:45 / 340-361:15/
        / 351-352:33 / Ch. 8:90 /

Searle, Shayle R. (1982). <u>Matrix Algebra Useful for
    Statistics</u>. New York: John Wiley & Sons, Inc.
    Chapter 8, p. 212-226.
    / --:13 / 217:78 / 221-222:84 / 226:86 / Ch. 8:90 /

Searle, Shayle R. (1987). <u>Linear Models for Unbalanced
    Data</u>. New York: Wiley. p. 10-13, 242-243, 282-288,
    416-456 (Ch. 11), 457-483 (Ch. 12).
    / 28-29:B / 221:78/ 233-236:41 / 242:12 /
    / 287-288:78 / 416-456:12 / 438-441:23 /

Seber, George Author Rederick. (1977). <u>Linear Regression
    Analysis</u>. New York: MacMillan Publishing Co., Inc.
    p. 13-17, 21, 414, 52-54, 151-155, 279-290.
    /Ch.10 :12 /

Seber, George Author Rederick. (1980). <u>The Linear
    Hypothesis: A General Theory</u>. New York: MacMillan
    Publishing Co., Inc. p. 5-9, 24-29, 34-35, 38-39,
    61-65.
    / 5:A.3 / 9:A.4 / 35:41, A.4 / 61-65:12 /

Sheynin, O. B. (1966). "Origin of the Theory of Errors".
    <u>Nature</u>, Vol. 211, p. 1103-1104.
    / --:93 /

Shinozaki, Nobuo, and Masaaki Sibuya. (1974). "The
    Reverse Order Law $(AB)^-=B^-A^-$" [more precisely,
    $(AB)^+=B_{mr}^- A_{lr}^-$]. <u>Linear Algebra and its Applications</u>,
    Vol. 9, p.29-40.
    / 34:84, 89 /

Snedecor, George W. and William G. Cochran. (1980).
    <u>Statistical Methods, 7th Edition</u>. Ames, Iowa: Iowa
    State University Press, Ames, Iowa.
    7th Edition (1980): / 59-63:290 / 365-392:11 /
    6th Edition (1967): / 15:8 / 419-446:11 /

Steel, Robert G. D. and James H. Torrie. (1980).
    <u>Principles and Proceedures of Statistics, A Biometric
    Approach, Second Ed</u>. New York: McGraw-Hill Book
    Company. Chapter 17, p. 401-437.
    / 401:10 / 401-437:11 /

Stewart, Margaret Jill. (1980). "Randomization Analysis
    of Replicated Randomized Complete Block Designs".
    Unpublished dissertation, Blacksburg, Virginia:
    Virginia Polytechnic Institute and State University.
    See Richards, Winston (1980).

Tang, P. C. (1938). "The Power Function of the Analysis
    of Variance Tests with Tables and Illustrations of
    Their Use". Statistical Research Memoirs.
    University of London: Department of Statistics. Vol.
    II, p. 126-157.
     / --:A.3 /

Thornett, M. L. (1982). "The Role of Randomization in
    Model-Based Inference". Australian Journal of
    Statistics, Vol. 24, p. 137-145.
     / 138-140:21, 119 /

Tiku, M. L. (1967). "Tables of the Power of the F-test".
    The Journal of the American Statistical Association,
    Vol. 62, p. 525-534.
     / --:A.3 /

Verrill, Steve. (1993). "Predictor Sort Sampling, Tight
    T's and the Analysis of Covariance". The Journal of
    the American Statistical Association, Vol. 88,
    p. 119-124.
     / --:116 / 119:116 /

Warner, S. L. (1965). "Randomized Response: A Survey
    Technique for Eliminating Evasive Answer Bias".
    The Journal of the American Statistical Institute,
    Vol. 60, p. 66-69.
     / --:104 /

Weisberg, Sanford (1980), Applied Linear Regression.
    Wiley and Sons, New York, p. 88-89.
     / 89:44 /

Welch, B. L. (1937). "On the z-Test in Randomized Blocks
    and Latin Squares". Biometrika, Vol. 29, p. 21-52.
     / --:125, 224, 227 / 22-23:139 / 47-48:164, 225 /

Welch, William. J. (1990). "Construction of Permutation
    Tests." The Journal of the American Statistical
    Association, Vol. 85, p. 693-698.
     / 693:168 /

White, Robert F. (1975). "Randomization and the Analysis
    of Variance". Biometrics, Vol. 31, p. 555-571.
     / --:149 /

Wibker, E. A., R. B. Howe, and J. D. Gilbert. (1979).
    "Explicit Solutions to the Reverse Order Law
    $(AB)^+ = B^-_{mr} A^-_{lr}$". Linear Algebra and its Applications,
    Vol. 25, p. 107-114.
     / 112:84,85 / 109:86 /

Wildt, Albert R. and Olli Ahtola. (1978). Analysis of
    Covariance. Beverly Hills, Calif. Sage Publications.
    p. 58-69.
     / 58-69:11 /

Wilk, M. B. (1955). "The Randomization Analysis of a
    Generalized Randomized Block Design". Biometrika,
    Vol. 42, p. 70-79.
     / 70,71:124, 125, 126 / 72:128, 133, 135 /
     / 78-79:169 /

Wilk, M. B. (1957). "Non-Additives in a Latin Square
    Design". The Journal of the American Statistical
    Society, Vol. 52, p. 218-236.
     / 202:8 / 224:124 /

Wilk, M. B. and Oscar Kempthorne. (1955). "Fixed, Mixed,
    and Random Models". The Journal of the American
    Statistical Society, Vol. 52, p. 1144-1167.
     / 1148:124 / 1149:134 / 1150:135 /

Wu, Chien-Fu. (1981A). "Iterative Construction of Nearly
    Balanced Assignments 1: Categorical Covariates".
    Technometrics, Vol. 23, p. 37-44.
     /--:116 /

Wu, Chien-Fu. (1981B). "On the Robustness and Efficiency
    of Some Randomized Designs". The Annals of
    Statistics, Vol. 9, p. 1168-1177.
     / --:120 /

Yates, F. (1975). "The Early History of Experimental
    Designs". In A Survey of Statistical Design and
    Linear Models, Jagdish N. Srivastava, Ed.,
    New York:North-Holland. p. 581-592.
     / 586:119 /

Zhou, Hong-Hao, Richard P. Koshakji, David J.
    Silberstein, Grant R. Wilkinson, and Alastair J. J.
    Wood. (1989). "Racial Differences in Drug Response".
    The New England Journal of Medicine, Vol. 320,
    p. 565-570. Related editorial on p. 588-589, and
    letters in Vol. 321, p. 257-259.
     / --:236 / See also Good, Phillip above.

Zielke, G. (1978). "*Zur historischen entwicklung von verallgemeinerten inversen Matrizen*". Wiss. Z. Martin-Luther-Univ. Halle-Wittenberg Math.-Natur. Reihe, Vol 27, p. 109-118. (Bibliography of g-inverses.)
/ --:91 /

Zuskind, G. and Oscar Kempthorne. (1960). "Treatment Errors in Comparative Experiments". Aeronautical Research Laboratory, Technical Note 59-19, Wright Patterson Air Force Base, Ohio. WADC Tech.
/ --:232 /

Zuskind, G., Oscar Kempthorne, R. F. White, E. E. Dayhoff, and T. E. Doerfler. (1964). "Research on Analysis of Variance and Related Topics". Aerospace Research Laboratories, ARL 64-193. Wright-Patterson Air Force Base, Ohio.
/ --:232 / 129:75 / 139-141:27 / 171-175:168 /

Zuskind, G., Oscar Kempthorne, F. B. Martin, E. J. Carney, and E. N. West. (1968). "Linear Models and Analysis of Variance Research Procedures". Aerospace Research Laboratories, ARL 68-0119. Wright-Patterson Air Force Base, Ohio.
See also Kempthorne, Zuskind, *et al.* (1967).
/ --:232, 318 / 11:318 /

END of BIBLIOGRAPHY

APPENDICES

APPENDIX A

DISTRIBUTIONS AND THEIR MOMENTS

A.1 The Multivariate Normal Distribution

The notation indicating that the random vector Y has the multivariate normal distribution, with mean $\Sigma\beta$, and variance-covariance matrix Ψ, is

$$Y \sim N_n(\Sigma\beta, \Psi), \text{ with}$$

$$Y \text{ n x 1}, \Sigma \text{ n x p}, \beta \text{ p x 1},$$
$$\Psi \text{ n x n and positive definite.}$$

Unless otherwise stated $\Psi = I\sigma^2$.
The density function is

$$f(x_1, x_2, x_3, \cdots , x_n \mid \Sigma\beta, \Psi)$$

$$= (2\pi)^{-1/(2n)}|\Psi|^{-1/2}\exp\{-(1/2)[Y-\Sigma\beta]\Psi^{-1}[Y-\Sigma\beta]\}.$$

The expectation and variance are

$$E_N(Y) = \Sigma\beta, \quad Var_N(Y) = \Psi.$$

The second central moment is $\Psi - (\Sigma\beta)(\Sigma\beta)'$; all other moments are zero. See Searle [1971, p. 43] for a development of distributions 1 - 3. For Ψ singular see Searle [1971 p. 66-77].

A.2   The Non-Central χ^2 Distribution

For $Y \sim N (\Sigma\beta, \Psi)$ and $A\Psi$ idempotent of rank n, $U = Y'AY$ has the non-central χ^2 distribution, with degrees of freedom n, and non-centrality parameter

$$\lambda = (1/2)(\Sigma\beta)'A (\Sigma\beta).$$

Typically, A is a projector matrix divided by σ^2, such as $\Sigma(\Sigma'\Sigma)^-\Sigma'/\sigma^2$, as in Searle. Also, $\Psi = I\sigma^2$, making $A\Psi$ idempotent. The density function is

$$f(u \mid rank(A\Psi), A, \Sigma\beta)$$

$$= \sum_{j = 0}^{\infty} \left[\frac{e^{-\lambda} \lambda^j}{j!} \frac{u^{(1/2)n+j-1}}{2^{(1/2)n+j}} \frac{e^{-(1/2)u}}{\Gamma[(1/2)n+j]} \right],$$

with $u > 0$, $n > 0$, and $\lambda \geq 0$.
When $\lambda = 0$, the summation is dropped. When $\lambda \neq 0$, the non-central χ^2 is a weighted sum of central χ^2 density functions with a Poisson weight factor. The gamma function, $\Gamma[a]$, is

$$\Gamma[a] = \begin{cases} (a - 1) \; ! & a = \text{integer} \\ \dfrac{[1 \cdot 3 \cdot 5, \cdots, (2a-1) \pi^{-(1/2)}}{2/n} & a = \text{integer} + (1/2) \\ \int_0^\infty x^{a-1} e^{-x} \, dx & a > 0, \text{ real} \end{cases}$$

See Mood, Graybill, Boas [1974], p. 534.

The χ^2 mean is $E_{\chi^2}(U) = n + 2\lambda$; the variance is $\mathrm{Var}_{\chi^2}(U) = 2n + 8\lambda$.   Moments two through four are

$$m_2 = 2(n + 2\lambda), \quad m_3 = 8(n + 3\lambda),$$

$$m_4 = 48(n + 4\lambda) + 4(n + 2\lambda)^2 .$$

[Kotz and Johnson, Vol. 6, p. 277.]

A.3   The Non-Central F Distribution

For $U_n \sim \chi^2(n, \lambda)$ and $U_d \sim \chi^2(d)$, and U_n and U_d statistically independent, the ratio

$$F = \frac{U_n / n}{U_d / d} \sim F(n, d, \lambda)$$

has the non-central F distribution.   The density function is

$$f(F) = \sum_{j=0}^{\infty} \left\{ \left[\frac{e^{-\lambda} \lambda^j}{j!} \; n^{(1/2)n+j} \; d^{(1/2)d} \; F^{(1/2)n+j-1} \right. \right.$$

$$\left. \left. \div \left[(d + nF)^{(1/2)(n+d)+k} \; B[(1/2)n+j, (1/2)d] \right] \right] \right\},$$

with $F \gtrless 0$, $n \gtrless 1$, $d \gtrless 1$, $\lambda \gtrless 0$, and $B[a,b] = \Gamma(a) \, \Gamma(b) / \Gamma(a+b)$.   When $\lambda = 0$, the summation is dropped.   The mean is $E_F(F) = [d/(d-2)][1 + 2\lambda/n]$ for $d > 2$, and is undefined otherwise.   The variance is

$$\mathrm{Var}_F(F) = \begin{cases} \dfrac{2d^2}{n^2(d-2)} \left[\dfrac{(n+2\lambda)^2}{(d-2)(d-4)} + \dfrac{n + 4\lambda}{d - 4} \right] & d > 4 \\ \text{undefined} & d \leq 4. \end{cases}$$

Other authors, SAS [1990, p. 578], Seber [1980, p. 5], and Pearson and Hartley [1976, p. 64] preface λ in the density function with (1/2), thereby eliminating the need to repeat the (1/2) when equating the non-centrality parameter to a quadratic form.

Pearson and Hartley [1976, p. 66-77] show examples of λ for one- and two-way classifications.   They include and show use of the Tang [1938] and Tiku [1967] charts of the power function for various values of α, n, d, and a function of λ.   Morrison [1983, p. 127-131] has an example of the assumptions needed to use the charts to determine sample size.   Duncan [1957] has graphs of

operating characteristics curves; Feldt and Mahnmoud
[1958] have charts for determining the sample size needed
to detect alternative values of the null hypothesis.
Scheffe' [1959, p. 42 and 446-454] presents the Fox power
charts. Both sets of charts are easier to read in
certain regions. We are unable to find the Lehmer charts
promised in the note ending page 42 of Scheffe'. Mann
[1949, p. 74-75] discusses Hsu's [1941] result on the
power of alternative tests.

A.4   The Non-central Beta Distribution

The transformation $R = (nF)/(nF + d)$ yields the
non-central Beta probability density function,
$f(R \mid n,d,\lambda)$

$$= \sum_{j=0}^{\infty} \left\{ \frac{e^{-\lambda} \lambda^j}{j!} \ R^{(1/2)n+j-1} \ (1-R)^{(1/2)d-1} \ \div \ B[\ (1/2)n+j, \ (1/2)d \ ] \right\},$$

with $R \geq 0$, $n \geq 1$, $d \geq 1$, $\lambda \geq 0$, and
$B[a,b] = \Gamma(a) \ \Gamma(b) \ / \ \Gamma(a+b)$.
When $\lambda = 0$, the summation is dropped. The reverse
transformation is

$$F = (dR)/[n(1-R)].$$

The value of R may be obtained directly from the
numerator and denominator χ^2 random variables used to
compute F, by substitution into the above definition for
R. Note that R equals the coefficient of determination,
usually denoted R^2.

$$R = U_n \ / \ [\ U_n + U_d \ ],$$

$$= \frac{\text{Sum of Squares due to the Model}}{\text{Total Sum of Squares}}.$$

The original derivation of the F-test used R
[Fisher, 1928, p. 671, distribution C]. Snedecor first
used the transformation from R to F, enabling him to
condense the tables of probabilities from those developed
for $z = (1/2)\ln [nR/(1-R)]$. See Seber [1980, p. 9 and
35] for this distribution.
The advantage of the R transformation is that for
some experimental designs, the denominator of R is
constant for all randomizations, which is not the case
for F. Unfortunately, the analysis of covariance has a
denominator which changes with each randomization for
both F and R.

APPENDIX B

THE R(·|·) NOTATION

The notation R(· , · , ···) denotes the reduction
in sum of squares due to fitting the model equation
containing those terms within the parentheses. The
reduction is from the total sum of squares $Y'Y$.

The sum of squares may be partitioned as follows:

R(·,·,···) = R(·,·,··· | ·,·,···) + R(·,·,···),

where terms in the last R(·) are those following the bar
(|) in the preceding term, R(·|·). The term R(·|·) is
the further reduction in the sum of squares achieved by
adding to the model equation those terms to the left of
the bar after the model equation already contains those
terms to the right of the bar. One reads the R(·|·)
notation as "the reduction in sums of squares due to
adding the terms ... (those to the left of the bar) to
the model equation which already contains the terms ...
(those to the right of the bar)". The terms to the right
of the bar comprise the "reduced" model equation. When
the terms to the left of the bar are added to those on
the right of the bar, the combined terms comprise the
"full" model equation.

In all cases R(·,·,···) has the form

$$Y' [X(X'X)^- X'] Y.$$

The X matrix contains columns for those terms indicated
within the parentheses of R(·,·,···). For example, the X
matrix corresponding to R(μ,b) is [X_μ | X_b], and the X
corresponding to R(μ,b,β) is [X_μ | X_b | Z]. (See Searle,
1971, pages 246-249 for a more complete discussion of
this notation. Searle [1987, p. 28-29] introduces a
different reduction notation, $\mathbb{R}$(·|·), not to be confused
with the above R(·|·).

APPENDIX C

SUMMATION SUM OF SQUARES TO PROJECTORS

VIA KRONECKER PRODUCTS

The $X(X'X)^{-1}X'$ matrix for the block and treatment sum of squares may be generated directly using Kronecker products, denoted as $\otimes$, and using J_c as the c x c matrix of ones. The summation notation for the centered sums of squares guides the choice of terms as in the following example. For

$$\overset{a}{\underset{i}{\Sigma}} \quad \overset{b}{\underset{j}{\Sigma}} \quad \overset{c}{\underset{k}{\Sigma}} \quad \overset{d}{\underset{l}{\Sigma}} \ (y_{i \cdot k \cdot} - y_i \ldots)^2 \ ,$$

$$X(X'X)^{-1}X' = I_a \otimes (1/b)J_b \otimes (I_c - (1/c)J_c) \otimes (1/d)J_d \ .$$

The subscripts of $y\ldots$ and $\bar{y}\ldots$ determine the form of the term of the Kronecker product as follows:
   1. When the same letter subscript appears in

       both $y\ldots$ and $\bar{y}\ldots$, as for i above, then use I_a.

   2. When both $y\ldots$ and $\bar{y}\ldots$ have been averaged over the same subscript, as is the case for j, then use $(1/b)J_b$; for l use $(1/l)J_l$.
   3. When the subscript letter appears in $y\ldots$,

       but not in $\bar{y}\ldots$, as for k above, then use $I_c - (1/c)J_c$.

The subscripts of $y\ldots$ must augment fastest from right to left in the same order as in the summation expression. A little manipulation will verify these rules.

   These rules were developed by Craig Jefferson and the author.

APPENDIX D

SUMMARY OF CONTRIBUTIONS TO THE DISCIPLINE

1.   Develops a permutation matrix technique to generate randomizations. This technique is more flexible than the $\Sigma\ Y_{ijk}\mathcal{S}_{ij}^{k}$ method.

2.   Recasts the noncentrality parameters into two parts, one changing with each randomization, the other constant for all randomizations. The first contains the the design matrix, which changes with each randomization. The second contains those terms typically not under the control of the experimentor. These terms are constant for all randomizations.

3.   Proves that, for certain models, different randomizations will provide identical values for certain estimates, analysis of variance terms, and noncentrality parameters, for experimental unit errors.

4.   Describes the clustering and dispersing effect of different types of errors on the values of certain estimates, analysis of variance terms, and noncentrality parameters.

5.   Proves a linkage between the noncentrality parameters in the single covariate case, useful in restricting randomizations to those randomizations which avoid low power.

6.   Provides an if and only if proof detailing when two matrices, such as design matrices, will have equal projectors, and hence equal sums of squares for a response variable.

7.   Develops the probability space for the randomization model, due in the main to Drs. Sahadeb Sarkar and Ignacy Kotlarski.

8.   Develops the model equation and probability space for a model combining unit errors and the classical normally distributed random errors.

9.   Shows the algebraic relationship between the errors of Dr. Kempthorne and those of Dr. Neyman.

10. Presents a two-stage randomization process which allows unequal selection probabilities, such as restricted randomization, to be easily incorporated into the defining and counting of the number of randomizations assigned nonzero selection probability.

11. Extends the randomization test and power computation to the case of unequal selection probabilities, and defines such under simulation.

12. Makes a modest start in linking the randomization model with concepts of survey sampling.

13. Links the model assumptions with the model equations for the randomized block design with covariates, in both matrix and summation notations.

14. Offers a more accessible presentation of the "Model Σ".

15. Programs in SAS a method of generating simulated errors which are orthogonal to the covariate.

16. Develops a computer-generated F-distribution analogue of normal probability paper, and other graphic means of analyzing numerical simulation results.

17. Briefly reviews the early assumptions as to the characteristics of the error term in a regression model equation.

18. Lists accessible literature on the analysis of covariance, the noncentrality parameter, generalized inverses, experimental unit and experimental unit-treatment interaction errors, exchangeability, and limiting theorems in the analysis of covariance.

19. Gears its presentation to rapidly introduce one familiar with the general linear model to the terminology and topics of the randomization model. Concepts from diverse sources are logically organized and presented in a uniform notation. Frequent internal and external references are provided.

APPENDIX E

COLLECTED SUGGESTED EXTENSIONS

This work has suggested a variety of extensions, some of which may appeal to those with interests and strengths in matrix algebra, SAS computer programming, survey sampling, limit theorems, etc. Some are suitable for a statistics major's senior paper, others require course work taken at the master's level, and others are suitable for doctoral theses. This dissertation will help beginners grasp the terminology and issues of research into the randomization model. Their progress, and the progress of the discipline, will be aided by a logical sequence of problems to solve. Such a sequence will span several students and years. A researcher who remains current in this area of research will need to alter the sequence to fit developments and available talent and interests. This author believes such an effort would be fruitful.

I.   First Steps – Generation of Errors

1.   Resolve the role of the singular normal distribution (singular Ψ) in the case of normal errors forced to sum to zero for each randomization. Begin the literature review with the sources in Searle [1971, p. 69]. Construct a set of normally distributed random errors which are standardized to variance 30, but not to mean zero within each randomization. Test this set of errors with the F-distribution probability plots of Section 4.6.1, (p. 290). This may isolate the off-diagonality seen for (n) in Graph 4.6.1, (p. 295) to the forcing of a zero sum of errors within each randomization. Examine Ψ alternative to that of p. 317. Examine the difference of (4.6.1), (p. 317), for patterns within g-groups or elsewhere.

2.   Once the case for the normally distributed error is resolved, proceed to the experimental unit error case, again examining alternative Ψ. Possibly, such "sum to zero" errors invalidate use of the singular F-distribution.

3.   Discover those conditions on the errors and design matrix which ensure a zero matrix in (4.6.1), (p.318).

4.   With a singular Ψ as the variance-covariance matrix, use generalized least squares to compute the estimates and F-ratios; see Searle [1971], p. 220-223, and if necessary citations found there.

5.   Review the question of changing the value of σ^2 with each randomization for the experimental unit-treatment interaction error cases.

6.   Consider using two normally distributed random error data sets. The one with mean zero and standard deviation constant in all randomizations would be for comparison with the experimental unit error cases. The second would have its mean and standard deviation in each randomization equal to that of the experimental unit-treatment interaction error case.

7.   Review the construction of the unit errors. Examine the implications of forcing each set to be orthogonal to the covariate. See Section 4.2.1, (p. 243).

8.   Discover if orthogonality with the covariate can be maintained for experimental unit-treatment interaction errors which also sum to zero over treatments (as these do) and over units within a block (these do not for k=4). See Section 4.2.2, (p. 248).

II   Further Understanding of the Projectors

1.   Discover the cause(s) of the extreme values of $\hat{\beta}$ for experimental unit error seen in Graph 4.5.3, (p.277).

2.   Discover the reason(s) for the constant shifts in the mean $\hat{\tau}$ for the experimental unit-treatment interaction and normally distributed random error cases as noted in Table 4.4.1, (p. 259).

3.   Discover the reason for the multiple observations at points on Graph 4.7.1 for $\tau_1 - \hat{\tau}_1$ vs Cox's ration two for experimental unit errors, (p. 331); see p. 214.

5.   Discover why the means (Table 4.4.1, p. 259) and standard deviations (Table 4.4.2, p. 262) for estimators and sums of squares for the normally distributed random error case are not what normal theory would predict.

6.   Disentangle the matrix which is under the experimentor's control from the expressions for the sums of squares as was done for the noncentrality parameters. See Theorems 3.7.10 and 3.7.11, (p. 211-216).

7.   Prove the relationships between
(a) the Cox ratios and the noncentrality parameter for
the adjusted covariate (see Graph 4.8.3, p. 340), and
(b) balancing on the covariate, g-groups and the
serpentine method (see Section 3.2.6, p. 114) and 4.8,
p. 335-362). Begin with a clearer understanding of Cox's
results.

III.   Further Simulations Using this Paper's Model

1.   Streamline the computer programs to permit one
submission to compute all 12 cases and analyze them.
Input would be the six sets of errors, the two values of
τ, and the values of mean, block, and covariate values.
Add one or more of the methods of power computation
discussed in Section 3.6.5 (p. 176-177). Analysis could
resemble Tables 4.4.1, (p. 259), 4.4.2, (p. 263), 482,
(p. 361); Graphs 4.5.1, (p. 272), 4.6.1, (p. 295), 4.6.7,
(p. 302); and Displays 4.8.7, (p. 350) and 4.8.10,
(p. 356). Consider formal tests of the equality of two
distributions. The analysis should have the capacity
for alternative model equations and assumptions.

2.   Investigate the three methods of power
computation of Section 3.6.5, (p. 176-177).

3. Examine other values of errors and model
equation parameters to obtain typical outcomes of the
simulation. Vary the relative magnitude of errors,
treatment effects, and covariates, and their variances.

4.   Examine the difference in dispersion effects of
the experimental unit-treatment interaction and normally
distributed random errors. See p. 200.

5.   Combine error types. See Sections 3.5, (p.
150) and 4.2.4, (p. 253).

6.   Examine the effects of a missing observation,
both missing experimental unit and missing treatment. A
brief DATA step prior to the PROC GLM step of SAS
program glm will do this. Adjust the reading of the glm
output to obtain the desired types of sums of squares.

7.   Bootstrap the randomization distribution, see
p. 133.

IV. Alternative Model Equations - Derivations

1. Repeat the derivations of such Theorems as 2.6.4 - 2.6.22 (p. 48-62), and 3.7.9-3.7.11 (p. 202-214) for:
 a) multiple experimental units per treatment per block, r>1.
 b) block x treatment interaction and placement of same when not statistically significant, see Bingham and Fienberg [1982].
 c) multiple factors, with and without factor interaction.
 d) a covariate which changes with the treatment.
 e) multiple covariates, with and without covariate interaction, constant for all treatments and changing with the treatment or with one factor. Orthogonality to all covariates will require a new method of generating the error terms.
 f) the general linear hypothesis (p. 41).

2. Derive the randomization model with mixed factors. The values of the covariates need to be considered as randomaly sampled from a larger population.

3. Examine comparing model equations via the ratio of their respective noncentrality parameters. Such a ratio cancels the unknown variance term when it is known to be the same for both models. See p. 40.

V. Alternative Model Equations - Simulations

1. With two blocks, add levels of the treatment variable to provide numerator and denominator degrees of freedom sufficient for the existence of the first four moments for the F-distribution. Compute values for them from the simulation. Construct Pearson curves to compare with the F-distribution. For the cases with nonzero τ and/or β, the noncentrality parameter changes with each randomization, as will the moments. A method is needed to compare such a set of moments with the single moment supplied by the F-distribution. The number of randomizations increases rapidly. See p. 265 and Cook [1951B].

2. Examine the robustness to erroneous model equations and variance-covariance matrices of the randomization model, with equal and unequal assignment probabilities. See (3) of Section 3.2.8, p. 120.

VI. Other Concepts

1. Merge concepts from survey sampling and the randomization model. The super-population model (Section 3.5.5, p. 163) appears to allow the experimental unit of the actual experiment to be viewed as a sample from a

larger population (p. 104). See Extensions III.5 and IV.2.

2. The concept of a strategy (Section 3.4.4, p. 147) permits developing estimators and test statistics tailored to particular sets of selection probabilities, such as those of Section 4.7, (p. 320) and 4.8, (p. 335). See Section 3.4.4, p. 147.

3. Develop the role exchangeability and partial exchangeability play in the randomization model, Section 3.5.6, p. 164). Develop the implications of de Finetti's representation theorem (p. 165).

4. Develop an analysis of covariance analog of the regression technique of analysis of pure error. See p. 198.

5. Consider models which reuse experimental units. The crossover model appears to use randomizations in the same g-group for both (all) experiments. Repeating the same treatment on some units may permit estimation of the magnitude of both types of unit errors.

6. Drop Assumption 10 (p. 158) of independence of randomization and normal distribution. This alters the probability space of (3.5.3), (p. 157).

7. Proving the conjecture concerning Theorem 3.2.1(c), (p. 113-114), may be quite easy.

8. The papers by Roux (p. 137) appear to suggest other simulations and relationships than those herein.

9. Sections 3.8.1-3.8.3 suggests some challenging approaches to finding the randomization expectation of the mean square errors in the analysis of covariance.

APPENDIX F

LISTING OF COMPUTER PROGRAMS

F.1 SUMMARY OF COMPUTER PROGRAMS

Program Name	ASCII File Name	Program Function
meu	spmeu	Create experimental unit error
meuxti	spmeuxti	Create experimental unit - treatment error
cterms	spcterms	Generate observed Y values for various error terms
glm	spglm	Generate estimates and ANOVA terms; produce means, variances, histograms, and plots
noncent	spnoncen	Generate noncentrality parameters
testf	sptestf	Test F-distribution probability plots
snormal	spsnorm	Test for singular normal distribution
probp	spprobp	Combine estimates, anova terms, non-centrality parameters, error variances; compute OSL's, generate F-distribution probability plots.
cox	spcox	Analyze best and worst cases.
gmeans	spgmeans	Display means of estimator by g-group.

    This copy is supplied with a 3.5 inch hard diskette containing the above SAS programs. The ASCII files are on a standard DOS formatted diskette with file names in the second column above.

    An IBM catridge containing transportable versions of the SAS data sets and the above SAS programs is available from the author.

F.2   Flow of Computer Programs

```
   ┌─────────┐         ┌──────────┐
   │  meu    │         │ meuxti   │
   └────┬────┘         └────┬─────┘
        │                   │
        ▼                   ▼
   ┌─────────┐         ┌──────────┐
   │ EU_ERR  │         │ EUTI_ERR │
   └────┬────┘         └────┬─────┘
        └────────┐  ┌───────┘
                 ▼  ▼
            ┌──────────┐
            │ cterms   │
            └────┬─────┘
                 │
                 ▼
            ┌──────────┐
            │ OBS_Y    │
            └────┬─────┘
                 │
                 ▼
            ┌───────┐        ┌────────┐
            │ glm   │───────▶│ YHAT   │
            └───┬───┘        └────────┘
                │
                ▼
            ┌──────────────┐
            │ F- RATIOS    │
            │ ESTIMATES    │
            └──────┬───────┘
                   │
                   ▼
   ┌──────────┐       ┌────────────┐    ┌──────────┐   ┌─────────┐
   │ sing_nor │◀──────│ λ VALUES   │◀───│ noncent  │◀──│ g_table │
   └────┬─────┘       └────────────┘    └──────────┘   └─────────┘
        │
        ▼
   ┌──────────┐
   │ probp    │
   └────┬─────┘
        │
        ▼
   TABLES, PLOTS, ETC.
```

```
U13293A.SASPROG.CNTL(MEU)
VPSPRINT 5.1.002   WEDNESDAY NOVEMBER 4,1992   14:59:12 U13293A  MVS1          **
VPSPRINT SASPROG.CNTL(*) LOCAL FORM(9001) CLASS(E)

//U13293AA  JOB (*),'ROBERT WILSON',
//  TIME=(1,0),MSGLEVEL=(1,1),MSGCLASS=X,CLASS=2
/*JOBPARM  ROOM=Y,FORMS=9001,COPIES=1
//S1  EXEC  SAS,OPTIONS='MACRO,DQUOTE'
//SASDATA    DD  DSN=U13293A.SASDATA.DATA,DISP=OLD
 *
 * Program File:  sasprog.cntl(meu)
 * Input File(s): none
 * Output File(s): sasdata.data(eu_err)
 *    Experimental Unit error with within-block variance equal and not.
 *
 *   Produces error terms --  e (sub ij).
 *   Type = experimental unit error.
 *           Unique for each experimental unit.
 *           Sums to zero within each block.
 *           Block variance=(block sum of squares)/(block num.of e.u.).
 *           Linearly independent with the vector of covariates.
 *
 *   Number of versions = 2.
 *           1. Block 1 variance EQuals block 2 variance = b_var.
 *               Variable: euEQ
 *           2. Block 2 variance is v_ratio times block 1 variance.
 *               Variable: euNE
 *        The sum of squares as pooled over all blocks is equal in both
 *        versions and equals b_var * NROW(eu)  * n_blocks.
 *;

        %LET n_blocks =  2;     *number of blocks ;
        %LET b_var     = 30;    *within-block variance when eq.;
        %LET v_ratio   = 4;     *var(block2)/var(block1) when not eq.;
        %LET max_iter  = 100;   *maximum number of iterations;
        %LET toler = .00000001; *maximum {tan(beta) - tan(betahat)};
                *For orthogonality of error and covariate,
                *toler should be less than .000 000 1  .
                *For this, max_iter should be about 100.;
 *
 *        Values for the slope coefficient and covariates
 *        are defined in data set initial, below.
 *
 *;
   ************* Debug Print Switch Macro and Switches   **********;
 * Macro for debug printing follows:;
   %MACRO  db(db_site,db_act);
      %IF %UPCASE(&db_site) = TRUE  | %UPCASE(&db_site) = T
          %THEN   %QUOTE(&db_act);
   %MEND db;
 *
*Debug print switches (dbps#):
 True or t = on (in upper and/or lower case), anything else = off.;
        %LET dbps1 = f    ;  *Echo print slope and covariates;
        %LET dbps2 = f    ;  *Uniform random numbers;
        %LET dbps3 = f    ;  *Initial errors and observed y;
        %LET dbps4 = f    ;  *Means and centered data;
        %LET dbps5 = f    ;  *Beta and rotation angle for centered data;
        %LET dbps6 = f    ;  *Rotation and trig function values;
        %LET dbps7 = f    ;  *Iteration results;
        %LET dbps8 = f    ;  *Postiteration betahat and eu;
        %LET dbps9 = f    ;  *Postiteration - characteristics of eu;
        %LET dbps10= f    ;  *Adjust to eq and ne block variances;
```

```
        %LET dbpsii= f       ;  *Final characteristics of eu;


*************** Values of Slope Coefficient and Covariates ******;
*
*   Each row of data set initial contains the block, the slope then the
*     covariate values for the experimental values from the same block.
*     Each row is one block,   i=1,2,....,b.
*     Variable beta = covariate slope coefficient.
*     Variables (columns) z1, z2,... = covariates, Z(sub ij).
*;
        DATA initial;   INPUT block_id    beta  z1-z4 ;
*   Debug Print option follows: ;
        %db(&dbpsi, PUT block_id= beta=  z1= z2= z3= z4= ; );
        CARDS;
           1      2      1   2  11  14
           2      2      5   9  16  19
           ;


************ Logic of Program ************************************;
*
*     For block i, posit a line,
*         Y(ij) =  beta*Z(ij) = beta*covariate(ij).
*     Consider the intercept (=mu+block+treatment) as zero.
*     Y values of this line are the error-free superpopulation points.
*     They are also the true Y(ij) values without error.
*     Construct a vector of initial errors which sum to zero.
*     Add the initial errors to the vector of superpopulation points.
*     This sum is the initial observed y vector.
*     Until the least squares slope of observed y vs covariates=beta
*       iteratively:
*         Center the y vector and covariates about the origin.
*         Rotate the observed y until its least squares slope = beta.
*         Subtract out the superpopulation y.
*         Adjust the new errors to sum to zero.
*         Repeat the iteration.
*     Adjust the errors to have the desired variances within each block.
*     Do this for equal and not equal block variances.
*;


******************** Logic of rotation ************************;
*     theta = tan(beta) for centered data.
*     Centered theta hat = c_thetah = angle of slope coefficient
*         of least squares fit of data centered on origin.
*     rotate = theta - c_thetah = angle (radians) to rotate data
*         to obtain proper angle of theta.
*
*     Angle of thetahat is smaller than angle of theta.
*       "Smaller" is less positive or more negative.
*       +betahat too shallow or -betahat too steep.
*       Therefore, rotate the data counterclockwise to increase its
*       slope, betahat value and thetahat value.
*
*     theta - thetahat = rotate < 0:
*     Angle of thetahat is larger than angle of theta.
*       "Larger" is more positive or less negative.
*       +betahat too steep or -betahat too shallow.
*       Therefore, rotate the data clockwise to decrease its
*       slope, betahat value and thetahat value.
*
*     Note,  SAS takes sin(neg angle) = -sin( abs(angle)).
*;

PROC IML;
****************** Begin looping through blocks ******************;
START mblock;
```

```
DO block = 1 TO &n_blocks;
   *move covariates and beta from data set initial into proc iml.;
      USE initial;  READ NEXT  INTO  indata;
      block_id = indata(| 1, 1:1 |);
      beta = indata(| 1, 2:2 |);
      z = indata(| 1, 3:6 |);    z = SHAPE(z,1,4); z=z¢;
  *************** Initial error vector (X axis, variable z) ********;
     *create vector of random variables.;
      eu=J(4,1,0);
      eu = uniform(repeat( 9284372828, 4, 1) );
      *It appears impossible to use a variable as the seed.;
*   Debug Print options follow: ;
    %db(&dbps2, PRINT "initial random eu" block eu;);
     *Make errors sum to zero.;
       mean_eu = eu(| : |);
       eu      =  eu - mean_eu;
*   Debug Print options follow: ;
    %db(&dbps3, PRINT  "centered eu"  block mean_eu  eu;);
  ****************** Initial observed y-values ******************;
     *sup_popy= superpopulation (error-free) values of Y.;
      sup_popy = beta*z;
      obs_y    = sup_popy + eu;
*   Debug Print options follow: ;
    %db(&dbps3, PRINT  "initial obs_y" block sup_popy  eu obs_y ;);
  ************* Center data about origin of XY plane *************;
       mean_y = obs_y(| : |);         cent_y = obs_y - mean_y;
       mean_z = z(| : |);             cent_z = z  -  mean_z;
       cent_dat = cent_y || cent_z;
*   Debug Print options follow: ;
    %db(&dbps4, PRINT block obs_y mean_y  z mean_z   cent_dat;);
  ************* Compute initial betahat for centered data *********;
       c_zpart = INV(cent_z¢*cent_z)*cent_z¢;
       c_betah = c_zpart*obs_y;    *Beta hat for centered data;
       *The centered data gives beta=tan(angle of l.sq. line.);
       *and centered theta hat.  theta = angle in radians.;
       theta = ATAN(beta);         c_thetah = ATAN(c_betah);
       rotate = theta - c_thetah;  *radians need to rotate.;
       s_rotate = sin(rotate);     c_rotate = cos(rotate);
    %db(&dbps5, PRINT  block c_betah theta    c_thetah;);
    %db(&dbps6, PRINT  block rotate  s_rotate  c_rotate;);
  ****************** Begin data rotation ********************;
      iter = 0;
      START mrotate;
      DO WHILE ( ABS(rotate ) > &toler   &   iter < &max_iter);
         cent_y = (cent_dat * (c_rotate // s_rotate) ) ;
      ******* Reset errors to sum to zero *******************;
         eu = cent_y - sup_popy;
         mean_eu = eu(| : |);
         eu = eu - mean_eu;
      ******* Recompute centered y vector *******************;
         obs_y = sup_popy + eu;
         mean_y = obs_y(| : |);
         cent_y = obs_y - mean_y;
         cent_dat = cent_y || cent_z;
      ******* Recompute centered beta hat and theta *******************;
         c_betah = c_zpart*cent_y; *was obs_y;
         c_thetah = ATAN(c_betah);
         rotate = theta - c_thetah;  *radians needed to rotate.;
         s_rotate = sin(rotate);     c_rotate = cos(rotate);
         iter = iter + 1;
*   Debug Print options follow: ;
    %db(&dbps7, PRINT  block toler  &max_iter iter
                    beta  c_betah   theta c_thetah rotate;);
     **************** Return to top of while loop ****************;
      END;  FINISH;  RUN mrotate;
```

```
******************* Iteration for one block complete ************;
    *return post-iteration observed y values to original position.;
    *obtain post-iteration betahat.;
        obs_y = obs_y + sup_popy(| : |);     *original eu had mean=0.;
        betahat = inv(z¢*z)*z¢*obs_y;
        eu= obs_y - sup_popy;
*   Debug Print options follow: ;
    %db(&dbps8, PRINT  "postiteration" block  iter sum_eu
                       beta  c_betah  betahat eu;);
    ************ Report characteristics of post-iteration eu ********;
        mean_eu = eu(| : |);
        eu = eu - mean_eu;
        sum_eu   = eu(| : |) * NROW(eu);
        var_eu  = eu(| ## |) / NROW(eu); *mean = zero;
        orthog = eu¢*z;
*   Debug Print options follow: ;
    %db(&dbps9, PRINT  "postiter " block   iter  rotate betahat eu
                       sum_eu var_eu  orthog;);
    *************** Adjust eu to desired variances ***************;
        eu_eq = SQRT(&b_var / var_eu) # eu;
        *For not equal errors, obtain total sum of squares for equal
         variances in all blocks.  Divide into v_ratio + 1 parts.
         Multiply block 1 by 1 part and block 2 by v_ratio parts.
         Divide by existing block variance.
         Square root the above.  Multiply by existing eu.;
        IF block = 1 THEN eu_ne = SQRT(
           &b_var * NROW(eu)  * &n_blocks / (&v_ratio +1) / NROW(eu)
                    / var_eu )  # eu;
        IF block = 2 THEN eu_ne = SQRT(
           &b_var * NROW(eu) * &n_blocks / (&v_ratio +1) / NROW(eu)
             * &v_ratio  / var_eu  )  # eu;
        *Compute final observed y ;
        obs_yeq = sup_popy + eu_eq;
        obs_yne = sup_popy + eu_ne;
*   Debug Print options follow: ;
    %db(&dbps10, PRINT "final eq & ne"
                 eu_eq obs_yeq    sup_popy    obs_yne  eu_ne ;);
    ******** Report characteristics of final eu_eq and eu_ne ********;
        betah_eq= INV(z¢*z)*z¢*obs_yeq;
        mean_eq = eu_eq(| : |);
        eu = eu_eq - mean_eq;
        sum_eq  = eu_eq(| : |) * NROW(eu_eq);
        var_eq  = eu_eq(| ## |) / NROW(eu_eq); *mean = zero;
        orthogeq  = eu_eq¢*z;

        betah_ne= INV(z¢*z)*z¢*obs_yne;
        mean_ne = eu_ne(| : |);
        eu = eu_ne - mean_ne;
        sum_ne  = eu_ne(| : |) * NROW(eu_ne);
        var_ne  = eu_ne(| ## |) / NROW(eu_ne); *mean = zero;
        orthogne  = eu_ne¢*z;
*   Debug Print options follow: ;
    %db(&dbps11, PRINT  "char. of final eu"
            block  iter   sum_eq  var_eq  orthogeq  betah_eq
                          sum_ne  var_ne  orthogne  betah_ne ;);
************ Move errors to output file  one e.u. per row ********;
    plot = ( DO(1, NROW(z),1) )¢; *DO yields a row vector;
    b_num = J(4,1,1) # block;
    IF block = 1 then
        CREATE eu_err
        VAR{ b_num  plot z   eu_eq  eu_ne sup_popy   obs_yeq obs_yne };
    APPEND;
****************** End of One Block *****************************;
END; FINISH; RUN mblock;
CLOSE eu_err;
*
```

```
********************* End of All Blocks ***************************;
RUN;
*


************** Label and Print Data Set eu_err ********************;
DATA eu_err; SET eu_err;
  covar = z;  block = b_num;  drop z b_num;
  IF plot  = 1 THEN plabel = 'A';       *for graphic output;
  IF plot  = 2 THEN plabel = 'B';
  IF plot  = 3 THEN plabel = 'C';
  IF plot  = 4 THEN plabel = 'D';
  LABEL block= 'block'
        plot = 'experimental unit'
        covar ='covariate'
        eu_eq ='e.u. error var-eq'
        eu_ne ='e.u. error var-ne'
        sup_popy='1st superpopulation y'
        obs_yeq ='obs. y  var-eq'
        obs_yne ='obs. y  var-ne'
        plabel = 'plot id label'
        ;
  PROC PRINT   DATA=eu_err;
    TITLE1 'Experimental Unit Errors  for cases: ';
    TITLE2 '(a) block variances are equal and';
    TITLE3 '(b) block variance of block 2 is 4 times that of block 1';
************** Plot Data for Visual Examination *********************;
PROC PLOT  DATA=eu_err  UNIFORM;    BY block;
    PLOT sup_popy * covar = plabel
                   obs_yeq * covar ='E'
                   obs_yne * covar ='N'
          / OVERLAY HPOS=50 HAXIS=0 TO 19 BY 1  HREF=0.0  VREF=0.0;
    PLOT eu_eq * covar ='E'   eu_ne * covar ='N'
          / OVERLAY HPOS=50 HAXIS=0 TO 19 BY 1  VREF=0.0;
PROC PLOT  DATA=eu_err  UNIFORM;
    PLOT sup_popy * covar = plabel
                   obs_yeq * covar ='E'
                   obs_yne * covar ='N'
          / OVERLAY HPOS=50 HAXIS=0 TO 19 BY 1  HREF=0.0  VREF=0.0;
    PLOT eu_eq * covar ='E'   eu_ne * covar ='N'
          / OVERLAY HPOS=50 HAXIS=0 TO 19 BY 1  VREF=0.0;
*
************** Move temporary data to permanent SAS save file. *******;

DATA sasdata.eu_err;  SET eu_err;
RUN;
```

```
U13293A.SASPROG.CNTL(MEUXTI)
VPSPRINT 5.1.002   WEDNESDAY NOVEMBER 4,1992   14:59:14 U13293A   MVS1          **
VPSPRINT SASPROG.CNTL(*) LOCAL FORM(9001) CLASS(E)

//U13293AA   JOB (*),'ROBERT WILSON',
//   TIME=(1,0),MSGLEVEL=(1,1),MSGCLASS=X,CLASS=2
/*JOBPARM   ROOM=R,FORMS=9001,COPIES=1
//S1  EXEC  SAS,OPTIONS='MACRO,DQUOTE'
//SASDATA   DD  DSN=U13293A.SASDATA.DATA,DISP=OLD
 *
 * Program File:  sasprog.cntl(meuxti)
 * Input File(s): none
 * Output File(s):  sasdata.data(euti_err)
 *     Experimental Unit-treatment interaction error
 *     with within-block variances equal and not.
 *
 *     Produces error terms --  e (sub ij).
 *     Internal to this and program meu, both error types are named eu.
 *     Type = experimental unit - treatment interaction error.
 *             Unique for each experimental unit.
 *             Sums to zero over each plot, for each treatment.
 *             Sums to near zero over each treatment, for each plot.
 *             Errors for last treatment are -(sum of errors for rest).
 *             Block variance=(block sum of squares)/(block num.of e.u.).
 *             Linearly independent with the vector of covariates.
 *
 *     Number of versions = 2.
 *          1. Block 1 variance EQuals block 2 variance = b_var.
 *             Variable: euEQ for each treatment
 *          2. Block 2 variance is v_ratio times block 1 variance.
 *             Variable: euNE for each treatment              .
 *          The sum of squares as pooled over all blocks is equal in both
 *          versions and equals b_var * NROW(eu)  * n_blocks.
 *;

        %LET n_blocks =  2;     *number of blocks ;
        %LET n_treats =  4;     *number of treatments;
        %LET b_var    = 30;     *within-block variance when eq.;
        %LET v_ratio  =  4;     *var(block2)/var(block1) when not eq.;
        %LET max_iter = 100;    *maximum number of iterations;
        %LET toler = .00000001; *maximum {tan(beta) - tan(betahat)};
             *For orthogonality of error and covariate,
             *toler should be less than .000 000 1  .
             *For this, max_iter should be about 100.;
 *
 *          Values for the slope coefficient and covariates
 *          are defined in data set initial, below.
 *
 *;
   ************* Debug Print Switch Macro and Switches   **********;
 * Macro for debug printing follows:;
     %MACRO  db(db_site,db_act);
        %IF %UPCASE(&db_site) = TRUE  | %UPCASE(&db_site) = T
           %THEN   %QUOTE(&db_act);
     %MEND db;
 *
*Debug print switches (dbps#):
 True or t = on (in upper and/or lower case), anything else = off.;
        %LET dbps1 = false ;  *Echo print slope and covariates;
        %LET dbps2 = f     ;  *Uniform random numbers;
        %LET dbps3 = f     ;  *Initial errors and observed y;
        %LET dbps4 = f     ;  *Means and centered data;
        %LET dbps5 = f     ;  *Beta and rotation angle for centered data;
```

```
        %LET dbps6 = f     ;   *Rotation and trig function values;
        %LET dbps7 = f     ;   *Iteration results;
        %LET dbps8 = f     ;   *Postiteration betahat and eu;
        %LET dbps9 = f     ;   *Postiteration - characteristics of eu;
        %LET dbps10= f     ;   *Adjust to eq and ne block variances;
        %LET dbps11= f     ;   *Final characteristics of eu;
        %LET dbps12= f     ;   *Initial eu for last treatment;
        %LET dbps13= f     ;   *Final   eu for last treatment;
        %LET dbps14= f     ;   *Final characteristics of eu-last treat;


*************** Values of Slope Coefficient and Covariates *******;
*
*   Each row of data set initial contains the block, the slope then the
*     covariate values for the experimental values from the same block.
*     Each row is one block,  i=1,2,....,b.
*     Variable beta = covariate slope coefficient.
*     Variables (columns) z1, z2,... = covariates, Z(sub ij).
*;
        DATA initial;   INPUT block_id    beta   z1-z4 ;
*   Debug Print option follows: ;
        %db(&dbps1, PUT block_id= beta=  z1= z2= z3= z4= ; );
        CARDS;
            1      2    1   2  11  14
            2      2    5   9  16  19
            ;


************* Logic of Program ***********************************;
*
* Repeat the following for the first n_treats-1.
* Errors for the last treatment are defined to be the
*  negative of the sum of errors for the first n_treats-1 treatments.
*
*   For block i, posit a line,
*      Y(ij) =  beta*Z(ij) = beta*covariate(ij).
*   Consider the intercept (=mu+block+treatment) as zero.
*   Y values of this line are the error-free superpopulation points.
*   They are also the true Y(ij) values without error.
*   Construct a vector of initial errors which sum to zero.
*   Add the initial errors to the vector of superpopulation points.
*   This sum is the initial observed y vector.
*   Until the least squares slope of observed y vs covariates=beta
*     iteratively:
*.      Center the y vector and covariates about the origin.
*       Rotate the observed y until its least squares slope = beta.
*       Subtract out the superpopulation y.
*       Adjust the new errors to sum to zero.
*       Repeat the iteration.
*   Adjust the errors to have the desired variances within each block.
*   Do this for equal and not equal block variances.
*;


******************** Logic of rotation ***********************;
*    theta = tan(beta) for centered data.
*    Centered theta hat = c_thetah = angle of slope coefficient
*         of least squares fit of data centered on origin.
*    rotate = theta - c_thetah = angle (radians) to rotate data
*         to obtain proper angle of theta.
*
*    Angle of thetahat is smaller than angle of theta.
*     "Smaller" is less positive or more negative.
*     +betahat too shallow or -betahat too steep.
*     Therefore, rotate the data counterclockwise to increase its
*     slope, betahat value and thetahat value.
*
*    theta - thetahat = rotate < 0:
```

```
*    Angle of thetahat is larger than angle of theta.
*     "Larger" is more positive or less negative.
*     +betahat too steep or -betahat too shallow.
*     Therefore, rotate the data clockwise to decrease its
*     slope, betahat value and thetahat value.
*
*    Note,  SAS takes sin(neg angle) = -sin( abs(angle)).
*;

PROC IML;
****************** Begin looping through blocks ********************;
START mblock;
DO block = 1 TO &n_blocks;
   *Create column vector of block id number.;
    b_num = J(4,1,1) # block;
   *initialize totals of error vectors;
    t_eu_eq = J(4,1,0);   t_eu_ne = J(4,1,0);
   *move covariates and beta from data set initial into proc iml.;
       USE initial;  READ NEXT  INTO  indata;
       block_id = indata(| 1, 1:1 |);
       beta = indata(| 1, 2:2 |);
       z = indata(| 1, 3:6 |);     z = SHAPE(z,1,4); z=z¢;
   *Create column vector of plot id numbers.;
    plot = ( DO(1, NROW(z),1) )¢; *DO yields a row vector;
   *************** Begin looping through treatments ****************;
START mtreat;
DO treat = 1 to &n_treats-1;                        ,
   *Create column vector of treatment id numbers.;
    t_num = J(4,1,1) # treat;
   ***************** Initial error vector (X axis, variable z) ********;
      *create vector of random variables.;
   *    eu=J(4,1,0);
        eu = uniform(repeat( 5382743916, 4, 1) );
        *It appears impossible to use a variable as the seed.;
*   Debug Print options follow: ;
    %db(&dbps2, PRINT "initial random eu" block eu;);
      *Make errors sum to zero.;
         mean_eu = eu(| : |);
         eu      =  eu - mean_eu;
*   Debug Print options follow: ;
    %db(&dbps3, PRINT  "centered eu"  block mean_eu  eu;);
   ******************** Initial observed y-values ******************;
      *sup_popy= superpopulation (error-free) values of Y.;
         sup_popy = beta*z;
         obs_y    = sup_popy + eu;
*   Debug Print options follow: ;
    %db(&dbps3, PRINT  "initial obs_y" block sup_popy  eu obs_y ;);
   ************* Center data about origin of XY plane **************;
            mean_y = obs_y(| : |);          cent_y = obs_y - mean_y;
            mean_z = z(| : |);              cent_z = z  -  mean_z;
            cent_dat = cent_y || cent_z;
*   Debug Print options follow: ;
    %db(&dbps4, PRINT block obs_y mean_y  z mean_z   cent_dat;);
   ************* Compute initial betahat for centered data *********;
            c_zpart = INV(cent_z¢*cent_z)*cent_z¢;
            c_betah = c_zpart*obs_y;    *Beta hat for centered data;
            *The centered data gives beta=tan(angle of l.sq. line.);
            *and centered theta hat.  theta = angle in radians.;
            theta = ATAN(beta);         c_thetah = ATAN(c_betah);
            rotate = theta - c_thetah;  *radians need to rotate.;
            s_rotate = sin(rotate);     c_rotate = cos(rotate);
    %db(&dbps5, PRINT  block c_betah theta     c_thetah;);
    %db(&dbps6, PRINT  block rotate s_rotate c_rotate;);
   ****************** Begin data rotation ********************;
       iter = 0;
       START mrotate;
```

```
        DO WHILE ( ABS(rotate ) > &toler   &   iter < &max_iter);
            cent_y = (cent_dat * (c_rotate // s_rotate) ) ;
        ******* Reset errors to sum to zero ***************************;
            eu = cent_y - sup_popy;
            mean_eu = eu(| : |);
            eu = eu - mean_eu;
        ******* Recompute centered y vector **************************;
            obs_y = sup_popy + eu;
            mean_y = obs_y(| : |);
            cent_y = obs_y - mean_y;
            cent_dat = cent_y || cent_z;
        ******* Recompute centered beta hat and theta *****************;
            c_betah = c_zpart*cent_y; *was obs_y;
            c_thetah = ATAN(c_betah);
            rotate = theta - c_thetah;  *radians needed to rotate.;
            s_rotate = sin(rotate);     c_rotate = cos(rotate);
            iter = iter + 1;
*    Debug Print options follow: ;
     %db(&dbps7, PRINT  block toler  &max_iter iter
                        beta  c_betah  theta c_thetah rotate;);
     ***************** Return to top of while loop *****************;
     END;  FINISH;  RUN mrotate;

   ******************** Iteration for one error complete ************;|
     *return post-iteration observed y values to original position.;
     *obtain post-iteration betahat.;
         obs_y = obs_y + sup_popy(| : |);     *original eu had mean=0.;
         betahat = inv(z¢*z)*z¢*obs_y;
         eu= obs_y - sup_popy;
*    Debug Print options follow: ;
     %db(&dbps8, PRINT  "postiteration" block  iter sum_eu
                        beta  c_betah  betahat eu;);
     ************ Report characteristics of post-iteration eu ********;
         mean_eu = eu(| : |);
         eu = eu - mean_eu;
         sum_eu  = eu(| : |) * NROW(eu);
         var_eu  = eu(| ## |) / NROW(eu); *mean = zero;
         orthog  = eu¢*z;
*    Debug Print options follow: ;
     %db(&dbps9, PRINT  "postiter " block   iter  rotate betahat eu
                        sum_eu var_eu  orthog;);
     *************** Adjust eu to desired variances  ***************;
         eu_eq = SQRT(&b_var / var_eu) # eu;
         *For not equal errors, obtain total sum of squares for equal
          variances in all blocks.  Divide into v_ratio + 1 parts.
          Multiply block 1 by 1 part and block 2 by v_ratio parts.
          Divide by existing block variance.
          Square root the above.  Multiply by existing eu.;
         IF block = 1 THEN eu_ne = SQRT(
          &b_var * NROW(eu)  * &n_blocks / (&v_ratio +1) / NROW(eu)
                      / var_eu )  # eu;
         IF block = 2 THEN eu_ne = SQRT(
          &b_var * NROW(eu) * &n_blocks / (&v_ratio +1) / NROW(eu)
          * &v_ratio  / var_eu  )  # eu;
       *Compute final observed y ;
       obs_yeq = sup_popy + eu_eq;
       obs_yne = sup_popy + eu_ne;
*    Debug Print options follow: ;
     %db(&dbps10, PRINT "final eq & ne"
                   eu_eq obs_yeq    sup_popy    obs_yne  eu_ne ;);
     ******** Report characteristics of final eu_eq and eu_ne ********;
         betah_eq= INV(z¢*z)*z¢*obs_yeq;
         mean_eq = eu_eq(| : |);
         eu = eu_eq - mean_eq;
         sum_eq  = eu_eq(| : |) * NROW(eu_eq);
         var_eq  = eu_eq(| ## |) / NROW(eu_eq); *mean = zero;
```

```
        orthogeq  = eu_eq¢*z;

        betah_ne= INV(z¢*z)*z¢*obs_yne;
        mean_ne = eu_ne(| : |);
        eu = eu_ne - mean_ne;
        sum_ne  = eu_ne(| : |) * NROW(eu_ne);
        var_ne  = eu_ne(| ## |) / NROW(eu_ne); *mean = zero;
        orthogne  = eu_ne¢*z;
*   Debug Print options follow: ;
    %db(&dbps11, PRINT  "char. of final eu"
        block  treat iter   sum_eq  var_eq  orthogeq  betah_eq
                            sum_ne  var_ne  orthogne  betah_ne ;);
************ Move errors to output file  one e.u. per row *********;
    IF block = 1 &  treat = 1 then
        CREATE euti_err
        VAR{b_num  plot t_num z .eu_eq eu_ne sup_popy obs_yeq obs_yne };
    APPEND;
    t_eu_eq = t_eu_eq + eu_eq; *Cumulative totals for last treatment.;
    t_eu_ne = t_eu_ne + eu_ne;
****************** End of One Treatment ***********************;
    END; FINISH; RUN mtreat;
*************** End of All but Last Treatment ********************;

******************** Create errors for final treatment ************;
    t_num = J(4,1,1) # treat;
    eu_eq = - t_eu_eq;       *mean already equals zero;
    eu_ne = - t_eu_ne;
    var_eq  = eu_eq(| ## |) / NROW(eu_eq); *mean = zero;
    var_ne  = eu_ne(| ## |) / NROW(eu_ne); *mean = zero;
*   Debug Print options follow: ;
    %db(&dbps12, PRINT  "initial last treat"
        block  treat  eu_eq  var_eq  eu_ne  var_ne;);
    *************** Adjust eu to desired variances  ****************;
        eu_eq = SQRT(&b_var / var_eq) # eu_eq;
        IF block = 1 THEN eu_ne = SQRT(
          &b_var * NROW(eu)  * &n_blocks / (&v_ratio +1) / NROW(eu)
                    / var_ne ) # eu_ne;
        IF block = 2 THEN eu_ne = SQRT(
          &b_var * NROW(eu) * &n_blocks / (&v_ratio +1) / NROW(eu)
          * &v_ratio  / var_ne ) # eu_ne;
        *Compute final observed y for last treatment;
        obs_yeq = sup_popy + eu_eq;
        obs_yne = sup_popy + eu_ne;
    ***** Report characteristics of errors for final treatment *****;
        betah_eq= INV(z¢*z)*z¢*obs_yeq;
        mean_eq = eu_eq(| : |);
        sum_eq  = eu_eq(| : |) * NROW(eu_eq);
        var_eq  = eu_eq(| ## |) / NROW(eu_eq); *mean = zero;
        orthogeq  = eu_eq¢*z;

        betah_ne= INV(z¢*z)*z¢*obs_yne;
        mean_ne = eu_ne(| : |);
        sum_ne  = eu_ne(| : |) * NROW(eu_ne);
        var_ne  = eu_ne(| ## |) / NROW(eu_ne); *mean = zero;
        orthogne  = eu_ne¢*z;
*   Debug Print options follow: ;
    %db(&dbps13, PRINT  "final eu-last treat"
        block treat iter   eu_eq  eu_ne ;) ;
*   Debug Print options follow: ;
    %db(&dbps14, PRINT  "char. of final eu-last treat"
        block treat iter   sum_eq  var_eq  orthogeq  betah_eq
                            sum_ne  var_ne  orthogne  betah_ne ;);

    APPEND;

******************** End of One Block ***************************;
 END; FINISH; RUN mblock;
```

```
********************* End of All Blocks ****************************;
  CLOSE euti_err;
*
RUN;  *End of creation of errors.;
* exit proc iml;
************** Label and Print Data Set euti_err *******************;
TITLE1 'Experimental Unit x Treatment Interaction Errors';
DATA euti_err; SET euti_err;
  covar = z;  block = b_num;  treat = t_num;  drop z b_num t_num;
  IF plot  = 1 THEN plabel = 'A';      *for graphic output;
  IF plot  = 2 THEN plabel = 'B';
  IF plot  = 3 THEN plabel = 'C';
  IF plot  = 4 THEN plabel = 'D';
  IF treat = 1 THEN tlabel = 'A';      *for graphic output;
  IF treat = 2 THEN tlabel = 'B';
  IF treat = 3 THEN tlabel = 'C';
  IF treat = 4 THEN tlabel = 'D';
  LABEL block= 'block'
        plot = 'experimental unit'
        treat= 'treatment'
        covar ='covariate'
        eu_eq ='e.u. error var-eq'
        eu_ne ='e.u. error var-ne'
        sup_popy='superpopulation y'
        obs_yeq ='obs. y  var-eq'
        obs_yne ='obs. y  var-ne'
        plabel = 'plot id label'
        tlabel = 'treatment id label'
        ;

  PROC PRINT   DATA=euti_err;
    VAR  block plot treat covar eu_eq  eu_ne   sup_popy
         obs_yeq obs_yne  plabel tlabel;
  TITLE2 '(a) within-block variances are equal and';
  TITLE3 '(b) block variance of block 2 is 4 times that of block 1';
  PROC MEANS   DATA=euti_err  N SUM MEAN VAR MIN MAX  VARDEF=N;
      BY block treat;
      VAR eu_eq  eu_ne;

  TITLE2 ;
  TITLE3 ;

  PROC SORT    DATA=euti_err;  BY block plot;
  PROC PRINT   DATA=euti_err;
    VAR  block plot treat covar eu_eq  eu_ne  sup_popy
         obs_yeq obs_yne  plabel tlabel;
    TITLE1 'Experimental Unit Errors  for cases: ';
    TITLE2 ;
    TITLE3 ;
  PROC MEANS   DATA=euti_err  N SUM MEAN VAR MIN MAX  VARDEF=N;
      BY block plot ;
      VAR eu_eq  eu_ne;

  PROC MEANS   DATA=euti_err  N SUM MEAN VAR MIN MAX  VARDEF=N;
      BY block;
      VAR  eu_eq  eu_ne;

  PROC MEANS   DATA=euti_err  N SUM MEAN VAR MIN MAX  VARDEF=N;
      VAR eu_eq  eu_ne;

  PROC SORT    DATA=euti_err;  BY block treat;
************** Plot Data for Visual Examination *******************;
PROC PLOT  DATA=euti_err  UNIFORM;    BY block treat;
    PLOT sup_popy *  covar = plabel
                     obs_yeq * covar ='E'
                     obs_yne * covar ='N'
```

```
              / OVERLAY  HPOS=50    HAXIS=O TO 19 BY 1   HREF=O.O VREF=O.O;
      PLOT eu_eq * covar ='E'      eu_ne * covar ='N'
              / OVERLAY  HPOS=50    HAXIS=O TO 19 BY 1   VREF=O.O;
PROC PLOT  DATA=euti_err  UNIFORM;    BY block;
      PLOT sup_popy * covar = plabel
                      obs_yeq * covar ='E'
                      obs_yne * covar ='N'
              / OVERLAY  HPOS=50    HAXIS=O TO 19 BY 1   HREF=O.O VREF=O.O;
      PLOT eu_eq * covar =tlabel    eu_ne * covar =tlabel
              / HPOS=50  HAXIS=O TO 19 BY 1   VREF=O.O;
PROC PLOT  DATA=euti_err  UNIFORM;
      PLOT sup_popy * covar = plabel
                      obs_yeq * covar ='E'
                      obs_yne * covar ='N'
              / OVERLAY  HPOS=50    HAXIS=O TO 19 BY 1   HREF=O.O VREF=O.O;
      PLOT eu_eq * covar =tlabel    eu_ne * covar =tlabel
              / HPOS=50  HAXIS=O TO 19 BY 1   VREF=O.O;

************** Move temporary data to permanent SAS save file. *******;
DATA sasdata.euti_err;  SET euti_err;
RUN; *End of program;
```

```
U13293A.SASPROG.CNTL(CTERMS)
VPSPRINT 5.1.002   WEDNESDAY NOVEMBER 4,1992   14:59:05 U13293A   MVS1          **
VPSPRINT SASPROG.CNTL(*) LOCAL FORM(9001) CLASS(E)

//U13293AA  JOB (*),'ROBERT WILSON',
//  TIME=(1,0),MSGLEVEL=(1,1),MSGCLASS=X,CLASS=2
/*JOBPARM  ROOM=Y,FORMS=9001,COPIES=1
//S1  EXEC  SAS,OPTIONS='MACRO,DQUOTE'
//SASDATA   DD  DSN=U13293A.SASDATA.DATA,DISP=OLD


PROC CONTENTS  DATA=sasdata._ALL_ ;



*
* Program File:  sasprog.cntl(cterms)   Collect TERMS
* Step One:
*     Input File(s): eu_err and euti_err
*     Output File:   rm_error   one obs per eu per randomization
* Step Two:
*     Create standardized normal errors, store in file nor_err.
* Step Three:
*     Merge eu, euti, and normal error, store in file all_err.
* Step Four:
*     Combine (sum) errors, store in file comb_err.
* Step Five:
*     Standardize combined sums, store in file std_err.
* Step Six:
*     Create observed y's for combined errors, store in file obs_y.
* Step Seven:
*     Attach variable g_group to the file.
*
*
*
* Output File(s):  sasdata.data(obs_y)
*    One record per experimental unit for each
*    experimental-randomization = {(t!)**b}*{bt}
*
*   Collects model parameter terms (mu, blocks, treats, beta, covar)
*        and error terms (e.u.,  eu-ti, and normal).
*
*   With the above, construct the various observed y values.
*   Two super population y values are defined:
*            s_popynt = mu + b_val + beta*z   (no treatment)
*            s_popywt = mu + b_val + beta*z + trt_val (with treat.)
*
*
*
*************** Debug Print Switch Macro and Switches   **********;
* Macro for debug printing follows::;
     %MACRO  db(db_site,db_act);
        %IF %UPCASE(&db_site) = TRUE  | %UPCASE(&db_site) = T
           %THEN    %QUOTE(&db_act);
     %MEND db;
*
*Debug print switches (dbps#):
 True or t = on (in upper and/or lower case), anything else = off.;
        %LET dbps1 = f     ; *data set eu_err & euti_err as into IML;
        %LET dbps2 = f     ; *output to rm_err within loop of IML;
        %LET dbps3 = f     ; *data set rm_err - first observations;
        %LET dbps4 = f     ; *data nor_err eq var only - 1st obs;
        %LET dbps5 = f     ; *data nor_err eq & ne var - 1st obs;
```

```
         %LET dbps6 = f    ;  *data all_err - eu, euti & normal-1st obs;
         %LET dbps7 = f    ;  *data comb_err before standardizing;
         %LET dbps8 = f    ;  *data std_err after 1st standardizing;
         %LET dbps9 = f    ;  *data std_err after 2nd standardizing;
         %LET dbps10= f    ;  *data obs_y  observed y values;
         %LET dbps11= f    ;  *;


**********************  End of Debug Print Swithces ****************;


*******************************************************************;
*Step One:
 Produce data set of one obs. per eu, per randomization, with
     experimental unit and eu-treatment interaction errors.;

 /*  *Successful step turned off. ;

PROC IML;
******************* Enter Model Equation Terms and Errors **********;
*Define constants;
 mu_val = {10.0};
 b_valm = {-1.5, 1.5};
 trt_valm = {-6.5, -3.5, 2.5, 7.5};
 beta = {2.0};
 labels = { "A" , "B",  "C",  "D" };


*Collect experimental unit errors;
 *   vector name eu_eq and eu_ne  appear in two input data sets
 *   Two super population y values are defined:
 *           s_popynt = mu_val + b_val + beta*z   (no treatment)
 *           s_popywt = mu_val + b_val + beta*z + trt_val (with treat.)
 *
 *
eu_eqm = J(8,1,0);    eu_nem = J(8,1,0);    zm = J(8,1,0);
sup_ym = j(8,1,0);    *Without mean, blocks, or treatments;
USE sasdata.eu_err ;
   READ ALL  VAR{eu_eq}  INTO  eu_eqm;
   READ ALL  VAR{eu_ne}  INTO  eu_nem;
   READ ALL  VAR{covar}  INTO  zm;

*Collect experimental unit-treatment interaction errors;
euti_eq1 = J(32,1,0);     euti_ne1 = J(32,1,0);
euti_eqm = J( 8,4,0);     euti_nem = J( 8,4,0);
USE sasdata.euti_err ;
   READ ALL  VAR{eu_eq}  INTO  euti_eq1;
   READ ALL  VAR{eu_ne}  INTO  euti_ne1;
*Reshape to block-plot rows by treatments columns.;
euti_eqm( | 1:4  , 1 | ) = euti_eq1( |  1:4  | );
euti_eqm( | 1:4  , 2 | ) = euti_eq1( |  5:8  | );
euti_eqm( | 1:4  , 3 | ) = euti_eq1( |  9:12 | );
euti_eqm( | 1:4  , 4 | ) = euti_eq1( | 13:16 | );
euti_eqm( | 5:8  , 1 | ) = euti_eq1( | 17:20 | );
euti_eqm( | 5:8  , 2 | ) = euti_eq1( | 21:24 | );
euti_eqm( | 5:8  , 3 | ) = euti_eq1( | 25:28 | );
euti_eqm( | 5:8  , 4 | ) = euti_eq1( | 29:32 | );

euti_nem( | 1:4  , 1 | ) = euti_ne1( |  1:4  | );
euti_nem( | 1:4  , 2 | ) = euti_ne1( |  5:8  | );
euti_nem( | 1:4  , 3 | ) = euti_ne1( |  9:12 | );
euti_nem( | 1:4  , 4 | ) = euti_ne1( | 13:16 | );
euti_nem( | 5:8  , 1 | ) = euti_ne1( | 17:20 | );
euti_nem( | 5:8  , 2 | ) = euti_ne1( | 21:24 | );
euti_nem( | 5:8  , 3 | ) = euti_ne1( | 25:28 | );
euti_nem( | 5:8  , 4 | ) = euti_ne1( | 29:32 | );

*   Debug Print options follow: ;
     %db(&dbps1. PRINT   "Data input from eu err and euti err  (dbps1)";
```

```
                    PRINT   mu_val  b_valm  trt_valm    beta  zm
                            labels
                            eu_eqm eu_nem     euti_eqm   euti_nem; );

************************ Create Output Data Set ********************;
block = J(1,1,0);   plot = J(1,1,0);
*Data set  Randomization Model ERRor;
CREATE rm_err
    VAR{ exp_rand  b1_rand b2_rand   block plot plabel    treat tlabel
         mu_val  b_val   trt_val beta  z
         eu_eq eu_ne    euti_eq euti_ne   s_popynt s_popywt };

*********************** Define Randomization Pattern **************;
rpattern =
  {1 2 3 4 , 1 2 4 3 , 1 3 2 4 , 1 3 4 2 , 1 4 2 3 , 1 4 3 2,
   2 1 3 4 , 2 1 4 3 , 2 3 1 4 , 2 3 4 1 , 2 4 1 3 , 2 4 3 1,
   3 1 2 4 , 3 1 4 2 , 3 2 1 4 , 3 2 4 1 , 3 4 1 2 , 3 4 2 1,
   4 1 2 3 , 4 1 3 2 , 4 2 1 3 , 4 2 3 1 , 4 3 1 2 , 4 3 2 1};

********** Begin Looping through Experimental-Randomizations ********;
START mb1_rand;
DO b1_rand = 1 TO 24;           *All randomizations for block 1, (t!).;
   START mb2_rand;
   DO b2_rand = 1 TO 24;        *All randomizations for block 2, (t!).;
      START mblock;
      DO block = 1 TO 2;
      b_val = b_valm(| block |);
      exp_rand = (b1_rand - 1) # 24   +   b2_rand;
         START mplot;
         DO plot = 1 TO 4;
            plabel= labels(| plot |);
            s_popynt=sup_ym(| (block-1)#4 + plot |);
            eu_eq =  eu_eqm(| (block-1)#4 + plot |);
            eu_ne =  eu_nem(| (block-1)#4 + plot |);
            z     =      zm(| (block-1)#4 + plot |);
            sup_popy=   sup_ym(| (block-1)#4 + plot |);
            IF block = 1 THEN
               treat = rpattern(| b1_rand , plot |);
               ELSE
               treat = rpattern(| b2_rand , plot |);
            tlabel  =   labels(| treat |);
            trt_val = trt_valm(| treat |);
            s_popynt = mu_val + b_val + (beta * z );
            s_popywt = mu_val + b_val + (beta * z ) + trt_val;
            euti_eq = euti_eqm(| (block-1)#4 + plot , treat |);
            euti_ne = euti_nem(| (block-1)#4 + plot , treat |);

*    Debug Print options follow: ;
     %db(&dbps2, PRINT  "Data set rm_err  within loop  (dbps2)";
                 PRINT  exp_rand b1_rand b2_rand   block  plot
                        plabel  tlabel
                        mu_val  b_val  trt_val     beta  z
                        s_popynt  s_popywt
                        eu_eq eu_ne      euti_eq   euti_ne ; );
             APPEND;

         END; FINISH;  RUN mplot;
      END; FINISH;  RUN mblock;
   END; FINISH;  RUN mb2_rand;
END; FINISH;  RUN mb1_rand;
RUN;


********* Test Print Initial Version of Data Set obs_y *****;

*    Debug Print options follow: ;
```

```
        %db(&dbps3, TITLE1 "Observations from  rm_err  (dbps3)";
                    PROC PRINT DATA=rm_err (OBS=32);
                    PROC PRINT DATA=rm_err (FIRSTOBS=4577);
                    PROC MEANS DATA=rm_err (OBS=32)
                         N MEAN VAR STD  VARDEF=N;
                      BY  exp_rand  block;
                      VAR eu_eq euti_eq  eu_ne euti_ne;
                    PROC MEANS DATA=rm_err (OBS=32)
                         N MEAN VAR STD  VARDEF=N;
                      BY exp_rand;
                      VAR eu_eq euti_eq  eu_ne euti_ne;
                   PROC MEANS DATA=rm_err
                         N MEAN VAR STD  VARDEF=N;
                      VAR eu_eq euti_eq  eu_ne euti_ne;
                    );
DATA sasdata.rm_err;  SET rm_err;     *Randomization Model ERRor;

  End of successful step.  */
  ********************************************************************;
  *Step Two:
  Generate Normally Distributed Random Errors;

    /*   *Successful step turned off.;

*Normal random error terms have various variances;
DATA nor_err      (DROP= sq03  sq10  sq30  sq90);
   *Obtain very accurate values for standard deviation.;
   sq03 = SQRT(03);
   sq10 = SQRT(10);
   sq30 = SQRT(30);
   sq90 = SQRT(90);
   PUT sq03= sq10= sq30= sq90=;

   DO exp_rand = 1 TO 576  BY 1; *576;
      DO block = 1 TO 2    BY 1;
         DO plot = 1 TO 4    BY 1;
            nre03_eq =  RANNOR(31415927);
            nre10_eq =  RANNOR(31415927);  *Subsequent seeds are;
            nre30_eq =  RANNOR(31415927);  *ignored.;
            nre90_eq =  RANNOR(31415927);
            OUTPUT nor_err;
         END;
      END;
   END;

*Force set of normal errors for each randomization to have
  zero mean  and  specified variance.;
PROC STANDARD   DATA = nor_err    OUT = nor_err
     MEAN = 0.0   STD=1.7320508076      VARDEF=N;
   BY exp_rand block;
   VAR nre03_eq;
PROC STANDARD   DATA = nor_err    OUT = nor_err
     MEAN = 0.0   STD=3.1622776602      VARDEF=N;
   BY exp_rand block;
   VAR nre10_eq;
PROC STANDARD   DATA = nor_err    OUT = nor_err
     MEAN = 0.0   STD=5.4772255751      VARDEF=N;
   BY exp_rand block;
   VAR nre30_eq;
PROC STANDARD   DATA = nor_err    OUT = nor_err
     MEAN = 0.0   STD=9.4868329805      VARDEF=N;
   BY exp_rand block;
   VAR nre90_eq;

*   Debug Print options follow: ;
     %db(&dbps4, TITLE1 "Observations from nor_err eq var only(dbps4)";
```

```
                    PROC PRINT DATA=nor_err (OBS=32);
                    PROC PRINT DATA=nor_err (FIRSTOBS=4577);
                    PROC MEANS DATA=nor_err (OBS=32)
                       MEAN VAR STD  VARDEF=N;
                       BY  exp_rand  block;
                       VAR nre03_eq--nre90_eq;
                    PROC MEANS DATA=nor_err (OBS=32)
                       MEAN VAR STD  VARDEF=N;
                       BY exp_rand;
                       VAR nre03_eq--nre90_eq;
                    );

     *For not equal variances:
       Obtain total sum of squares for equal variances, over all blocks.
       Divide into var ratio + 1 parts (4+1).
       Obtain new variance, divide by denominator = number of observations.
       Multiply block 1 by 1 part and block 2 by var ratio parts (4).
       Divide by existing block variance.
       Square root the above.
       Multiply be the existing normal error.
        nre##_ne = SQRT( (SS for 8 obs)/(var ratio + 1) / (n obs 1 block)
                     * ( (1 or var ratio)  / (initial variance) ) )
                 * (current error term)
       Experiment-wide variance will be the initial variance.
       ;
     DATA nor_err;   SET nor_err;
        IF block = 1  THEN DO;
           nre03_ne = SQRT( ( 3 *4*2)/(4+1)/4 * ( 1/ 3) )  * nre03_eq;
           nre10_ne = SQRT( (10 *4*2)/(4+1)/4 * ( 1/10) )  * nre10_eq;
           nre30_ne = SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) )  * nre30_eq;
           nre90_ne = SQRT( (90 *4*2)/(4+1)/4 * ( 1/90) )  * nre90_eq;
        END;
        IF block = 2  THEN DO;
           nre03_ne = SQRT( ( 3 *4*2)/(4+1)/4 * ( 4/ 3) )  * nre03_eq;
           nre10_ne = SQRT( (10 *4*2)/(4+1)/4 * ( 4/10) )  * nre10_eq;
           nre30_ne = SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) )  * nre30_eq;
           nre90_ne = SQRT( (90 *4*2)/(4+1)/4 * ( 4/90) )  * nre90_eq;
        END;


     *    Debug Print options follow: ;
         %db(&dbps5, TITLE1 "Observations from nor_err eq & ne var (dbps5)";
                    PROC PRINT DATA=nor_err (OBS=32);
                    PROC PRINT DATA=nor_err (FIRSTOBS=4577);
                    PROC MEANS DATA=nor_err (OBS=32)
                       MEAN VAR STD  VARDEF=N;
                       BY  exp_rand  block;
                       VAR nre03_eq--nre90_ne;
                    PROC MEANS DATA=nor_err (OBS=32)
                       MEAN VAR STD  VARDEF=N;
                       BY exp_rand;
                       VAR nre03_eq--nre90_ne;
                    );


     DATA sasdata.nor_err;    SET nor_err;


      End of successful step.  */

     ***************************************************************;
     *Step Three:
       Merge eu, euti and normally distributed errors.;

      /*  *Successful step turned off. ;
```

```
DATA all_err;       *ALL three ERRor types;
  MERGE sasdata.rm_err    sasdata.nor_err;      BY exp_rand;
  covar = z;      drop z ;
  LABEL
        exp_rand= 'experiment-wise rando'
        b1_rand = 'block 1 randomization'
        b2_rand = 'block 2 randomization'

        block    = 'block id'
        b_val    = 'block effect value'
        plot     = 'experimental unit id'
        plabel   = 'plot label'
        treat    = 'treatment id'
        trt_val  = 'treat effect value'
        tlabel   = 'treatment label'
        covar    = 'covariate (z)'
        beta      = 'covariate slope coef.'

        eu_eq    = 'e.u. error var-eq'
        eu_ne    = 'e.u. error var-ne'
        euti_eq  = 'e.u.t.i. error var-eq'
        euti_ne  = 'e.u.t.i. error var-ne'

        nre03_eq = 'normal err v=3 -eq'
        nre03_ne = 'normal err v=3 -ne'
        nre10_eq = 'normal err v=10 -eq'
        nre10_ne = 'normal err v=10 -ne'
        nre30_eq = 'normal err v=30 -eq'
        nre30_ne = 'normal err v=30 -ne'
        nre90_eq = 'normal err v=90 -eq'
        nre90_ne = 'normal err v=90 -ne'

        s_popynt= 'superpop. no treat'
        s_popywt= 'superpop. with treat'
        ;

*    Debug Print options follow: ;
     %db(&dbps6, TITLE1 "Observations from all_err (eu & nor) (dbps6)";
               PROC PRINT DATA=all_err (OBS=32);
               PROC PRINT DATA=all_err (FIRSTOBS=4577);
               PROC MEANS DATA=all_err (OBS=32)
                    N MEAN VAR STD  VARDEF=N;
                 BY  exp_rand  block;
                 VAR eu_eq euti_eq  eu_ne euti_ne nre03_eq--nre90_ne;
               PROC MEANS DATA=all_err (OBS=32)
                    N MEAN VAR STD  VARDEF=N;
                 BY exp_rand;
                 VAR eu_eq euti_eq  eu_ne euti_ne nre03_eq--nre90_ne;
               PROC MEANS DATA=all_err
                    N MEAN VAR STD  VARDEF=N;
                 VAR eu_eq euti_eq  eu_ne euti_ne nre03_eq--nre90_ne;
               );

DATA sasdata.all_err;    Set all_err; *ALL three ERRor types;

  End of successful step.  */

****************************************************************;
*Step Four:
  Combine (sum) three types of error terms in various combinations. ;

 /*  *Successful step turned off. ;


  *Within exp-randomization variances are not standardized.
```

```
        Errors are divided by SQRT(num of summed terms) to bring var of
           sum of one to three error terms within same range.;

     *The error terms are E01 - E81.
      The 81 are {euti (3)} x {normal err (1+ (4*2))} x {eu (3)} = 81.
      The 3 are: O, block var eq, and block var ne (eu and euti).
      The 1+4 are var=0, 5, 10, 30, 90. The 2 are block var eq and ne.
      Each error term may be one to three of: e.u., e.u.-t.i.,normal err.;

    *Data set COMBined ERRors  one, two, or three types summed;
    DATA       comb_err;   SET sasdata.all_err;

     *Obtain divisor for sums;
      RETAIN invsq2 invsq3;
      IF _N_ = 1 THEN DO;   invsq2=1/SQRT(2); invsq3=1/SQRT(3); END;

    ****************** Create error terms ************************;
    *Var=0            normal      euti         eu ;
      e01 =       0* ( 0    +     0       +    0      );
      e02 =       1* ( 0    +     0       +    eu_eq  );
      e03 =       1* ( 0    +     0.      +    eu_ne  );
      e04 =       1* ( 0    +     euti_eq +    0      );
      e05 = invsq2* ( 0    +     euti_eq +    eu_eq  );
      e06 = invsq2* ( 0    +     euti_eq +    eu_ne  );
      e07 =       1* ( 0    +     euti_ne +    0      );
      e08 = invsq2* ( 0    +     euti_ne +    eu_eq  );
      e09 = invsq2* ( 0    +     euti_ne +    eu_ne  );

    *Var=3            normal      euti         eu ;
      e10 =       1* (nre03_eq +    0     +   0      );
      e11 = invsq2* (nre03_eq +    0     +   eu_eq );
      e12 = invsq2* (nre03_eq +    0     +   eu_ne );
      e13 = invsq2* (nre03_eq + euti_eq +   0      );
      e14 = invsq3* (nre03_eq + euti_eq +   eu_eq );
      e15 = invsq3* (nre03_eq + euti_eq +   eu_ne );
      e16 = invsq2* (nre03_eq + euti_ne +   0      );
      e17 = invsq3* (nre03_eq + euti_ne +   eu_eq );
      e18 = invsq3* (nre03_eq + euti_ne +   eu_ne );

    *                 normal      euti         eu ;
      e19 =       1* (nre03_ne + 0      +    0      );
      e20 = invsq2* (nre03_ne + 0      +    eu_eq );
      e21 = invsq2* (nre03_ne + 0      +    eu_ne );
      e22 = invsq2* (nre03_ne + euti_eq +  0      );
      e23 = invsq3* (nre03_ne + euti_eq +  eu_eq );
      e24 = invsq3* (nre03_ne + euti_eq +  eu_ne );
      e25 = invsq2* (nre03_ne + euti_ne +  0      );
      e26 = invsq3* (nre03_ne + euti_ne +  eu_eq );
      e27 = invsq3* (nre03_ne + euti_ne +  eu_ne );

    *Var=10           normal      euti         eu ;
      e28 =       1* (nre10_eq + 0      +    0      );
      e29 = invsq2* (nre10_eq + 0      +    eu_eq );
      e30 = invsq2* (nre10_eq + 0      +    eu_ne );
      e31 = invsq2* (nre10_eq + euti_eq +  0      );
      e32 = invsq3* (nre10_eq + euti_eq +  eu_eq );
      e33 = invsq3* (nre10_eq + euti_eq +  eu_ne );
      e34 = invsq2* (nre10_eq + euti_ne +  0      );
      e35 = invsq3* (nre10_eq + euti_ne +  eu_eq );
      e36 = invsq3* (nre10_eq + euti_ne +  eu_ne );

    *                 normal      euti         eu ;
      e37 =       1* (nre10_ne + 0      +    0      );
      e38 = invsq2* (nre10_ne + 0      +    eu_eq );
      e39 = invsq2* (nre10_ne + 0      +    eu_ne );
      e40 = invsq2* (nre10_ne + euti_eq +  0      );
```

```
    e41 = invsq3* (nre10_ne + euti_eq + eu_eq  );
    e42 = invsq3* (nre10_ne + euti_eq + eu_ne  );
    e43 = invsq2* (nre10_ne + euti_ne +  O     );
    e44 = invsq3* (nre10_ne + euti_ne + eu_eq  );
    e45 = invsq3* (nre10_ne + euti_ne + eu_ne  );

*Var=30          normal     euti        eu ;
    e46 =        1* (nre30_eq + O      +  O     );
    e47 = invsq2* (nre30_eq + O        + eu_eq  );
    e48 = invsq2* (nre30_eq + O        + eu_ne  );
    e49 = invsq2* (nre30_eq + euti_eq +  O     );
    e50 = invsq3* (nre30_eq + euti_eq + eu_eq  );
    e51 = invsq3* (nre30_eq + euti_eq + eu_ne  );
    e52 = invsq2* (nre30_eq + euti_ne +  O     );
    e53 = invsq3* (nre30_eq + euti_ne + eu_eq  );
    e54 = invsq3* (nre30_eq + euti_ne + eu_ne  );

*                normal     euti        eu ;
    e55 =        1* (nre30_ne + O      +  O     );
    e56 = invsq2* (nre30_ne + O        + eu_eq  );
    e57 = invsq2* (nre30_ne + O        + eu_ne  );
    e58 = invsq2* (nre30_ne + euti_eq +  O     );
    e59 = invsq3* (nre30_ne + euti_eq + eu_eq  );
    e60 = invsq3* (nre30_ne + euti_eq + eu_ne  );
    e61 = invsq2* (nre30_ne + euti_ne +  O     );
    e62 = invsq3* (nre30_ne + euti_ne + eu_eq  );
    e63 = invsq3* (nre30_ne + euti_ne + eu_ne  );

*Var=90          normal     euti        eu ;
    e64 =        1* (nre90_eq + O      +  O     );
    e65 = invsq2* (nre90_eq + O        + eu_eq  );
    e66 = invsq2* (nre90_eq + O        + eu_ne  );
    e67 = invsq2* (nre90_eq + euti_eq + O      );
    e68 = invsq3* (nre90_eq + euti_eq + eu_eq  );
    e69 = invsq3* (nre90_eq + euti_eq + eu_ne );
    e70 = invsq2* (nre90_eq + euti_ne + O      );
    e71 = invsq3* (nre90_eq + euti_ne + eu_eq );
    e72 = invsq3* (nre90_eq + euti_ne + eu_ne );

*                normal     euti        eu ;
    e73 =        1* (nre90_ne + O      +  O     );
    e74 = invsq2* (nre90_ne + O        + eu_eq  );
   ·e75 = invsq2* (nre90_ne + O        + eu_ne  );
    e76 = invsq2* (nre90_ne + euti_eq + O      );
    e77 = invsq3* (nre90_ne + euti_eq + eu_eq );
    e78 = invsq3* (nre90_ne + euti_eq + eu_ne );
    e79 = invsq2* (nre90_ne + euti_ne + O      );
    e80 = invsq3* (nre90_ne + euti_ne + eu_eq );
    e81 = invsq3* (nre90_ne + euti_ne + eu_ne );


*    Debug Print options follow: ;
     %db(&dbps7,
     TITLE1 "Observations from comb_err before standardizing (dbps7)";
                PROC PRINT DATA=comb_err (OBS=16);
                PROC PRINT DATA=comb_err (FIRSTOBS=4593);
                PROC MEANS DATA=comb_err (OBS=16)
                       N MEAN VAR STD  MIN MAX  VARDEF=N;
                   BY exp_rand block;
                   VAR e01-e81;
                PROC MEANS DATA=comb_err (OBS=16)
                       N MEAN VAR STD  MIN MAX  VARDEF=N;
                   BY exp_rand;
                   VAR e01-e81;
                PROC MEANS DATA=comb_err
                       N MEAN VAR STD  MIN MAX  VARDEF=N;
```

```
                        VAR e01-e81;
                  );

DATA  sasdata.comb_err;      SET   comb_err;

 End of successful step.   */

********************************************************************;
*Step Five:
  Standardize the sumed error terms;

 /*  *Successful step turned off. ;


*All combined terms are standardizd to mean zero and experiment-wide
    variance of 30, the same values as single-source errors.
   When all combined terms have {equal - not equal} within-block
    variances, then their sum has {equal - not equal} within-block
    variances.
   When the combined terms are a mix of eq and ne within-block
    variance terms, they are treated as if all terms had no equal
    within-block variances.  They are standardized within each block
    to have the usual ratio of block variances.

   Thus, for comparisons, all single-source error or combined error
    term have zero mean and the same experiment-wide variance.
  ;


*Data set: STanDardize combined ERRors;
PROC STANDARD    DATA = sasdata.comb_err         OUT = std_err
        MEAN = 0.0    STD=5.4772255751      VARDEF=N;
   BY exp_rand block;
   VAR e01-e81;

*    Debug Print options follow: ;
     %db(&dbps8,
     TITLE1 "Observations from std_err - 1st standardizing (dbps8)";
               PROC PRINT DATA=std_err (OBS=16);
               PROC PRINT DATA=std_err (FIRSTOBS=4593);
               PROC MEANS DATA=std_err (OBS=16)
                     N MEAN VAR STD  MIN MAX  VARDEF=N;
                  BY exp_rand block;
                  VAR e01-e81;
               PROC MEANS DATA=std_err (OBS=16)
                     N MEAN VAR STD  MIN MAX  VARDEF=N;
                  BY exp_rand;
                  VAR e01-e81;
               PROC MEANS DATA=std_err
                     N MEAN VAR STD  MIN MAX  VARDEF=N;
                  VAR e01-e81;
               );

DATA std_err;    SET std_err;
  *Set all sums involving a term with unequal block variances to
   also have the standard (1:4) unequal block variance. ;

   IF block = 1  THEN DO;
     *Var=00;
        e03s  = e03  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
        e06s  = e06  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
        e07s  = e07  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
        e08s  = e08  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
        e09s  = e09  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );

     *Var=03;
        e12s  = e12  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
```

```
   e15s  = e15  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e16s  = e16  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e17s  = e17  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e18s  = e18  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );

   e19s  = e19  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e20s  = e20  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e21s  = e21  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e22s  = e22  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e23s  = e23  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e24s  = e24  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e25s  = e25  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e26s  = e26  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e27s  = e27  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );

*Var=10;
   e30s  = e30  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e33s  = e33  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e34s  = e34  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e35s  = e35  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e36s  = e36  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );

   e37s  = e37  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e38s  = e38  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e39s  = e39  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e40s  = e40  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e41s  = e41  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e42s  = e42  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e43s  = e43  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e44s  = e44  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e45s  = e45  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );

*Var=30;
   e48s  = e48  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e51s  = e51  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e52s  = e52  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e53s  = e53  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e54s  = e54  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );

   e55s  = e55  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e56s  = e56  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e57s  = e57  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e58s  = e58  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e59s  = e59  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e60s  = e60  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e61s  = e61  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e62s  = e62  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e63s  = e63  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );

*Var=90;
   e66s  = e66  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e69s  = e69  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e70s  = e70  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e71s  = e71  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e72s  = e72  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );

   e73s  = e73  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e74s  = e74  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e75s  = e75  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e76s  = e76  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e77s  = e77  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e78s  = e78  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e79s  = e79  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e80s  = e80  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
   e81s  = e81  * SQRT( (30 *4*2)/(4+1)/4 * ( 1/30) );
END;
```

```
IF block = 2  THEN DO;
 *Var=00;
   e03s  = e03  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e06s  = e06  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e07s  = e07  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e08s  = e08  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e09s  = e09  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );

 *Var=03;
   e12s  = e12  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e15s  = e15  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e16s  = e16  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e17s  = e17  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e18s  = e18  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );

   e19s  = e19  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e20s  = e20  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e21s  = e21  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e22s  = e22  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e23s  = e23  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e24s  = e24  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e25s  = e25  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e26s  = e26  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e27s  = e27  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );

 *Var=10;
   e30s  = e30  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e33s  = e33  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e34s  = e34  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e35s  = e35  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e36s  = e36  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );

   e37s  = e37  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e38s  = e38  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e39s  = e39  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e40s  = e40  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e41s  = e41  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e42s  = e42  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e43s  = e43  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e44s  = e44  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e45s  = e45  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );

 *Var=30;
   e48s  = e48  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e51s  = e51  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e52s  = e52  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e53s  = e53  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e54s  = e54  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );

   e55s  = e55  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e56s  = e56  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e57s  = e57  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e58s  = e58  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e59s  = e59  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e60s  = e60  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e61s  = e61  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e62s  = e62  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e63s  = e63  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );

 *Var=90;
   e66s  = e66  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e69s  = e69  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e70s  = e70  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e71s  = e71  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
   e72s  = e72  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
```

```
        e73s  = e73  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
        e74s  = e74  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
        e75s  = e75  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
        e76s  = e76  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
        e77s  = e77  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
        e78s  = e78  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
        e79s  = e79  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
        e80s  = e80  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );
        e81s  = e81  * SQRT( (30 *4*2)/(4+1)/4 * ( 4/30) );

    END;



*    Debug Print options follow: ;
    %db(&dbps9,
    TITLE1 "Observations from std_err - 2nd standardizing (dbps9)";
            PROC PRINT DATA=std_err (OBS=16);
            PROC PRINT DATA=std_err (FIRSTOBS=4593);
            PROC MEANS DATA=std_err (OBS=16)
                    N MEAN VAR STD  MIN MAX  VARDEF=N;
                BY exp_rand block;
                VAR e03s--e81s;
            PROC MEANS DATA=std_err (OBS=16)
                    N MEAN VAR STD  MIN MAX  VARDEF=N;
                BY exp_rand;
                VAR e03s--e81s;
            PROC MEANS DATA=std_err
                    N MEAN VAR STD  MIN MAX  VARDEF=N;
                VAR e03s--e81s;
            );

    DATA  sasdata.std_err;   SET std_err;
        DROP  e01-e81  invsq2 invsq3;

 End of successful step.  */

*************************************************************************;
*Step Six:
  Compute the observed y's for each of the single and combined error
  terms;


 /*  *Successful step turned off. ;

*Create various observed y's.
  Each is an error added to:  mu + block + beta*covariate + one of
    (zero or the appropriate treatment effect).
  yobs001 - yobs081 have treatments eq 0.
  yobs101 - yobs181 have treatments ne 0.
  The 1 inidicates the first (and only) nonzero treatment vector.
  yobs###s use an error term with one or more terms originally having
    unequal block variances.  The sum has been standardized to have
    the usual unequal block variances (1:4 ratio).;


**** Create macro too generate formula for ***************************;
************* the observed population values yobs(#)(##)  ***********;
    %MACRO  m_yobs;
        %DO  err_id = 1 %TO 9 ;
            yobs00&err_id =  s_popynt +  e0&err_id;
            yobs10&err_id =  s_popywt +  e0&err_id;
        %END;
        %DO  err_id = 10 %TO 81;
```

```
            yobs0&err_id =  s_popynt +  e&err_id;
            yobs1&err_id =  s_popywt +  e&err_id;
        %END;
    %MEND;
************ End of macro m_yobs ********************************;

DATA  obs_y;
    merge sasdata.comb_err   sasdata.std_err;
 *Data set comb_err attaches original e##.;
 *Data set std_err attaches the standardized e##s.;


 *Call macro  m_yobs to generate code creating the observed y values
  for the various non-standardized sums of error terms.;
      %m_yobs;

*Compute the observed y values for the various standardized sums of
 error terms. ;
  *All treatment values are zero. ;
    *Var=00;
        yobs003s = e03 + s_popynt ;
        yobs006s = e06 + s_popynt ;
        yobs007s = e07 + s_popynt ;
        yobs008s = e08 + s_popynt ;
        yobs009s = e09 + s_popynt ;

    *Var=03;
        yobs012s = e12 + s_popynt ;
        yobs015s = e15 + s_popynt ;
        yobs016s = e16 + s_popynt ;
        yobs017s = e17 + s_popynt ;
        yobs018s = e18 + s_popynt ;

        yobs019s = e19 + s_popynt ;
        yobs020s = e20 + s_popynt ;
        yobs021s = e21 + s_popynt ;
        yobs022s = e22 + s_popynt ;
        yobs023s = e23 + s_popynt ;
        yobs024s = e24 + s_popynt ;
        yobs025s = e25 + s_popynt ;
        yobs026s = e26 + s_popynt ;
        yobs027s = e27 + s_popynt ;

    *Var=10;
        yobs030s = e30 + s_popynt ;
        yobs033s = e33 + s_popynt ;
        yobs034s = e34 + s_popynt ;
        yobs035s = e35 + s_popynt ;
        yobs036s = e36 + s_popynt ;

        yobs037s = e37 + s_popynt ;
        yobs038s = e38 + s_popynt ;
        yobs039s = e39 + s_popynt ;
        yobs040s = e40 + s_popynt ;
        yobs041s = e41 + s_popynt ;
        yobs042s = e42 + s_popynt ;
        yobs043s = e43 + s_popynt ;
        yobs044s = e44 + s_popynt ;
        yobs045s = e45 + s_popynt ;

    *Var=30;
        yobs048s = e48 + s_popynt ;
        yobs051s = e51 + s_popynt ;
        yobs052s = e52 + s_popynt ;
        yobs053s = e53 + s_popynt ;
        yobs054s = e54 + s_popynt ;
```

```
        yobs055s = e55 + s_popynt ;
        yobs056s = e56 + s_popynt ;
        yobs057s = e57 + s_popynt ;
        yobs058s = e58 + s_popynt ;
        yobs059s = e59 + s_popynt ;
        yobs060s = e60 + s_popynt ;
        yobs061s = e61 + s_popynt ;
        yobs062s = e62 + s_popynt ;
        yobs063s = e63 + s_popynt ;

    *Var=90;
        yobs066s = e66 + s_popynt ;
        yobs069s = e69 + s_popynt ;
        yobs070s = e70 + s_popynt ;
        yobs071s = e71 + s_popynt ;
        yobs072s = e72 + s_popynt ;

        yobs073s = e73 + s_popynt ;
        yobs074s = e74 + s_popynt ;
        yobs075s = e75 + s_popynt ;
        yobs076s = e76 + s_popynt ;
        yobs077s = e77 + s_popynt ;
        yobs078s = e78 + s_popynt ;
        yobs079s = e79 + s_popynt ;
        yobs080s = e80 + s_popynt ;
        yobs081s = e81 + s_popynt ;


*Treatment values are nonzero. ;
    *Var=00;
        yobs103s = e03 + s_popywt ;
        yobs106s = e06 + s_popywt ;
        yobs107s = e07 + s_popywt ;
        yobs108s = e08 + s_popywt ;
        yobs109s = e09 + s_popywt ;

    *Var=03;
        yobs112s = e12 + s_popywt ;
        yobs115s = e15 + s_popywt ;
        yobs116s = e16 + s_popywt ;
        yobs117s = e17 + s_popywt ;
        yobs118s = e18 + s_popywt ;

        yobs119s = e19 + s_popywt ;
        yobs120s = e20 + s_popywt ;
        yobs121s = e21 + s_popywt ;
        yobs122s = e22 + s_popywt ;
        yobs123s = e23 + s_popywt ;
        yobs124s = e24 + s_popywt ;
        yobs125s = e25 + s_popywt ;
        yobs126s = e26 + s_popywt ;
        yobs127s = e27 + s_popywt ;

    *Var=10;
        yobs130s = e30 + s_popywt ;
        yobs133s = e33 + s_popywt ;
        yobs134s = e34 + s_popywt ;
        yobs135s = e35 + s_popywt ;
        yobs136s = e36 + s_popywt ;

        yobs137s = e37 + s_popywt ;
        yobs138s = e38 + s_popywt ;
        yobs139s = e39 + s_popywt ;
        yobs140s = e40 + s_popywt ;
        yobs141s = e41 + s_popywt ;
```

```
        yobs142s = e42 + s_popywt ;
        yobs143s = e43 + s_popywt ;
        yobs144s = e44 + s_popywt ;
        yobs145s = e45 + s_popywt ;

    *Var=30;
        yobs148s = e48 + s_popywt ;
        yobs151s = e51 + s_popywt ;
        yobs152s = e52 + s_popywt ;
        yobs153s = e53 + s_popywt ;
        yobs154s = e54 + s_popywt ;

        yobs155s = e55 + s_popywt ;
        yobs156s = e56 + s_popywt ;
        yobs157s = e57 + s_popywt ;
        yobs158s = e58 + s_popywt ;
        yobs159s = e59 + s_popywt ;
        yobs160s = e60 + s_popywt ;
        yobs161s = e61 + s_popywt ;
        yobs162s = e62 + s_popywt ;
        yobs163s = e63 + s_popywt ;

    *Var=90;
        yobs166s = e66 + s_popywt ;
        yobs169s = e69 + s_popywt ;
        yobs170s = e70 + s_popywt ;
        yobs171s = e71 + s_popywt ;
        yobs172s = e72 + s_popywt ;

        yobs173s = e73 + s_popywt ;
        yobs174s = e74 + s_popywt ;
        yobs175s = e75 + s_popywt ;
        yobs176s = e76 + s_popywt ;
        yobs177s = e77 + s_popywt ;
        yobs178s = e78 + s_popywt ;
        yobs179s = e79 + s_popywt ;
        yobs180s = e80 + s_popywt ;
        yobs181s = e81 + s_popywt ;


*   Debug Print options follow: ;
    %db(&dbps10,
    TITLE1 "Observations from obs_y  (dbps10)";
                PROC PRINT DATA=obs_y (OBS=16);
                PROC PRINT DATA=obs_y (FIRSTOBS=4593);
                PROC MEANS DATA=obs_y (OBS=16)
                        N MEAN VAR STD  MIN MAX   VARDEF=N;
                    BY exp_rand block;
                    VAR yobs001--yobs181s;
                PROC MEANS DATA=obs_y (OBS=16)
                        N MEAN VAR STD  MIN MAX   VARDEF=N;
                    BY exp_rand;
                    VAR yobs001--yobs181s;
                PROC MEANS DATA=obs_y
                        N MEAN VAR STD  MIN MAX   VARDEF=N;
                    VAR yobs001--yobs181s;
                );


DATA   sasdata.obs_y;    SET obs_y;

 End of successful step.   */

*****************************************************************;
*Step Seven:
```

Add the variable g_roup to the file of observed y's. ;


```
  DATA   sasdata.obs_y;
     MERGE   sasdata.obs_y    sasdata.gr_index;
     BY   exp_rand;
 PROC PRINT  DATA = sasdata.obs_y  (OBS=16);
     VAR exp_rand--beta   s_popynt s_popywt covar g_group;
 PROC PRINT  DATA = sasdata.obs_y  (FIRSTOBS=4593);
     VAR exp_rand--beta   s_popynt s_popywt covar g_group;
//
```

```
U13293A.SASPROG.CNTL(GLM)
VPSPRINT 5.1.002   WEDNESDAY NOVEMBER 4,1992   14:59:09 U13293A   MVS1        **
VPSPRINT SASPROG.CNTL(*) LOCAL FORM(9001) CLASS(E)

//U13293AA  JOB (*),'ROBERT WILSON',
//  TIME=(1,0),MSGLEVEL=(1,1),MSGCLASS=X,CLASS=2
/*JOBPARM  ROOM=Y,FORMS=9001,COPIES=1
//S1  EXEC  SAS,OPTIONS='MACRO,DQUOTE'
//SASDATA  DD  DSN=U13293A.SASDATA.DATA,DISP=OLD
//SASGLM   DD  DSN=U13293A.SASGLM2.DATA,DISP=OLD
//FT22F001 DD  DSN=U13293A.GLMOUT.DATA,DISP=(NEW,KEEP),
//     SPACE=(TRK,(800,400)),UNIT=STORAGE
//SYSIN  DD *

 *
 * Program File:  sasprog.cntl(glm)
 *               Run General Linear Models on all randomizations.
 *
 *    //FT22F001  DD  DSN=U13293A.GLMOUT56.DATA,DISP=OLD
 *
 * Step One:
 *     Enter the treatment/no treatment code, the error id, and
 *        s (if errors are standardized) into the macro let statements.
 *
 * Step Two:
 *    Run General Linear Models on all randomizations.
 *    Loop (via BY exp_rando) through all experimental randomizations.
 *    For each, perform PROC GLM, obtaining ss for usual estimates.
 *    Write usual printed output to scratch file.
 *    Save predicted y_hats, residuals, & incoming data in sas file.
 *        Only selected variables from sasdata.obs_y are kept.
 *
 * Step Three:
 *    Merge g_group labels to file of y_hats, residuals, & input data.
 *    Label new variables in this file.
 *    Add g_group variable to this file.
 *
 * Step Four:
 *    Read file of usually printed GLM output, selecting variables.
 *    Save in sas file.
 *
 * Step Five:
 *    Add g_group variable to this file.
 *    Add usual constraints estimators.
 *    Compute additional sums of squares, ms, F-ratios, & osls.
 *    Begin computation of ratios 1 - 3.
 *
 * Step Six:
 *    Obtain means over all randomizations.
 *
 * Step Seven:
 *    Complete computation of ratios 1 - 3.
 *    Compute two types of errors.
 *    Label for printing and for permanent file.
 *    Print initial observations.
 *
 * Step Eight:
 *   Analysis - means and graphic display.
 *
 * Output Files   --  One pair per error type:
 *   sasglm.&yhat_ds
 *     One record per experimental randomization per experimentsl unit.
 *   sasglm.&glm_ds
```

```
*     One record per experimental randomization
*


*************** Debug Print Switch Macro and Switches    **********;
* Macro for debug printing follows:;
    %MACRO  db(db_site,db_act);
       %IF %UPCASE(&db_site) = TRUE  | %UPCASE(&db_site) = T
           %THEN   %QUOTE(&db_act);
    %MEND db;
*
*Debug print switches (dbps#):
 True or t = on (in upper and/or lower case), anything else = off.;
          %LET dbps1 = f     ; *Printed output & yhat_ds from test glm;
          %LET dbps2 = f     ; *Y hat output from glm;
          %LET dbps3 = f     ; *Read anova output from glm;
          %LET dbps4 = f     ; *;

********************** End of Debug Print Switches ***************;

******************************************************************;
*Step One:
 Select set of observed y's for analysis.
          Enter ### or ###s into function %STR(...), with ### equalling
            a 0 for no treatment effects or a 1 for treatments, then
            two digits indicating the error combination.
            Follow ### with an "s" when using the standardized errors.;

          %LET yobs    = %STR(yobs155);    *0 or 1, then below message;
          %LET error   =    %STR(e55);    *00 to 81 with or w/o an s;
          %LET ypred   = %STR(ypre155);
          %LET resid   = %STR(resi155);
          %LET yhat_ds = %STR(yhat155);   *Output data set of obser.;
          %LET glm_ds  =  %STR(glm155);   *Output data set of anova terms;
          TITLE1  "Observed Y uses Error Combination  &yobs ";

*Output is stored in sas library  sasglm.data under the member name
     yhat_ds  and  glm_ds  .
     &yhat_ds     Has  1 record per y_hat by exp_rand = {T!**B}X{TB}.
     &glm_ds      Has  1 record per  anova table  = {T!**B}.

******************************************************************;
*Step Two:
   Use PROC GLM to compute the analysis of variance.
   Send GLM output to scratch file, then read it, creating a sas file.
   Augment data set of individual responses with the value of the
        predicted y, y_hat and the residual values.;

*     Debug Print options follow: ;
     %db(&dbps1,
          TITLE1 "GLM anova output for &yobs  (dbps1)";
          OPTIONS  NONOTES;
          PROC PRINT DATA = sasdata.obs_y
            (OBS=16  KEEP=exp_rand--beta s_popynt s_popywt  covar
                 &yobs    &error );
          PROC GLM  DATA = sasdata.obs_y
            (OBS=16  KEEP=exp_rand--beta s_popynt s_popywt  covar
                 &yobs    &error );
          BY exp_rand;
          CLASS   block treat;
          MODEL   &yobs = block treat covar / SS1 SS4  SOLUTION;
          MEANS   block  treat;
          OUTPUT  OUT = yhat_ds
                  P= &ypred   R= &resid ;
          PROC PRINT DATA = yhat_ds  (OBS=16);
          OPTIONS  NOTES;
       );
```

```
   PROC PRINTTO  UNIT=22 NEW; *Sends output from PROC GLM to a temp file.;
   PROC GLM  DATA = sasdata.obs_y
        (         KEEP=exp_rand--beta s_popynt s_popywt  covar
                  &yobs    &error );
      BY exp_rand;
      CLASS   block treat;
      MODEL   &yobs = block treat covar / SS1 SS4  SOLUTION;
      MEANS   block  treat;
      OUTPUT  OUT = &yhat_ds
           P= &ypred   R= &resid ;
   OPTIONS  NONOTES;

   PROC PRINTTO;     *Returns output to the line printer.;
   OPTIONS  NOTES;


   *********************************************************************;

   *Step Three:
     Label new variables  and add variable g_group;

   PROC SORT  DATA= sasdata.gr_index;  BY exp_rand;
   DATA   sasglm.&yhat_ds;
      MERGE  &yhat_ds  sasdata.gr_index;     BY exp_rand;
      LABEL exp_rand = 'Experimental Randomization'
           g_group  = 'Eq. projector group'
           &ypred   = 'Predicted y'
           &resid   = 'Residual'
               ;
   *    Debug Print options follow: ;
        %db(&dbps2, TITLE1 "Observations from glm output data  (dbps2)";
           PROC PRINT DATA = sasglm.&yhat_ds (OBS=16) ;
                   );



   *********************************************************************;
   *Step Four:
     Read the GLM output from each experimantal randomization.
       Extract quantities of interest from the three sections of output:
          anova table (sums of squares), parmaeter estimates, and means.
       Add g_group and non-centrality variables.
       Compute additional variables.;

   DATA   sasglm.&glm_ds;
      DROP     location  skip;
      INFILE   FT22F001   MISSOVER      EOF = all_ran;
      LENGTH   location  $19;

   *Loop through all experimental randomizations;
   DO   exp_rand = 1 TO   576;

     *Skip through page 1 of GLM output - listing of class, levels, values;
     *Use //, not #n, to permit changes in number of title lines.
        INPUT  10n;
        DO UNTIL  (location='Dependent') ;
           INPUT location ; END;
        *PUT location= ;     *For testing;

     *Having arrived at page 2, column 1 of (blank) line following
        'Dependent Variable',skip to line beginning 'Model' and read vars.;
        INPUT //;
        INPUT    @30 dfmod ssmod msmod fmod oslmod
           //    @30 dfres ssres msres
           //    @30 dftot sstot
```

```
    ////  @20 rsq coefvar   rootmse mean_y
    ////  ;
    *Read blocks given mean, treatments given mean and blocks, and
     covariate (z) given mean (m), blocks (b), and treatments (t);
     *Type I Sum of Squares;
  INPUT      @30 dfb     ssb_m     msb_m     fb_m     oslb_m
      /      @30 dft     sst_mb    mst_mb    ft_mb    oslt_mb
      /      @30 dfz     ssz_mbt   msz_mbt   fz_mbt   oslz_mbt ;
     *Type IV Sum of Squares;
   INPUT
      ///    @30 skip    ssb_mtz   msb_mtz   fb_mtz   oslb_mtz
      /      @30 skip    sst_mbz   mst_mbz   ft_mbz   oslt_mbz ;


     *Read parameter estimates;
     DO UNTIL  (location = 'Parameter');  INPUT  location $ ;   END;
     * PUT location= ;   *For testing;
        INPUT / @22 muhatg  / @22 bl1hatg
              // @22 tr1hatg / @22 tr2hatg / @22 tr3hatg
              // @22 betahat;

  *Read means from third page of glm output;
     DO UNTIL  (location = 'Level');  INPUT location  ;   END;
     * PUT location= ;      *For testing;
        INPUT //  skip skip bl1ybar skip bl1zbar
             /  skip skip bl2ybar skip bl2zbar
         /////  skip skip tr1ybar skip tr1zbar
             /  skip skip tr2ybar skip tr2zbar
             /  skip skip tr3ybar skip tr3zbar
             /  skip skip tr4ybar skip tr4zbar;
   OUTPUT;
 *Return to the top of the loop and read glm output
  from the next experimental randomization.;
 END;
 all_ran: location=location; *Control passes to here upon end of file.;
         *label: needs a statement to activate it, thus loc=loc.;
 PUT exp_rand=;

 *   Debug Print options follow: ;
     %db(&dbps3,
     TITLE1 "Observations from  glm output,  sasdata.&glm_ds  (dbps3)";
        PROC PRINT DATA = sasglm.&glm_ds (OBS=2);
                 );

 ******************************************************************;
 *Step Five:
    Merge equal projector group id variable, g_group
    Obtain the 'usual constraints' aov estimates and add to data set.
       See Leyman Ott, Edition One, page 524 for details.
          The aov estimates = ('cards')transpose * (glm estimates).
          The covariate slope estimate is unchanged.
          The arrays mimic matrix multiplication.
    Add remaining sums of squares, ms, F-ratios, and osls to data set.
    Compute ratio1, a balancing function of the covariates.;


 DATA  temp;
    INPUT  tma1-tma5     tb1a1-tb1a5
           tt1a1-tt1a5   tt2a1-tt2a5   tt3a1-tt3a5;
 *  The sas-estimates to usual-constraints conversion "matrix"
    follows;
    CARDS;
           1.0     .5      .25     .25     .25
           0.0     .5     0.0     0.0     0.0
           0.0    0.0      .75    -.25    -.25
           0.0    0.0     -.25     .75    -.25
```

```
          0.0    0.0   -.25   -.25    .75
     ;
DATA  sasglm.&glm_ds;
   MERGE  sasglm.&glm_ds    sasdata.gr_index;    BY exp_rand;

DATA  sasglm.&glm_ds;     SET  sasglm.&glm_ds;
    IF _N_ = 1 THEN  SET temp;
    RETAIN tma1-tma5     tb1a1-tb1a5  k
           tt1a1-tt1a5  tt2a1-tt2a5  tt3a1-tt3a5;
    DROP   tma1-tma5     tb1a1-tb1a5  k
           tt1a1-tt1a5  tt2a1-tt2a5  tt3a1-tt3a5;
       ARRAY  glma  {5} muhatg bl1hatg tr1hatg tr2hatg tr3hatg;
       ARRAY  mua   {5} tma1-tma5;
       ARRAY  b1a   {5} tb1a1-tb1a5;
       ARRAY  t1a   {5} tt1a1-tt1a5;
       ARRAY  t2a   {5} tt2a1-tt2a5;
       ARRAY  t3a   {5} tt3a1-tt3a5;
       muhata = 0;  bl1hata= 0;  tr1hata = 0;  tr2hata = 0; tr3hata = 0;
    DO k = 1 TO 5 BY 1;
       muhata = muhata  + glma{k} * mua{k};
       bl1hata= bl1hata + glma{k} * b1a{k};
       tr1hata= tr1hata + glma{k} * t1a{k};
       tr2hata= tr2hata + glma{k} * t2a{k};
       tr3hata= tr3hata + glma{k} * t3a{k};
    END;

*The "usual constraint" is that each sum of estimators equals zero.;
   bl2hata = -bl1hata;
   tr4hata = -SUM(tr1hata, tr2hata, tr3hata);

*Compute remaining ss, ma, f_statistics, and olss. ;
       ssb_mt = ssb_m;
       ssb_mz = ssb_mtz;
       sst_m  = sst_mb;
       sst_mz  = sst_mbz;
       ssz_m  = ssmod - ssb_mz  - sst_mz;
       ssz_mb = ssmod - sst_mbz - ssb_m;
       ssz_mt = ssmod - ssb_mtz - sst_m;
     msb_mt =ssb_mt /dfb;
     msb_mz =ssb_mz /dfb;
     mst_m  =sst_m  /dft;
     mst_mz =sst_mz /dft;
     msz_m  =ssz_m  /dfz;
     msz_mb =ssz_mb /dfz;
     msz_mt =ssz_mt /dfz;
      fb_mt=msb_mt/msres;  fb_mz =msb_mz /msres;
      ft_m =mst_m /msres;  ft_mz =mst_mz /msres;
      fz_m =msz_m /msres;  fz_mb =msz_mb /msres;  fz_mt =msz_mt /msres;
           oslb_mt  = 1 - PROBF(fb_mt,  dfb, dftot);
           oslb_mz  = 1 - PROBF(fb_mz,  dfb, dftot);
           oslt_m   = 1 - PROBF(ft_m ,  dft, dftot);
           oslt_mz  = 1 - PROBF(ft_mz,  dft, dftot);
           oslz_m   = 1 - PROBF(fz_m,   dfz, dftot);
           oslz_mb  = 1 - PROBF(fz_mb,  dfz, dftot);
           oslz_mt  = 1 - PROBF(fz_mt,  dfz, dftot);

*Compute ratio1 :
   ratio1 = for z: adj.(mean & blocks) treatment ms/residual ss
   Ratio1 is the block equivalent to the following:
   Finney (1946) page 54, equation 3, A / A*(T-1).
   Lucas (1950).
   Greenberg (1952) page 698, T sub ZZ / (P-1)*(S sub XX).
   Cox (1956) page 1146, equation 7, rightmost pate B sub/(T-1)*R sub.
   Cox (1957) page 153,  equation 6B, rightmost part T sub/(T-1)*R sub.
   Cox (1982) page 198,  equation 3,trace(B sub Z *INV(S sub Z))/(T-1).;
```

```
*Begin computation of omega, the mean of omega_df = S sub Z / DF sub W.
     omega = expectation over all randomizations of (Z residual MS).;

   zsst_mb  = 2 * USS(tr1zbar, tr2zbar, tr3zbar, tr4zbar);
   *zss_mbt = residual sum of sq for covar given mean, blocks, treats.
            = total ss - block ss - treatment ss + correction factor.
            covariates are: 1 2 11 14   5 9 16 19;
   zss_mbt  = 1045 - 748.25 - zsst_mb + 741.125;
   IF zss_mbt < .0000001 THEN  PUT  exp_rand=  zss_mbt= ;

   ratio1   = zsst_mb / ( dft  *  zss_mbt) ;
   omega_df = zss_mbt / dfres;  * mean square residual;


 *   Debug Print options follow: ;
     %db(&dbps4,
       TITLE1 "Observations from  glm output,  sasdata.&glm_ds  (dbps4)";
         PROC PRINT DATA = sasglm.&glm_ds (OBS=2);
                 );

******************************************************************;
* Step Six:
   Obtain the randomization expectation of parameters and aov estimates
     from the output of the glm analysis.;

PROC MEANS   DATA = sasglm.&glm_ds   N MEAN STD VAR MIN MAX ;
   TITLE1 "Means, etc. from  glm output,  sasdata.&glm_ds";
   VAR   muhata  bl1hata bl2hata
         tr1hata tr2hata tr3hata tr4hata    betahat  zss_mbt omega_df
         msb_m      mst_mb      mst_mbz     msz_mb      msz_mbt
         fmod        fb_m        ft_mb       ft_mbz      fz_mb       fz_mbt
         ssb_m       sst_mb      sst_mbz     ssz_mb      ssz_mbt
         os1b_m     os1t_mb     os1t_mbz    os1z_mb     os1z_mbt
         os1mod fmod     msmod msres   sstot ssmod ssres
         rsq   coefvar   rootmse   mean_y ;
   OUTPUT   OUT=meanout
         MEAN = amuhata   abl1hata abl2hata
               atr1hata atr2hata atr3hata atr4hata
               abetahat   azss_mbt   omega
               amsb_m    amst_mb    amst_mbz    amsz_mb    amsz_mbt
               afmod      afb_m      aft_mb     aft_mbz   afz_mb afz_mbt;

******************************************************************;
* Step Seven:
   Compute two types of errors.
   Complete construction of ratio2 and ratio3. ;

DATA  sasglm.&glm_ds (DROP= omega_df   _TYPE_  _FREQ_);
   SET  sasglm.&glm_ds;
   IF _N_ = 1 THEN  SET meanout;

* The two error types subtract the estimate from:
   (Type r) The randomization expectation, the average over all
      experimental randomizations.
   (Type s) The superpopulation values as defined a priori.;

   rmu =  amuhata - muhata;
   rbl1 = abl1hata - bl1hata;        rbl2 = abl2hata - bl2hata;
   rtr1 = atr1hata - tr1hata;        rtr2 = atr2hata - tr2hata;
   rtr3 = atr3hata - tr3hata;        rtr4 = atr4hata - tr4hata;
   rbeta    = abetahat - betahat;
   rmst_mb = amst_mb - mst_mb;         rmst_mbz = amst_mbz - mst_mbz;
   rmsz_mb = amsz_mb - msz_mb;         rmsz_mbt = amsz_mbt - msz_mbt;
   rfmod    = afmod   - fmod;          rfb_m    = afb_m    - fb_m;
   rft_mb  = aft_mb - ft_mb;           rft_mbz  = aft_mbz  - ft_mbz;
   rfz_mb  = afz_mb - fz_mb;           rfz_mbt  = afz_mbt  - fz_mbt;
```

```
          smu  = 10   - muhata;              sbeta = 2.0 - betahat;
          sbl1 = -1.5 - bl1hata;             sbl2 = 1.5 - bl2hata;
          str1 =  0.0 - tr1hata;             str2 = 0.0 - tr2hata;
          str3 =  0.0 - tr3hata;             str4 = 0.0 - tr4hata;
*
*Compute ratio2,  Cox (1982) Page 200 TR(B sub Z * INV(omega sub Z))
    Z SS for treats adj. for mean & blocks / EXP(Z residual mean sq.)
*;
          ratio2 = zsst_mb / omega;
*compute ratio3,  Cox (1956) page 1148, equation 14,
*    w / randomization expectation of w.  Where w = R sub ZZ,
*    the residual sum of squares (given mean, blocks, & treatments).
*    Ratio3 is used as a weight factor in a 2nd supposedly approximately
*    unbiased expectation of the residual mean square.
*;
          ratio3 = zss_mbt / azss_mbt;
LABEL
    EXP_RAND='Exper Random.'
    B1_RAND ='Block 1 Random'  B2_RAND ='Block 2 Random'
    /* ZT11='Z BL1 TREAT 1'   ZT12='Z BL1 TREAT 2'   ZT13='Z BL1 TREAT 3'
    ZT14='Z BL1 TREAT 4'   ZT21='Z BL2 TREAT 1'   ZT22='Z BL2 TREAT 2'
    ZT23='Z BL2 TREAT 3'   ZT24='Z BL2 TREAT 4'
    */ SSMOD = 'SS MODEL'      SSRES  = 'SS Redisual'   SSTOT = 'SS TOTAL'
    MSMOD = 'MS MODEL'      MSRES  = 'MS Residual'
    FMOD  = 'F. MODEL'      OSLMOD = 'OSL MODEL'     ROOTMSE='ROOT MSE'
    DFMOD = 'D.F. MODEL'    DFRES  = 'D.F. RESIDUAL'  DFTOT ='D.F. TOTAL'
    COEFVAR ='COEF VAR'     RSQ ='R SQUARED'  MEAN_Y = 'MEAN Y RESPONSE'
  SSB_M    ='SS B_ M'      SST_M   ='SS T_ M'      SSZ_M    ='SS Z_ M'
  SSB_MT   ='SS B_ M,T'    SST_MB  ='SS T_ M,B'    SSZ_MB   ='SS Z_ M,B'
  SSB_MZ   ='SS B_ M,Z'    SST_MZ  ='SS T_ M,Z'    SSZ_MT   ='SS Z_ M,T'
  SSB_MTZ  ='SS B_ M,T,Z'  SST_MBZ ='SS T_ M,B,Z'  SSZ_MBT  ='SS Z_ M,B,T'
    MSB_M    ='MS B_ M'      MST_M   ='MS T_ M'      MSZ_M    ='MS Z_ M'
    MSB_MT   ='MS B_ M,T'    MST_MB  ='MS T_ M,B'    MSZ_MB   ='MS Z_ M,B'
    MSB_MZ   ='MS B_ M,Z'    MST_MZ  ='MS T_ M,Z'    MSZ_MT   ='MS Z_ M,T'
    MSB_MTZ  ='MS B_ M,T,Z'  MST_MBZ ='MS T_ M,B,Z' MSZ_MBT  ='MS Z_ M,B,T'
  FB_M   = 'F B_ M'      FT_M   = 'F T_ M'      FZ_M   = 'F Z_ M'
  FB_MT  = 'F B_ M,T'    FT_MB  = 'F T_ M,B'    FZ_MB  = 'F Z_ M,B'
  FB_MZ  = 'F B_ M,Z'    FT_MZ  = 'F T_ M,Z'    FZ_MT  = 'F Z_ M,T'
  FB_MTZ = 'F B_ M,T,Z'  FT_MBZ = 'F T_ M,B,Z'  FZ_MBT = 'F Z_ M,B,T'
    OSLB_M   ='OSL B_ M'      OSLT_M   ='OSL T_ M'      OSLZ_M   ='OSL Z_ M'
    OSLB_MT  ='OSL B_ M,T'    OSLT_MB  ='OSL T_ M,B'    OSLZ_MB  ='OSL Z_ M,B'
    OSLB_MZ  ='OSL B_ M,Z'    OSLT_MZ  ='OSL T_ M,Z'    OSLZ_MT  ='OSL Z_ M,T'
    OSLB_MTZ ='OSL B_ M,T,Z'  OSLT_MBZ ='OSL T_ M,B,Z'  OSLZ_MBT ='OSL Z_ M,B,T'
    DFB      ='DF B'             DFT      = 'DF T'           DFZ = 'DF Z'
  MUHATG  = 'MU HAT GLM'         BL1HATG  = 'BLOCK 1 HAT GLM'
  TR1HATG = 'TREAT 1 HAT GLM'    TR2HATG  = 'TREAT 2 HAT GLM'
  TR3HATG = 'TREAT 3 HAT GLM'    BETAHAT  = 'BETA HAT GLM=AOV'
    MUHATA  = 'Mu hat aov'
    BL1HATA = 'Block 1 Hat Aov'    bl2hata  = 'Block 2 Hat Aov'
    TR1HATA = 'Treat 1 Hat Aov'    tr2hata  = 'Treat 2 Hat Aov'
    TR3HATA = 'Treat 3 Hat Aov'    tr4hata  = 'TREAT 4 Hat Aov'
  BL1YBAR='BLOCK 1 Y BAR' TR1YBAR='TREAT 1 Y BAR' TR1ZBAR='TREAT 1 Z BAR'
  BL2YBAR='BLOCK 2 Y BAR' TR2YBAR='TREAT 2 Y BAR' TR2ZBAR='TREAT 2 Z BAR'
  BL1ZBAR='BLOCK 1 Z BAR' TR3YBAR='TREAT 3 Y BAR' TR3ZBAR='TREAT 3 Z BAR'
  BL2ZBAR='BLOCK 2 Z BAR' TR4YBAR='TREAT 4 Y BAR' TR4ZBAR='TREAT 4 Z BAR'
    AMUHATA ='R-EXPT MUHATA'
    ABL1HATA='R-EXPT BL1HATA'      ABL2HATA='R-EXPT BL2HATA'
    ATR1HATA='R-EXPT TR1HATA'      ATR2HATA='R-EXPT TR2HATA'
    ATR3HATA='R-EXPT TR3HATA'      ATR4HATA='R-EXPT TR4HATA'
    ABETAHAT='R-EXPT BETAHAT'
    AMSB_M ='R-EXPT MSB_M'
    AMST_MB='R-EXPT MST_MB'      AMST_MBZ='R-EXPT MST_MBZ'
    AMSZ_MB='R-EXPT MSZ_MB'      AMSZ_MBT='R-EXPT MSZ_MBT'
    AFMOD  ='R-EXPT F_MOD'       AFB_M   ='R-EXPT FB_M'
    AFT_MB ='R-EXPT FT_MB'       AFT_MBZ ='R-EXPT FT_MBZ'
    AFZ_MB ='R-EXPT FZ_MB'       AFZ_MBT ='R-EXPT FZ_MBT'
```

```
        AZSS_MBT='R-EXPT ZSS_MBT'    OMEGA='R-EXPT ZSS_MBT/DF'
    RMU ='R-ERR MUHATA'      RBL1= 'R-ERR BL1HATA'    RBL2='R-ERR BL2HATA'
    RTR1='R-err tr1hata'     RTR2= 'R-err tr2hata'    RTR3='R-err tr3hata'
    RTR4='R-err tr4hata'     RBETA='R-err betahat'
    RMST_MB='R-ERR MST_MB'   RMST_MBZ='R-err mst_mbz'
    RMSZ_MB='R-ERR MSZ_MB'   RMSZ_MBT='R-err msz_mbt'
    RFMOD ='R-ERR FMOD'      RFB_M  ='R-ERR FB_M'
    RFT_MB ='R-ERR FT_MB'    RFT_MBZ ='R-ERR FT_MBZ'
    RFZ_MB ='R-ERR FZ_MB'    RFZ_MBT ='R-ERR FZ_MBT'
    SMU ='S-ERR MUHATA'      SBL1='S-ERR BL1HATA'    SBL2='S-ERR BL2HATA'
    STR1='S-err tr1hata'     STR2='S-err tr2hata'    STR3='S-err tr3hata'
    STR4='S-err tr4hata'     SBETA='S-err betahat'
    ZSS_MBT = 'Z SS _MBT'
 /* RATIO1 = 'zsst_mb/ dft#zss_mbt'
    RATIO2 = 'zsst_mb/ omega'
    RATIO3 = 'zss_mbt/ azss_mbt' ;
                         1          2          3          4|
            123456789 123456789 123456789 123456789 <MAX
 */ RATIO1 = 'R1-for Z: Adj. Treat. ms/residual ss'
    RATIO2 = 'R2-for Z: Adj. Treat. ss/E(residual ms)'
    RATIO3 = 'R3-for Z: Residual ss/E(residual ss)'  ;
*
*
* Print aov variables from first few experimental radnomizations.
* Labeling is temporary for the proc print operation.;
PROC PRINT  DATA=sasglm.&glm_ds (OBS=26) LABEL    SPLIT = '*' ;
VAR exp_rand b1_rand  b2_rand  g_group
    dfmod -- tr4zbar   muhata -- ratio1   rmu--ratio3 ;
LABEL
  EXP_RAND='EXPERI*RANDOM'
  B1_RAND ='BLOCK 1*RANDOM'  B2_RAND ='BLOCK 2*RANDOM'
  /*
  ZT11='Z BL1*TREAT 1'  ZT12='Z BL1*TREAT 2'  ZT13='Z BL1*TREAT 3'
  ZT14='Z BL1*TREAT 4'  ZT21='Z BL2*TREAT 1'  ZT22='Z BL2*TREAT 2'
  ZT23='Z BL2*TREAT 3'  ZT24='Z BL2*TREAT 4'
  */
    SSMOD = 'SS*MODEL'    SSRES  = 'SS*RESIDUAL'  SSTOT = 'SS*TOTAL'
    MSMOD = 'MS*MODEL'    MSRES  = 'MS*RESIDUAL'
    FMOD  = 'F.*MODEL'    OSLMOD = 'OSL*MODEL'      ROOTMSE='ROOT*MSE'
    DFMOD = 'D.F.*MODEL'  DFRES  = 'D.F.*RESIDUAL'  DFTOT ='D.F.*TOTAL'
    COEFVAR ='COEF*VAR'   RSQ ='R*SQUARED'  MEAN_Y = 'MEAN Y*RESPONSE'
  SSB_M  ='SS B_*M'     SST_M  ='SS T_*M'     SSZ_M  ='SS Z_*M'
  SSB_MT ='SS B_*M,T'  SST_MB ='SS T_*M,B'   SSZ_MB ='SS Z_*M,B'
  SSB_MZ ='SS B_*M,Z'  SST_MZ ='SS T_*M,Z'   SSZ_MT ='SS Z_*M,T'
  SSB_MTZ ='SS B_*M,T,Z' SST_MBZ ='SS T_*M,B,Z' SSZ_MBT ='SS Z_*M,B,T'
    MSB_M  ='MS B_*M'     MST_M  ='MS T_*M'      MSZ_M  ='MS Z_*M'
    MSB_MT ='MS B_*M,T'   MST_MB ='MS T_*M,B'    MSZ_MB ='MS Z_*M,B'
    MSB_MZ ='MS B_*M,Z'   MST_MZ ='MS T_*M,Z'    MSZ_MT ='MS Z_*M,T'
    MSB_MTZ ='MS B_*M,T,Z' MST_MBZ ='MS T_*M,B,Z' MSZ_MBT ='MS Z_*M,B,T'
  FB_M  = 'F B_*M'     FT_M  = 'F T_*M'     FZ_M  = 'F Z_*M'
  FB_MT = 'F B_*M,T'   FT_MB = 'F T_*M,B'   FZ_MB = 'F Z_*M,B'
  FB_MZ = 'F B_*M,Z'   FT_MZ = 'F T_*M,Z'   FZ_MT = 'F Z_*M,T'
  FB_MTZ = 'F B_*M,T,Z' FT_MBZ = 'F T_*M,B,Z' FZ_MBT = 'F Z_*M,B,T'
    OSLB_M ='OSL B_*M'     OSLT_M ='OSL T_*M'     OSLZ_M ='OSL Z_*M'
    OSLB_MT ='OSL B_*M,T'  OSLT_MB ='OSL T_*M,B'  OSLZ_MB ='OSL Z_*M,B'
    OSLB_MZ ='OSL B_*M,Z'  OSLT_MZ ='OSL T_*M,Z'  OSLZ_MT ='OSL Z_*M,T'
    OSLB_MTZ='OSL B_*M,T,Z' OSLT_MBZ='OSL T_*M,B,Z' OSLZ_MBT='OSL Z_*M,B,T'
    DFB  ='DF*B'                DFT  = 'DF*T'           DFZ = 'DF*Z'
  MUHATG = 'MU HAT*GLM'              BL1HATG = 'BLOCK 1*HAT GLM'
  TR1HATG = 'TREAT 1*HAT GLM'     TR2HATG = 'TREAT 2*HAT GLM'
  TR3HATG = 'TREAT 3*HAT GLM'     BETAHAT = 'BETA HAT*GLM=AOV'
    MUHATA = 'MU HAT*AOV'
    BL1HATA = 'BLOCK 1*HAT AOV'    BL2HATA = 'BLOCK 2*HAT AOV'
    TR1HATA = 'TREAT 1*HAT AOV'    TR2HATA = 'TREAT 2*HAT AOV'
    TR3HATA = 'TREAT 3*HAT AOV'    TR4HATA = 'TREAT 4*HAT AOV'
  BL1YBAR='BLOCK 1*Y BAR' TR1YBAR='TREAT 1*Y BAR' TR1ZBAR='TREAT 1*Z BAR'
```

```
       BL2YBAR='BLOCK 2*Y BAR' TR2YBAR='TREAT 2*Y BAR' TR2ZBAR='TREAT 2*Z BAR'
       BL1ZBAR='BLOCK 1*Z BAR' TR3YBAR='TREAT 3*Y BAR' TR3ZBAR='TREAT 3*Z BAR'
       BL2ZBAR='BLOCK 2*Z BAR' TR4YBAR='TREAT 4*Y BAR' TR4ZBAR='TREAT 4*Z BAR'
       /* THE FOLLOWING CONSTANT VARIABLES ARE NOT PRINTED HERE.
          THEY ARE PRINTED AS OUTPUT FROM THE MEANS PROCEEDURE.
        AMUHATA ='R-EXPT*MUHATA'
        ABL1HATA='R-EXPT*BL1HATA'         ABL2HATA='R-EXPT*BL2HATA'
        ATR1HATA='R-EXPT*TR1HATA'         ATR2HATA='R-EXPT*TR2HATA'
        ATR3HATA='R-EXPT*TR3HATA'         ATR4HATA='R-EXPT*TR4HATA'
        ABETAHAT='R-EXPT*BETAHAT'
        AMSB_M  ='R-EXPT*MSB_M'
        AMST_MB='R-EXPT*MST_MB'     AMST_MBZ='R-EXPT*MST_MBZ'
        AMSZ_MB='R-EXPT*MSZ_MB'     AMSZ_MBT='R-EXPT*MSZ_MBT'
        AFMOD  ='R-EXPT*F_MOD'      AFB_M   ='R-EXPT*FB_M'
        AFT_MB ='R-EXPT*FT_MB'      AFT_MBZ ='R-EXPT*FT_MBZ'
        AFZ_MB ='R-EXPT*FZ_MB'      AFZ_MBT ='R-EXPT*FZ_MBT'
        AZSS_MBT='R-EXPT*ZSS_MBT'     OMEGA='R-EXPT*ZSS_MBT/DF'
       */
       RMU  ='R-ERR*MUHATA'      RBL1= 'R-ERR*BL1HATA'    RBL2='R-ERR*BL2HATA'
       RTR1='R-ERR*TR1HATA'      RTR2='R-ERR*TR2HATA'     RTR3='R-ERR*TR3HATA'
       RTR4='R-ERR*TR4HATA'      RBETA='R-ERR*BETAHAT'
       RMST_MB='R-ERR*MST_MB'    RMST_MBZ='R-ERR*MST_MBZ'
       RMSZ_MB='R-ERR*MSZ_MB'    RMSZ_MBT='R-ERR*MSZ_MBT'
       RFMOD  ='R-ERR*FMOD'      RFB_M   ='R-ERR*FB_M'
       RFT_MB ='R-ERR*FT_MB'     RFT_MBZ ='R-ERR*FT_MBZ'
       RFZ_MB ='R-ERR*FZ_MB'     RFZ_MBT ='R-ERR*FZ_MBT'
       SMU  ='S-ERR*MUHATA'      SBL1='S-ERR*BL1HATA'    SBL2='S-ERR*BL2HATA'
       STR1='S-ERR*TR1HATA'      STR2='S-ERR*TR2HATA'    STR3='S-ERR*TR3HATA'
       STR4='S-ERR*TR4HATA'      SBETA='S-ERR*BETAHAT'
         ZSS_MBT = 'Z SS*_MBT'
         RATIO1 = 'R1-ZSST_MB/*DFT#ZSS_MBT'  RATIO2= 'R2-ZSST_MB/*OMEGA'
         RATIO3 = 'R3-ZSS_MBT/*AZSS_MBT' ;

********************************************************************;
* Step Eight
   Analysis of aov estimates and responses.;

PROC MEANS DATA = sasglm.&glm_ds      N MEAN STD VAR MIN MAX;
   TITLE1 "Two types of errors R=randomization  S=population";
   TITLE2 "Data set is  sasglm.&glm_ds ";
   VAR   RMU     RBL1  RBL2    RTR1  RTR2  RTR3  RTR4     RBETA
         RMST_MB  RMST_MBZ  RMSZ_MB  RMSZ_MBT
           SMU    SBL1  SBL2    STR1  STR2  STR3  STR4     SBETA
         OMEGA AZSS_MBT   ZSS_MBT    RATIO1 RATIO2 RATIO3;

PROC MEANS DATA = sasglm.&glm_ds      N MEAN STD VAR MIN MAX;
   TITLE1 "Mean squares weighted with  ratio3";
   TITLE2 "Data set is  sasglm.&glm_ds ";
   WEIGHT RATIO3;
   VAR MSRES  MST_MBZ  MSZ_MBT  ft_mbz  fz_mbt
       BETAHAT  muhata  bl1hata  bl2hata
       tr1hata  tr2hata  tr3hata  tr4hata  ;

PROC CHART   DATA = sasglm.&glm_ds;
   TITLE1 "The Randomization Model with Blocks and a Covariate";
   TITLE2 "Data set is  sasglm.&glm_ds ";
   HBAR  MST_MBZ  MSZ_MBT MSRES
         FT_MBZ    FZ_MBT
         OSLT_MBZ OSLZ_MBT

         MUHATA  BL1HATA BL2HATA
         TR1HATA TR2HATA TR3HATA TR4HATA   BETAHAT
         SMU    SBL1    SBL2    STR1   STR2   STR3   STR4    SBETA
   / TYPE = PERCENT    LEVELS = 20 ;

      /*
```

```
              OSLMOD FMOD      MSMOD MSRES     SSTOT SSMOD SSRES
          SSB_M        SST_MB      SST_MBZ       SSZ_MB       SSZ_MBT
          MSB_M        MST_MB      MST_MBZ       MSZ_MB       MSZ_MBT
          FB_M         FT_MB       FT_MBZ        FZ_MB        FZ_MBT
          OSLB_M       OSLT_MB     OSLT_MBZ      OSLZ_MB      OSLZ_MBT
          RSQ   COEVAR  ROOTMSE   MEAN_Y
            ZSS_MBT RATIO1 RATIO2  RATIO3
            RMU      RBL1       RBL2      RTR1     RTR2     RTR3     RTR4     RBETA
            RMST_MB           RMST_MBZ       RMSZ_MB       RMSZ_MBT
            SMU      SBL1      SBL2      STR1     STR2     STR3     STR4      SBETA
      */


PROC PLOT     DATA = sasglm.&glm_ds;
     TITLE1 "The Randomization Model with Blocks and a Covariate";
     TITLE2 "Data set is  sasglm.&glm_ds ";
         PLOT B1_RAND * B2_RAND = MST_MBZ / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = MSZ_MBT / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = MSRES   / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = FT_MBZ  / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = FZ_MBT  / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = OSLT_MBZ/ HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = OSLZ_MBT/ HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = BETAHAT / HPOS=50   CONTOUR=10 ;
  /*   PLOT B1_RAND * B2_RAND = ZSS_MBT / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RSQ     / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = COEFVAR / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = MEAN_Y  / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RMU     / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RBL1    / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RTR1    / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RTR2    / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RTR3    / HPOS=50   CONTOUR=10 ;.
         PLOT B1_RAND * B2_RAND = RTR4    / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RBETA   / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RMST_MB / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RMST_MBZ/ HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RMSZ_MB / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RMSZ_MBT/ HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RFMOD   / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RFB_M   / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RFT_MB  / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RFT_MBZ / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RFZ_MB  / HPOS=50   CONTOUR=10 ;
         PLOT B1_RAND * B2_RAND = RFZ_MBT / HPOS=50   CONTOUR=10 ;
   */


PROC PLOT     DATA = sasglm.&glm_ds;
     TITLE1 "The Randomization Model with Blocks and a Covariate";
     TITLE2 "Data set is  sasglm.&glm_ds ";
     PLOT (msz_mbt msres  fz_mbt  betahat)  * g_group / HPOS=50;
     PLOT(SMU    SBL1  SBL2      STR1    STR2    STR3    STR4    SBETA )
         * RATIO2 / HPOS=50    VREF=0.0;
     PLOT (RATIO1 RATIO2) * RATIO3 / HPOS=50;
     PLOT RATIO1 * RATIO2 / HPOS=50;
 /*
   PLOT(RMU     RBL1   RBL2      RTR1    RTR2    RTR3    RTR4     RBETA
        RMST_MB          RMST_MBZ      RMSZ_MB      RMSZ_MBT
           SMU    SBL1  SBL2     STR1    STR2    STR3    STR4     SBETA )
        * RATIO2 / HPOS=50   VREF=0.0;
 */


PROC CHART    DATA = sasglm.&yhat_ds;
     TITLE1 "The Randomization Model with Blocks and a Covariate";
```

```
     TITLE2 "Data set is   sasdata.obs_y  ":
  HBAR    &yobs  &error  &ypred  &resid
           / TYPE = PERCENT      LEVELS = 20;



   /*
PROC PLOT    DATA = sasglm.&yhat_ds;
   TITLE1 "The Randomization Model with Blocks and a Covariate";
   TITLE2 "Data set is   sasdata.obs_y  ";
       PLOT &resid * (&pred covar)   / HPOS=50 VREF=0.0;
       PLOT &yobs  *  &pred          / HPOS=50;
   */



   //



   *    Debug Print options follow: ;
       %db(&dbps1, PRINT   "Data input from eu_err and euti_err   (dbps1)";
                   PRINT   mu_val  b_valm  trt_valm     beta  zm
                           labels
                           eu_eqm eu_nem       euti_eqm    euti_nem; );

   ********************** Create Output Data Set *********************;
   ********* Test Print Initial Version of Data Set obs_y *****;

   *    Debug Print options follow: ;
       %db(&dbps3, TITLE1 "Observations from   rm_err   (dbps3)";
                   PROC PRINT DATA=rm_err (OBS=32);
                   PROC PRINT DATA=rm_err (FIRSTOBS=4577);
                   PROC MEANS DATA=rm_err (OBS=32)
                       N MEAN VAR STD  VARDEF=N;
                    BY  exp_rand  block;
                    VAR eu_eq euti_eq  eu_ne euti_ne;
                   PROC MEANS DATA=rm_err (OBS=32)
                       N MEAN VAR STD  VARDEF=N;
                    BY exp_rand;
                    VAR eu_eq euti_eq  eu_ne euti_ne;
                   PROC MEANS DATA=rm_err
                       N MEAN VAR STD  VARDEF=N;
                    VAR eu_eq euti_eq  eu_ne euti_ne;
                   );
DATA sasdata.rm_err;  SET rm_err;     *Randomization Model ERRor;

DATA all_err;     *ALL three ERRor types;
  MERGE sasdata.rm_err    sasdata.nor_err;      BY exp_rand;
  covar = z;     drop z ;
  LABEL
       exp_rand= 'experiment-wise rando'
       b1_rand = 'block 1 randomization'
       b2_rand = 'block 2 randomization'

       block   = 'block id'
       b_val   = 'block effect value'
       plot    = 'experimental unit id'
       plabel  = 'plot label'
       treat   = 'treatment id'
       trt_val = 'treat effect value'
       tlabel  = 'treatment label'
       covar   = 'covariate (z)'
       beta    = 'covariate slope coef.'

       eu_eq   = 'e.u. error var-eq'
```

```
            eu_ne    = 'e.u. error var-ne'
            eut1_eq  = 'e.u.t.1. error var-eq'
            eut1_ne  = 'e.u.t.1. error var-ne'

            nre03_eq = 'normal err v=3 -eq'
            nre03_ne = 'normal err v=3 -ne'
            nre10_eq = 'normal err v=10 -eq'
            nre10_ne = 'normal err v=10 -ne'
            nre30_eq = 'normal err v=30 -eq'
            nre30_ne = 'normal err v=30 -ne'
            nre90_eq = 'normal err v=90 -eq'
            nre90_ne = 'normal err v=90 -ne'

            s_popynt= 'superpop. no treat'
            s_popywt= 'superpop. with treat'
            ;


**** Create macro too generate formula for ***************************;
************* the observed population values yobs(#)(##)  *************;
    %MACRO   m_yobs;
        %DO   err_id = 1 %TO 9 ;
            yobs00&err_id =  s_popynt +  e0&err_id;
            yobs10&err_id =  s_popywt +  e0&err_id;
        %END;
        %DO   err_id = 10 %TO 81;
            yobs0&err_id =  s_popynt +  e&err_id;
            yobs1&err_id =  s_popywt +  e&err_id;
        %END;
    %MEND;
************* End of macro m_yobs ************************************;
    /*  *Successful step turned off. ;
```

```
U13293A.SASPROG.CNTL(NONCENT)
VPSPRINT 5.1.002   WEDNESDAY NOVEMBER 4,1992   14:59:16 U13293A   MVS1          **
VPSPRINT SASPROG.CNTL(*) LOCAL FORM(9001) CLASS(E)

//U13293AA JOB (*),'R. S. WILSON',TIME=(0,9),CLASS=2,
//  MSGCLASS=X,MSGLEVEL=(1,1)
/*JOBPARM  ROOM=R,FORMS=9001,COPIES=1
//   EXEC SAS,REGION=918K
//SASDATA   DD   DSN=U13293A.SASDATA.DATA,DISP=OLD
//SYSIN  DD *

  /*************************************************************
  *  File:  sasprog.cntl(noncent)
  *
  *  Compute noncentrality parameter
  *      Use specified treatments, covariates, and slope coefficient
  *      Compute various parts of the noncentrality formulas.
  *
  *      Variance (sigma squared) and 1/2 are omitted from all
  *      noncentrality parameters.
  *
  *
  *************************************************************/

options  nodate;

proc iml;
 *Variable value matrices follow;
  * Treatment coefficients;
      bt = { -6.5, -3.5, 2.5};
  * Covariate coefficient(s);
      bz = {  2.0 };
  * Covariate values  with comments on equal-valued g-groups;
   Title1 'Noncentrality Parameters - less (1/2)(sigma squared)';
   Title2 'covariate = {1,2,11,14,  5,9,16,19}  coefficient = 2.0';
   Title3 'treatments = -6.5, -3.5, 2.5 (+7.5)}';
      xz = { 1, 2, 11, 14,   5,  9, 16, 19}; *NO dup 1 change from org;
    * xz = { 1, 2, 11, 14,   5,  9, 17, 19}; *original, reordered;
    * xz = { 1, 2, 11, 14,   7,  9, 15, 19}; *NO dup;
    * xz = { 1, 2, 11, 14,   7,  9, 15, 18}; *1 dup ;
    * xz = { 1, 2, 11, 14,  18, 17,  9,  5}; *1 dup;
    * xz = { 1, 2, 11, 14,   7,  9, 17, 19}; *NO dup  2 diff=10;
    * xz = { 1, 2, 11, 14,  19, 17,  9,  5}; *1 dup and eq following;
    * xz = { 1, 2, 11, 14,   5,  9, 17, 19}; *1 dup and eq above;
    * xz = { 2.63, 9.11, 16.19, 20.93,   30.21, 40.21, 41.09, 50.43};

  * Debug print switches    prints if true;
      dbps1 = 'false' ;      *escape;
      dbps2 = 'false';       *constant values;
      dbps3 = 'false';       *permutation matrix;

 *Constant valued matrices follow;
    xm  = {1, 1, 1, 1,   1,  1,  1,  1};
    xb  = {1, 1, 1, 1,  -1, -1, -1, -1};
    xt  = {1 0 0,  0 1 0,  0 0 1,  -1  -1  -1,
           1 0 0,  0 1 0,  0 0 1,  -1  -1  -1};
    w   = ( j(4,4, 1/4) || j(4,4,0) )  //  ( j(4,4,0) || j(4,4, 1/4) );
    lam_ft = bt¢ * xt¢*xt * bt;
    I = I(8);
    lam_fb = bz¢ * xz¢ * (I - w) * xz * bz;
    slope = lam_ft / lam_fb;
    xxxx_t = xt * inv(xt¢ * xt)        * xt¢;
    zzzz   = xz * inv(xz¢ * xz)        * xz¢;
```

```
   ziz    = xz * inv(xz¢* (I - w) * xz) * xz¢;
   m4 = J(4,4,0);
   j43= J(4,3,0);
   zero_esc = J(1,21,0);
   escape = zero_esc;

 * Debug print switch;
     if dbps2 =   'true' then print  w xxxx_t ziz  lam_ft lam_fb;

*Each row of Mpattern (randomization patterns) is one permitted
*  block-randomization or assignment of treatments (element value)
*  to experimental units (element position) within one block.
*;
  rpattern =
  { 1 2 3 4 ,  1 2 4 3 ,  1 3 2 4 ,  1 3 4 2 ,  1 4 2 3 ,  1 4 3 2,
    2 1 3 4 ,  2 1 4 3 ,  2 3 1 4 ,  2 3 4 1 ,  2 4 1 3 ,  2 4 3 1,
    3 1 2 4 ,  3 1 4 2 ,  3 2 1 4 ,  3 2 4 1 ,  3 4 1 2 ,  3 4 2 1,
    4 1 2 3 ,  4 1 3 2 ,  4 2 1 3 ,  4 2 3 1 ,  4 3 1 2 ,  4 3 2 1 } ;

  varname    = {exp_rand  b1_rand  b2_rand   g_group
             lam269t  lam2610t  lam2634t  lam_dt    lam_ft     lam_st
             lam269b  lam2610b  lam2630b  lam_db    lam_fb     lam_gb
             l_s3715      l_s          r_1       r_2      slope      };

   * lam269_t  =  Eq 2.6.9  noncent tau  given rest;
   * lam2610_t = Eq 2.6.10 noncent tau  given rest;
   * lam2634_t = Eq 2.6.34 noncent tau  given rest;
   * lam_dt =  noncent tau  given rest as (fixed - random);
   * lam_ft =  fixed part of tau given rest;
   * lam_st =  random part of tau  given rest;

   * lam269_b  =  Eq 2.6.9  noncent cov  given rest;
   * lam2610_b  = Eq 2.6.10 noncent cov  given rest;
   * lam2630_b  = Eq 2.6.30 noncent cov  given rest;
   * lam_db =  noncent beta given rest as (fixed - random);
   * lam_fb =  fixed part of beta given rest;
   * lam_gb =  group random part of beta given rest;

   * l_s3715=  (lam_fb * lam_st) / lam_ft;
   * l_s    =  (bz¢ * xz ¢ M_s * xt * bt) squared;
   * r_1    =  Right side for Theo. 3.4.13  gb * ft;
   * r_2    =  Right side for Theo. 3.4.14  st * fb;

   * slope =  lam_ft / lam_fb;

Gtable   =
{1   2    3   5   4   6    7   8  13  19  14 20   9  11  15  21  17  23  10  12  16  22  18 24,
 2   1    4   6   3   5    8   7  14  20  13  19  10  12  16  22  18  24   9  11  15  21  17 23,
 3   5    1   2   6   4    9  11  15  21  17  23   7   8  13  19  14  20  12  10  18  24  16 22,
 4   6    2   1   5   3  10  12  16  22  18  24   8   7  14  20  13  19  11   9  17  23  15 21,
 5   3    6   4   1   2  11   9  17  23  15  21  12  10  18  24  16  22   7   8  13  19  14 20,
 6   4    5   3   2   1  12  10  18  24  16  22  11   9  17  23  15  21   8   7  14  20  13 19,
 7   8   13  19  14  20   1   2   3   5   4   6  15  21   9  11  23  17  16  22  10  12  24 18,
 8   7   14  20  13  19   2   1   4   6   3   5  16  22  10  12  24  18  15  21   9  11  23 17,
 9  11   15  21  17  23   3   5   1   2   6   4  13  19   7   8  20  14  18  24  12  10  22 16,
10  12   16  22  18  24   4   6   2   1   5   3  14  20   8   7  19  13  17  23  11   9  21 15,
11   9   17  23  15  21   5   3   6   4   1   2  18  24  12  10  22  16  13  19   7   8  20 14,
12  10   18  24  16  22   6   4   5   3   2   1  17  23  11   9  21  15  14  20   8   7  19 13,
13  19    7   8  20  14  15  21   9  11  23  17   1   2   3   5   4   6  22  16  24  18  10 12,
14  20    8   7  19  13  16  22  10  12  24  18   2   1   4   6   3   5  21  15  23  17   9 11,
15  21    9  11  23  17  13  19   7   8  20  14   3   5   1   2   6   4  24  18  22  16  12 10,
16  22   10  12  24  18  14  20   8   7  19  13   4   6   2   1   5   3  23  17  21  15  11  9,
17  23   11   9  21  15  18  24  12  10  22  16   5   3   6   4   1   2  19  13  20  14   7  8,
18  24   12  10  22  16  17  23  11   9  21  15   6   4   5   3   2   1  20  14  19  13   8  7,
19  13   20  14   7   8  21  15  23  17   9  11  22  16  24  18  10  12   1   2   3   5   4  6,
20  14   19  13   8   7  22  16  24  18  10  12  21  15  23  17   9  11   2   1   4   6   3  5,
```

```
21 15 23 17  9 11 19 13 20 14  7  8 24 18 22 16 12 10  3  5  1  2  6  4,
22 16 24 18 10 12 20 14 19 13  8  7 23 17 21 15 11  9  4  6  2  1  5  3,
23 17 21 15 11  9 24 18 22 16 12 10 19 13 20 14  7  8  5  3  6  4  1  2,
24 18 22 16 12 10 23 17 21 15 11  9 20 14 19 13  8  7  6  4  5  3  2  1};

*G-table was copied in from output from g-table program, below. ;

******************** begin looping through randomizations *****;
create   temp  from escape  (|colname = varname |);

  start mb1_rand;
  do b1_rand = 1 to 24  by 1;
     b1 = j43;
     m1 = m4;
     start mp_set1;
     do p_set1 = 1 to 4 by 1;    *1 row of rpattern;
        one_col = rpattern(|b1_rand, p_set1 |);
        if one_col  = 4   then b1(|p_set1,      |)    = -1;
        if one_col ¬= 4   then b1(|p_set1, one_col |)= 1;
        m1(|p_set1, one_col |) = 1;
     end;
     finish;
     run mp_set1;

  start mb2_rand;
  do b2_rand = 1 to 24  by 1;
     *Compute experimental randomization number;
        exp_rand = (b1_rand - 1) # 24 + b2_rand;
     *Compute randomization group number from gtable;
        group = gtable(|b1_rand, b2_rand |);
        b2 = j43;
        m2 = m4;
     start mp_set2;
     do p_set2 = 1 to 4 by 1;
        one_col = rpattern(|b2_rand, p_set2 |);
        if one_col  = 4 then b2(|p_set2,      |)    = -1;
        if one_col ¬= 4 then b2(|p_set2, one_col |)= 1;
        m2(|p_set2, one_col |)= 1;
     end;
     finish;
     run mp_set2;


  xt_s = b1 // b2 ;
  xxxx_ts = xt_s * inv(xt_s¢ * xt_s) * xt_s¢;
  m_s = (m1 || m4) // ( m4 || m2) ;
*Debug print switch;
     if dbps3 = 'true' then print  xt_s xxxx_ts  m_s;

*Compute direct noncentralities;
  *Section 2.6  for cov given rest;
     x_p = xm || xb || xt_s || xz;
     xxxx_p = x_p * inv(x_p¢ * x_p) * x_p¢;
     b_h = bz;  x_h = xz;
     x_r = xm || xb || xt_s ;
     xxxx_r = x_r * inv(x_r¢ * x_r) * x_r¢;
     lam269b  = (x_h * b_h)¢ * ( xxxx_p - xxxx_r ) * (x_h * b_h);
     lam2610b = b_h¢*( (x_h¢*x_h) - (x_h¢ * xxxx_r * x_h) )*b_h;
     lam2630b = (bz¢ * xz¢ *  (i-w)  * xz * bz )
                -(bz¢ * xz¢ * xxxx_ts * xz * bz );

  *Section 2.6  for tau given rest;
     b_h = bt;  x_h = xt_s;
     x_r = xm || xb || xz ;
     xxxx_r = x_r * inv(x_r¢ * x_r) * x_r¢;
```

```
        lam269t  = (x_h * b_h)¢ * ( xxxx_p - xxxx_r ) * (x_h * b_h);
        lam2610t = b_h¢*( (x_h¢*x_h) - (x_h¢ * xxxx_r * x_h) )*b_h;
        lam2634t = (bt¢ * xt_s¢ *         xt_s * bt)
                  -(bt¢ * xt_s¢ * ziz  * xt_s * bt);

*Theorem numbers correspond to dissertation text.;
*Compute random parts of noncent using permutor  Theo 3.7.10 & 11;
   lam_gb = bz¢ * xz¢ *     m_s  * xxxx_t * m_s¢   * xz * bz;
   lam_st = bt¢ * xt¢ *     m_s¢ *  ziz   * m_s    * xt * bt;

*Compute the (Diference=fixed-random) noncentralities;
   lam_db = lam_fb - lam_gb;
   lam_dt = lam_ft - lam_st;

*Compute left side of Theo 3.7.15 which is < lam_gb;
   l_s3715  = (lam_fb * lam_st ) / lam_ft;

*Compute left and right-hand sides for Theo 3.7.13 & 14;
*  Since (1/2)sigma omitted from lamda, 2sigma is also omitted.;
   l_s = bz¢ * xz¢ * M_s * xt * bt;
   l_s = l_s * l_s;
   r_1 = lam_gb * lam_ft;
   r_2 = lam_st * lam_fb;

*Debug print switch;
   if dbps1 = 'true' then print
          exp_rand      b1_rand       b2_rand       group
            lam269t       lam2610t      lam2634t
            lam_dt        lam_ft        lam_st
            lam269b       lam2610b      lam2630b
            lam_db        lam_fb        lam_gb
        l_s3715      l_s           r_1          r_2           slope  ;
  /*
     if group = 24 then  t = m_s¢ * ziz * m_s;
     if group = 24 then  print
           b1_rand  b2_rand  (|format=2.0|)
           lam_gb lam_st  (|format=8.5|)      t  (|format=5.3|);
  */


* Output results;
        escape = zero_esc;
        escape =
          exp_rand ||  b1_rand  || b2_rand  ||  group     ||
             lam269t  || lam2610t || lam2634t ||
             lam_dt  || lam_ft  || lam_st   ||
           lam269b  || lam2610b ||  lam2630b ||
             lam_db  || lam_fb  || lam_gb   ||
         l_s3715 ||  l_s      || r_1     ||  r_2      || slope;
        setout  temp;
        append from escape;

  end;
  finish;
  run  mb2_rand;

  end;
  finish;
  run  mb1_rand;
*********************** end of iml ********************;


data temp;  set temp;
   lam_gb = round(lam_gb, 1);
   lam_db = round(lam_db, 1);
   label
```

```
            lam_dt ='Lamda: treatments given rest'
            lam_db ='Lamda: covariates given rest'
            lam_ft ='Fixed part: treatments given rest'
            lam_fb ='Fixed part: covariates given rest'
            lam_st ='Random part: treatments given rest'
            lam_gb ='Random part: covariates given rest'
            l_s3715='(Fix cov)/(Fix treat)*(Ran treat)'
            l_s    ='Cauchy-S left side'
            r_1    ='(Fix tre) * (Ran cov)'
            r_2    ='(Fix cov) * (Ran tre)'
            slope  ='(Fix tre) / (Fix cov)'
               ;

proc sort   data=temp;  by lam_gb;
proc means  data=temp  maxdec = 8   noprint
          n mean var std min max;
   by lam_gb;
   id g_group;
   var         lam_fb  lam_gb   lam_db     lam_ft    lam_st  lam_dt;
   output out=meanout
       mean= mean_fb mean_gb   mean_db    mean_ft  mean_st mean_dt
       var = var_fb  var_gb    var_db     var_ft   var_st  var_dt
         n =   n_fb    n_gb      n_db       n_ft     n_st    n_dt
       max = max_fb  max_gb    max_db     max_ft   max_st  max_dt
       ;

   proc print  data=meanout;
       var g_group n_gb  mean_fb mean_gb mean_db  var_db
                        mean_ft mean_st mean_dt  var_dt;

*Plot the two noncentrality parameters against one another.;
proc plot  data=temp;
   plot lam_dt * lam_db;
   plot lam_st * lam_gb;
   plot l_s * ( r_1  r_2) ;


*Plot the variance vs the mean of lamda treatments given rest;
proc plot  data=meanout;
   plot var_dt * mean_dt;

*Display distribution of noncentrality parameters, parts, and C.S.;
proc chart data=temp;
   hbar lam_dt lam_db   lam_st   lam_gb  l_s
       / type = percent   levels = 24;

*Test slope estimated from actual data vs slope in Cauchy-S. result;
proc reg  data=meanout;
   Title4 '*=predicted point  o=actual data points  ?=overlap';
   model max_st = max_gb;
   plot p. * max_gb='*'   max_st * max_gb='o' / overlay;


*Print randomizations with the largest and smallest lamdas;
   proc sort   data=temp; by lam_db  lam_dt;
   proc print  data=temp  (obs = 26);
     Title4
 'Data ordered by noncentrality of treatments given the rest (lam_dt)';
     Title5
 'within noncentrality of covariate given the rest (lam_db)            ';
   proc print  data=temp  (firstobs=551);

*Print first and last randomizations;
   proc sort   data=temp;  by exp_rand;
     Title4 'Data ordered by experimental randomization number';
     Title5 '';
```

```
     proc print  data=temp  (obs = 26);
     proc print  data=temp  (firstobs=551);


*Store data in permanent sas data file;
   data  sasdata.lamda;   set temp;

  /****************************************************************
  *  The following program computes the above matrix Gtable
  *
  *  Gtable provides the block randomization id number (1,2,...t!)
  *      for the off diagonal cell of m_s, the experiment-wide
  *      permutation matrix.  Randomizations with the same number
  *      have the same value of lamda: covariate given rest.
  * The table Gtable was computed in a prior run and copied into
  *      the above program.
  *
  *****************************************************************/
  /*
*Compute the transpose of each of the 24 permutation matrices
   of left;

start mbr;
do br = 1 to 24 by 1;
   start mvalue;
   do value = 1 to 4 by 1;
      start mrow;
      do row = 1 to 4 by 1;
         if left(|br,row |)= value then right_t(|br,value |)= row;
      end; finish;  run mrow;
   end; finish; run mvalue;
end; finish; run mbr;

print  left (|format=3.0(| mvalue right_t (|format=3.0 |);

 *Compute product  left * right_t;
  product = j(4,1,0);   table = j(24, 24, 0);
  start mlbr;
  do lbr = 1 to 24;
    start mrbr;
    do  rbr = 1 to 24;

    start mrow;
    do row = 1 to 4;
       temp = left(|lbr, row |);
       product(|row(| = right_t(| rbr, temp |);
    end;  finish;  run mrow;

    start mbr;
    do br = 1 to 24;
       if  left(|br,1(| = product(|1 |)&
           left(|br,2(| = product(|2 |)&
           left(|br,3(| = product(|3 |)&
           left(|br,4(| = product(|4 |)
       then  table(|lbr, rbr |)= br;
    end;  finish; run mbr;

   end; finish;  run mrbr;
end;  finish;  run mlbr;
print table (|format=2.0 |)"Table of M * M(transpose) Id";

 **************  end of  computation  of Gtable    ****************;
 */
```

```
U13293A.SASPROG.CNTL(TESTF)
VPSPRINT 5.1.002   WEDNESDAY NOVEMBER 4,1992   14:59:19 U13293A  MVS1        **
VPSPRINT SASPROG.CNTL(*) LOCAL FORM(9001) CLASS(E)

//U13293AA  JOB (*),'ROBERT WILSDN',
//  TIME=(1,0),MSGLEVEL=(1,1),MSGCLASS=X,CLASS=2
/*JOBPARM  ROOM=Y,FORMS=9001,COPIES=1
//S1  EXEC  SAS,OPTIONS='MACRO,DQUOTE'
//SYSIN  DD *
 *
 *
 * Program File:  sasprog.cntl(testf)
 *           Test F-distribution probability plots of osl's.
 * Steps:
 *     Generate normal distributed random variables.
 *     Square and sum to obtain 2 indep. chi-squared random variables.
 *     Divide by the degrees of freedom and each other to obtain
 *       f-distributed random variables.
 *     Obtain ranks.
 * Plot ranks vs osls.
 *
 *************** Debug Print Switch Macro and Switches   **********;
 * Macro for debug printing follows::
     %MACRO  db(db_site,db_act);
       %IF %UPCASE(&db_site) = TRUE  | %UPCASE(&db_site) = T
          %THEN   %QUOTE(&db_act);
     %MEND db;
 *
 *Debug print switches (dbps#):
  True or t = on (in upper and/or lower case), anything else = off.;
       %LET dbps1 = f      ; *Put data step values to log file;
       %LET dbps2 = t      ; *print final data set  6 obs;

*********************** End of Debug Print Switches ****************;

******************************************************************;

*Step One: Obtain f-distributed random variables and osl's.;

 DATA  f1(KEEP= f_var1 osl1 num_chi)
       N1(KEEP=var_num nor_var2);
    RETAIN seed1  20479   seed2 56193;
    DO  var_num= 1 TO 1000;
      num_chi = 0;
      DO  numdf = 1 TO 5  BY 1;
        *nor_var = RANNOR(5620479);
        CALL RANNOR(seed1   , nor_var1);
      num_chi = num_chi + nor_var1*nor_var1;
      END;

      den_chi = 0;
      DO  dendf = 1 TO 4  BY 1;
        *nor_var = RANNOR(562479);
        CALL RANNOR(seed2   , nor_var2);
        OUTPUT  n1;
      den_chi = den_chi + nor_var2*nor_var2;
      END;

      f_var1= (num_chi/5) / (den_chi/4 );

      osl1= 1 - PROBF( f_var1, 5, 4    );

    *Debut print switch follows;
```

```
        %db(&dbps1,  IF var_num < 6 THEN PUT
             nor_var1 nor_var2    num_chi= den_chi=   f_var1=  osl1=;);

           OUTPUT  f1;
      END;



  PROC STANDARD  DATA = N1   OUT=n2 std=1.0          vardef=n;
        by var_num;
        var  nor_var2;



  PROC MEANS DATA = n2 Uss  NOPRINT;
     by var_num;
     var  NOR_var2;
     output  out= n3  uss=den_chi2;

  data f1;  merge F1 n3;
        f_var2= (num_chi/5) / (den_chi2/4);

        osl2= 1 - PROBF( f_var2, 5, 4    );

        f_var3= (num_chi/5) / (den_chi2/ 3);

        osl3= 1 - PROBF( f_var3, 5, 3 );


     *Debut print switch follows;
      %db(&dbps2,  proc print  data=f1 (obs=6););


  ******************************************************************;
  *Rank osls. ;
  * Straight 45 degree line indicates actual f distribution.;

  PROC RANK  DATA = f1    OUT=f2
     TIES = MEAN  FRACTION;
     VAR     osl1     osl2     osl3;
     RANKS  r_osl1  r_osl2  r_osl3;

  DATA f2;    SET  f2;
     LABEL
        r_osl1  = "Rank central-OSL1"
        r_osl2  = "Rank central-OSL2"
        r_osl3  = "Rank central-OSL3"
      ;


  ******************************************************************;
  * Step:
     Analysize observed significance for test of f-distribution.;


  PROC PLOT  DATA = f2;
     TITLE1 "  Compare rank of osl with actual osl        ";
     TITLE2 "  Central F distribution  df(num)=5  df(den)=4   ";
     TITLE3 "  Normal errors not standardized prior to chi sq.";

     *Obs Sig Level for treatments, zero then nonzero ;
     PLOT  r_osl1  * osl1= '*'
          /  VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
             VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
             VREVERSE  HREVERSE   HPOS = 80   OVERLAY ;
```

```
      PLOT  r_os11 * os11 = '*'
            /  VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
               VREF  = 0 .05 .10 .15 .2    HREF  = 0 .05 .10 .15 .2
               VREVERSE  HREVERSE    HPOS = 80    OVERLAY ;


  PROC PLOT  DATA = f2;
    TITLE1 "  Compare rank of os1 with actual os1  not stand.";
    TITLE2 "  Central F distribution  df(num)=5  df(den)=4    ";
    TITLE3 "  Normal errors are standardized prior to chi sq.";

    *Obs Sig Level for treatments, zero then nonzero ;
    PLOT  r_os12   * os12= '*'
          /  VAXIS = 0 TO 1 BY .2       HAXIS = 0 TO 1 BY .2
             VREF  = 0 .2 .4 .6 .8 1.   HREF  = 0 .2 .4 .6 .8 1.
             VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

    PLOT  r_os12 * os12 = '*'
          /  VAXIS = 0 TO .2 BY .05     HAXIS = 0 TO .2 BY .05
             VREF  = 0 .05 .10 .15 .2   HREF  = 0 .05 .10 .15 .2
             VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

  PROC PLOT  DATA = f2;
    TITLE1 "  Compare rank of os1 with actual os1  not stand.";
    TITLE2 "  Central F distribution  df(num)=5  df(den)=3    ";
    TITLE3 "  Normal errors are standardized prior to chi sq.";

    *Obs Sig Level for treatments, zero then nonzero ;
    PLOT  r_os13   * os13= '*'
          /  VAXIS = 0 TO 1 BY .2       HAXIS = 0 TO 1 BY .2
             VREF  = 0 .2 .4 .6 .8 1.   HREF  = 0 .2 .4 .6 .8 1.
             VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

    PLOT  r_os13 * os13 = '*'
          /  VAXIS = 0 TO .2 BY .05     HAXIS = 0 TO .2 BY .05
             VREF  = 0 .05 .10 .15 .2   HREF  = 0 .05 .10 .15 .2
             VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;
//
  PROC CHART   DATA = f2;
       HBAR   f_var1 os11/ TYPE = PERCENT   LEVELS=20;
       TITLE1 "  Central F distribution  df(num)=3  df(den)=2)  ";

  PROC CHART   DATA = f2; WHERE f_var < 10000;
       HBAR f_var1  os11/ TYPE = PERCENT   LEVELS=20;
       TITLE1 "  Central F distribution  df(num)=3  df(den)=2)  ";

  PROC CHART   DATA = f2; WHERE f_var < 1000;
       HBAR f_var1  os11 / TYPE = PERCENT   LEVELS=20;
       TITLE1 "  Central F distribution  df(num)=3  df(den)=2)  ";

  PROC CHART   DATA = f2; WHERE f_var < 100;
       HBAR f_var1  os11/ TYPE = PERCENT   LEVELS=20;
       TITLE1 "  Central F distribution  df(num)=3  df(den)=2)  ";

  PROC CHART   DATA = f2; WHERE f_var < 10;
       HBAR f_var1  os11 / TYPE = PERCENT   LEVELS=20;
       TITLE1 "  Central F distribution  df(num)=3  df(den)=2    ";
//
```

```
U13293A.SASPROG.CNTL(SNORMAL)
VPSPRINT 5.1.002   WEDNESDAY NOVEMBER 4,1992   14:59:18 U13293A   MVS1          **
VPSPRINT SASPROG.CNTL(*) LOCAL FORM(9001) CLASS(E)

//U13293AA JOB (*),'ROBERT WILSON',TIME=(1,0),CLASS=2,
//  MSGCLASS=X,MSGLEVEL=(1,1)
/*JOBPARM  ROOM=Y,FORMS=9001,COPIES=1
//  EXEC SAS,REGION=918K
//SASGLM   DD  DSN=U13293A.SASGLM.DATA,DISP=OLD
//SASGLM2  DD  DSN=U13293A.SASGLM2.DATA,DISP=OLD
//SYSIN    DD *

 /*************************************************************
 *  File:  snormal   Singular Normal  estimate for sigma squared.
 *
 *  See Searle (1971) p. 69 Corollary 2s.2
 * Since L (below) does not have full column rank, the method of
 * Searle, p. 222, (132) and following, will not work.
 *
 *    The normally distributed errors were standardized to sum to
 *       zero within each randomization.  Prior to standardizing
 *       they were N(0, I*sigma-squared).  The standardizing involved
 *       pre-multiplying them by L=(I-(1/8)*J).  The resulting variance
 *       is V= L*L¢ = L since L is idempotent.  V is singular. The
 *       resulting distribution is singular normal, SN(0, V).  The
 *       residual sum of squares y¢*A*y is as before, but the degrees of
 *       of freedom are trace(AV), provided the) condition VAVAV=VAV is
 *       satisfied.                          below
 *
 *
 *
 *    By the method of construction, the experimental unit errors
 *       are the residuals from a least squares computation.  Thus,
 *       their distribution, assuming the errors of the model are normal,
 *       is the singular normal with V = LL' and L = I - Z*INV(Z¢*Z)*Z,
 *       as per below (2.3.3).  Thus, the residual sum of squares for
 *       models using only e.u. error is chi-squared with df=trace(AV),
 *       where A is the projector for the residual sum of squares.,
 *       A=I - P(X) - P(Rz) as below.
 *
 *
 *    For each randomization, this program verifies the condition on the
 *       theorem that VAVAV = VAV, calls in the sum of squares, computes
 *       the degrees of freedom = trace(AV), and outputs resutls.
 *
 *************************************************************/
proc iml;
 *Constant valued matrices follow:;
  xm  = {1, 1, 1, 1,   1, 1, 1,  1};   *Mean;
  xb  = {1, 1, 1, 1,  -1, -1, -1, -1};  *Blocks;
  z   = {1, 2, 11, 14,  5, 9, 16, 19};  *Covariates;
  j43 = J(4,3,0);

 *Constant projector follows:
  pz = z * INV(z¢*z) * z¢;

 *Output vector and variable names for output data set.
    sn_sigma = estimate of sigma squared using singular normal error.;
    escape = J(1, 6,  0);
    varname  = {exp_rand  max_val    min_val  tr_AV55
                sig55  sn_sig55};

 *  exp_rand =  experimental randomization id number;
```

```
*  max_va and min_va = max and min values of all elements of the
          difference matrix  VAVAV - VAV;
* tr_AV## = degrees of freedom for the variance estimate sn_sig##;
* sig_##   = variance estimate via glm using non-sing. normal df=2;
* sn_sig## = variance estimate using singular normal df=trace(AV);
* ## = O2 or O3 for the eq and ne block var. versions of e.u.   error;
* ## = 46 or 55 for the eq and ne block var. versions of normal error;


*Each row of Mpattern (randomization patterns) is one permitted
   block-randomization or assignment of treatments (element value)
   to experimental units (element position) within one block. ;
 rpattern =
   { 1 2 3 4 , 1 2 4 3 , 1 3 2 4 , 1 3 4 2 , 1 4 2 3 , 1 4 3 2,
     2 1 3 4 , 2 1 4 3 , 2 3 1 4 , 2 3 4 1 , 2 4 1 3 , 2 4 3 1,
     3 1 2 4 , 3 1 4 2 , 3 2 1 4 , 3 2 4 1 , 3 4 1 2 , 3 4 2 1,
     4 1 2 3 , 4 1 3 2 , 4 2 1 3 , 4 2 3 1 , 4 3 1 2 , 4 3 2 1 } ;


******************** begin looping through randomizations *****;
CREATE  temp  FROM  escape  (|COLNAME = varname |);

 START mb1_rand;
 DO b1_rand = 1 TO 24  BY 1;
    b1 = j43;
    START mp_set1;
    DO p_set1 = 1 TO 4 BY 1;    *1 row of rpattern;
       one_col = rpattern(|b1_rand, p_set1 |);
       IF one_col  = 4   THEN b1(|p_set1,   |)    = -1;
       IF one_col ¬= 4   THEN b1(|p_set1, one_col |)= 1;
    END;
    FINISH;
    RUN mp_set1;

    START mb2_rand;
    DO b2_rand = 1 TO 24  BY 1;
       *Compute experimental randomization number;
          exp_rand = (b1_rand - 1) # 24 + b2_rand;
       b2 = j43;
       START mp_set2;
       DO p_set2 = 1 TO 4 BY 1;
         one_col = rpattern(|b2_rand, p_set2 |);
         IF one_col  = 4 THEN b2(|p_set2,    |)    = -1;
         IF one_col ¬= 4 THEN b2(|p_set2, one_col |)= 1;
       END;
       FINISH;
       RUN mp_set2;

    *Construct full-model X matrix, X = xn ↑↑ xb ↑↑ xt_x ↑↑Z.;
       xt_s = b1 // b2 ;
       x_s = xm || xb || xt_s || z ;

    *Variable projectors follow:;
       px = x_s * INV(x_s¢ * x_s)* x_s¢;
       r  = (I(8) - px) * z;
       pr = r * GINV(r¢ * r) * r¢;

*PRINT exp_rand  px r pr;

    *Matrices A and V of Searle follow:;
       A = I(8) - px - pr;
      *For exper. unit. error, O2 & O3 ;
         *V = I(8) - px;
      *For normal dist. error, 46 & 55 ;
         V = ( I(4)-J(4,4,1/4)       ||     J(4,4,0)           )
          // (      J(4,4,0)         ||     I(4)-J(4,4,1/4)  );
```

```
              *Block diag form affects only ne block variances error sets;
          VAV = V * A * V;
          VAVAV = VAV * A * V;
          diff = VAVAV - VAV;
          max_val = MAX(diff);
          min_val = MIN(diff);
          tr_AV   = TRACE(A*V);


      *Read in sum of squares.;
          USE   sasglm2.glmO55;
          READ  NEXT 1  VAR{ssres}  INTO ssres;
      *Construct mean square estimator of variance.;
          sn_sig = ssres / tr_AV;
          sig = ssres / 2;

*PRINT exp_rand  A V  VAV  VAVAV  diff
      max_val min_val  tr_AV  sn_sig  sig;


      escape = exp_rand || max_val || min_val || tr_AV ||
                 sig ||  sn_sig;

    *Output results;
      SETOUT  temp;
      APPEND  FROM  escape;

   END;
   FINISH;
   RUN  mb2_rand;

  END;
  FINISH;
  RUN  mb1_rand;

***************************** END of IML ***********************;
TITLE1 "Singular normal and associated chi-square  of Searle, p. 69";
PROC MEANS  DATA=temp  N MIN MAX MEAN VAR  VARDEF=N;
   VAR tr_AV55  min_val  max_val sig55  sn_sig55;

PROC CHART DATA=temp;
   HBAR sn_sig55 sig55 tr_AV55  max_val/ TYPE = PERCENT    LEVELS = 24;

PROC PLOT DATA=temp;
    PLOT sn_sig55 * (sig55 tr_AV55  max_val) ;

*PROC PRINT  DATA=temp    (OBS=30);



DATA  sasglm2.sing_nor;
 MERGE    sasglm.sing_nor
          temp (KEEP = exp_rand  sn_sig55  tr_AV55);
 BY exp_rand;

PROC PRINT  DATA=sasglm2.sing_nor    (OBS=30);

//
```

```
U13293A.SASPROG.CNTL(PROBP)
VPSPRINT 5.1.002    WEDNESDAY NOVEMBER 4,1992    14:59:17 U13293A   MVS1          **
VPSPRINT SASPROG.CNTL(*) LOCAL FORM(9001) CLASS(E)

//U13293AA  JOB (*),'ROBERT WILSON',
//   TIME=(1,0),MSGLEVEL=(1,1),MSGCLASS=X,CLASS=2
/*JOBPARM  ROOM=Y,FORMS=9001,COPIES=1
//S1  EXEC  SAS,OPTIONS='MACRO,DQUOTE'
//SASGLM    DD  DSN=U13293A.SASGLM.DATA,DISP=OLD
//SASGLM2   DD  DSN=U13293A.SASGLM2.DATA,DISP=OLD
//SASDATA   DD  DSN=U13293A.SASDATA.DATA,DISP=OLD
//SYSIN DD *

    /*

 *
 * Program File:  sasprog.cntl(probp)
 *                Make F-distribution probability plots of osl's.
 * Step One:
 *     Select error types.  Extract f-ratios and null hypothesis
 *        osl's as provided by proc glm for each experimental
 *        randomizations. Extract the singular normal denominator
 *        mean square and degrees of freedom from sasglm2.sing_nor.
 *        These adjust e.u. and normal errors only.  The euti errors
 *        were not forced to sum to zero within each randomization.
 *
 *
 * Step Two:
 *     Produce experiment-wide variance for error terms.
 *     Atttach to data set fratios.   .
 *
 * Step Three:
 *    Use noncentrality parameters to obtain correct non-null
 *       hypothesis osl.
 *
 * Step Four:
 *    Rank all osls and attach to file fratios.
 *
 *
 * Step Five:
 *    Plot ranks vs actual osl values.
 *
 *
 * Output File:
 *   sasglm2.fratios
 *     One record per experimental randomization
 *

 ***************   Debug Print Switch Macro and Switches    **********;
 * Macro for debug printing follows:;
     %MACRO  db(db_site,db_act);
        %IF %UPCASE(&db_site) = TRUE  | %UPCASE(&db_site) = T
           %THEN   %QUOTE(&db_act);
     %MEND db;
 *
 *Debug print switches (dbps#):
  True or t = on (in upper and/or lower case), anything else = off.;
        %LET dbps1 = f    ;  *Extract frations from glm### data sets.;
        %LET dbps2 = f    ;  *Produce var of error terms - Step 2.;
        %LET dbps3 = f    ;  *Attach var(error) to fratios - Step 2.;
        %LET dbps4 = f    ;  *Attach lamda and noncent osl - Step 3.;
        %LET dbps5 = t    ;  *Attach rank(osl) to fratios - Step 4.;
        %LET dbps6 = f    ;  *;
```

```
*********************** End of Debug Print Switches ***************;

********************************************************************;

*Step One:
   Input f-ratios and osl's from various treatment-error combinations.
   Fz_mbt and oslz_mbt are = for glm0xx and glm1xx as nonzero treatments
   does not alter this central F statistic for the covariate.
   Key to new variable names:
      f=F-ratio   s=significance level    m=numerator mean square
      t=treatments z=covariate
      (for s only) c=central F    n=noncentral F
      0=zero treatment  1=nonzero treatment
      ##=error combination
      r prefix indicates rank of variable
      g suffix indicates f-ratio or sig level used singular normal.
                                          ;
DATA   sasglm2.fratios;
    MERGE
       sasglm2.sing_nor
       sasglm.glm002
             (KEEP= exp_rand  b1_rand  b2_rand  g_group
                    ft_mbz  oslt_mbz   fz_mbt  oslz_mbt
                    mst_mbz            msz_mbt
                RENAME = (ft_mbz=ft002  oslt_mbz=stc002   mst_mbz=mt002
                          fz_mbt=fz002  oslz_mbt=szc002   msz_mbt=mz002) )


       sasglm.glm102
             (KEEP= exp_rand
                    ft_mbz  oslt_mbz   fz_mbt  oslz_mbt
                    mst_mbz            msz_mbt
                RENAME = (ft_mbz=ft102  oslt_mbz=stc102   mst_mbz=mt102
                          fz_mbt=fz102  oslz_mbt=szc102   msz_mbt=mz102) )


       sasglm.glm003
             (KEEP= exp_rand
                    ft_mbz  oslt_mbz   fz_mbt  oslz_mbt
                    mst_mbz            msz_mbt
                RENAME = (ft_mbz=ft003  oslt_mbz=stc003   mst_mbz=mt003
                          fz_mbt=fz003  oslz_mbt=szc003   msz_mbt=mz003) )


       sasglm.glm103
             (KEEP= exp_rand
                    ft_mbz  oslt_mbz   fz_mbt  oslz_mbt
                    mst_mbz            msz_mbt
                RENAME = (ft_mbz=ft103  oslt_mbz=stc103   mst_mbz=mt103
                          fz_mbt=fz103  oslz_mbt=szc103   msz_mbt=mz103) )


       sasglm.glm004
             (KEEP= exp_rand
                    ft_mbz  oslt_mbz   fz_mbt  oslz_mbt
                RENAME = (ft_mbz=ft004  oslt_mbz=stc004
                          fz_mbt=fz004  oslz_mbt=szc004))

       sasglm.glm104
             (KEEP= exp_rand
                    ft_mbz  oslt_mbz   fz_mbt  oslz_mbt
                RENAME = (ft_mbz=ft104  oslt_mbz=stc104
                          fz_mbt=fz104  oslz_mbt=szc104))
```

```
    sasglm.glm005
        (KEEP= exp_rand
               ft_mbz   oslt_mbz    fz_mbt   oslz_mbt
          RENAME = (ft_mbz=ft005    oslt_mbz=stc005
                    fz_mbt=fz005    oslz_mbt=szc005))


    sasglm.glm105
        (KEEP= exp_rand
               ft_mbz   oslt_mbz    fz_mbt   oslz_mbt
          RENAME = (ft_mbz=ft105    oslt_mbz=stc105
                    fz_mbt=fz105    oslz_mbt=szc105))


    sasglm.glm007
        (KEEP= exp_rand
               ft_mbz   oslt_mbz    fz_mbt   oslz_mbt
          RENAME = (ft_mbz=ft007    oslt_mbz=stc007
                    fz_mbt=fz007    oslz_mbt=szc007))

    sasglm.glm107
        (KEEP= exp_rand
               ft_mbz   oslt_mbz    fz_mbt   oslz_mbt
          RENAME = (ft_mbz=ft107    oslt_mbz=stc107
                    fz_mbt=fz107    oslz_mbt=szc107))


    sasglm.glm046
        (KEEP= exp_rand
               ft_mbz   oslt_mbz    fz_mbt   oslz_mbt
               mst_mbz              msz_mbt
          RENAME = (ft_mbz=ft046  oslt_mbz=stc046    mst_mbz=mt046
                    fz_mbt=fz046  oslz_mbt=szc046    msz_mbt=mz046) )


    sasglm.glm146
        (KEEP= exp_rand
               ft_mbz   oslt_mbz    fz_mbt   oslz_mbt
               mst_mbz              msz_mbt
          RENAME = (ft_mbz=ft146  oslt_mbz=stc146    mst_mbz=mt146
                    fz_mbt=fz146  oslz_mbt=szc146    msz_mbt=mz146) )


    sasglm2.glm055
        (KEEP= exp_rand
               ft_mbz   oslt_mbz    fz_mbt   oslz_mbt
               mst_mbz              msz_mbt
          RENAME = (ft_mbz=ft055  oslt_mbz=stc055    mst_mbz=mt055
                    fz_mbt=fz055  oslz_mbt=szc055    msz_mbt=mz055) )


    sasglm2.glm155
        (KEEP= exp_rand
               ft_mbz   oslt_mbz    fz_mbt   oslz_mbt
               mst_mbz              msz_mbt
          RENAME = (ft_mbz=ft155  oslt_mbz=stc155    mst_mbz=mt155
                    fz_mbt=fz155  oslz_mbt=szc155    msz_mbt=mz155) )

    ;
    BY exp_rand;


*   Debug Print options follow: ;
    %db(&dbps1,
        TITLE1 "Data set  sasglm2.fratios   (dbps1)";
        PROC PRINT DATA = sasglm2.fratios   (OBS=26);
    );
```

```
******************************************************************;
*Step Two:
  Attach noncentrality parameters and within experimental randomization
     variance of error (not residual) term.;

PROC SORT  DATA = sasdata.lamda;  BY exp_rand;
DATA  sasglm2.fratios;
   MERGE  sasglm2.fratios
          sasdata.lamda
             (KEEP= exp_rand lam_dt lam_ft lam_st  lam_db lam_fb lam_gb
                    l_s);
   BY exp_rand;


*Since the eu errors are constant for all randomizations (but attached
  to different treatments) their error variance is constant and during
  construction was fixed at 30 (denom=n).  Likewise, the ndr errors had
  their variance fixed at 30 during construction.  The euti errors
  selected change with each randomization, thus their error variance
  changes with each randomization.  Their average, over all
  randomizations, is 30. For error conbinations included above,
  only 004, 005, and 007 involve euti error.  Non-zero treatments
  do not affect the error nor their variance.  For ease, all are
  attached.;


PROC MEANS  DATA = sasdata.obs_y ( KEEP= exp_rand   e02 -- e81s)
      VARDEF=N  NOPRINT;
      BY exp_rand;
      VAR  e02--e81;
      OUTPUT  OUT=sasglm2.err_var     VAR=var_e02-var_e81;

*   Debug Print options follow: ;
    %db(&dbps2,
        TITLE1 "Data set sasglm2.err_var -- var of errors (dbps2)";
        PROC PRINT DATA = sasglm2.err_var  (OBS=26) ;
    );



*Attach needed error (not residual) variances to fratios;
DATA  sasglm2.fratios;
   MERGE sasglm2.fratios
         sasglm2.err_var
            (KEEP= exp_rand var_e02  var_e03  var_e04   var_e05
                   var_e07  var_e46  var_e55);
   BY exp_rand;

*   Debug Print options follow: ;
    %db(&dbps3,
        TITLE1 "Data set sasglm2.fratios with var of errors (dbps3)";
        PROC PRINT DATA = sasglm2.fratios  (OBS=26);
    );



******************************************************************;
*Step Three:
  Use noncentrality parameters to obtain correct non-null
     hypothesis f-ratios for treatments and covariate.
  Use singular normal adjusted df to obtain adjusted error mean square,
     f-ratio, and osl for e.u. and normal error models.
  Variable szn0## = szni##  and  szc0## = szn0##.  The value of
     The treatments does not affect this observed significance level.;
```

```
DATA  sasglm2.fratios;    SET  sasglm2.fratios;
DROP  non_cent;
   *Compute the randomized observed significance level.;
   *SAS function PROBF appears not to accept non_cent over about 70
    for these f-ratios and degrees of freedom.;


   *Error 02;
    non_cent = (1/2) / var_e02 * lam_dt;
    IF non_cent > 70   THEN  non_cent = 70;
   *stc002  as per PROC GLM;
    stc002g= 1 - PROBF( mt002/sn_sig02, 3, tr_AV02                );
   *stc102  as per PROC GLM;
    stc102g= 1 - PROBF( mt102/sn_sig02, 3, tr_AV02                );
    stn102 = 1 - PROBF( ft102            , 3, 2,          non_cent );
    stn102g= 1 - PROBF( mt102/sn_sig02, 3, tr_AV02,    non_cent );

    non_cent = (1/2) / var_e02 * lam_db;
    IF non_cent > 70   THEN  non_cent = 70;
   *ms_z, ms_res and tr_AV## are the same for treat=0 or ne 0.;
   *tr_AV## is the same for equal and unequal block variances.;
   *Thus, szc0## = szc1##  and szn0## = szn1## - only 0 used.;
   *szc002  as per PROC GLM;
    szc002g= 1 - PROBF( mz002/sn_sig02, 1, tr_AV02                );
    szn002 = 1 - PROBF( fz002,          1, 2,          non_cent );
    szn002g= 1 - PROBF( mz002/sn_sig02, 1, tr_AV02,    non_cent );


   *Error 03;
    non_cent = (1/2) / var_e03 * lam_dt;
    IF non_cent > 70   THEN  non_cent = 70;
   *stc003  as per PROC GLM;
    stc003g= 1 - PROBF( mt003/sn_sig03, 3, tr_AV03                );
   *stc103  as per PROC GLM;
    stc103g= 1 - PROBF( mt103/sn_sig03, 3, tr_AV03                );
    stn103 = 1 - PROBF( ft103            , 3, 2,          non_cent );
    stn103g= 1 - PROBF( mt103/sn_sig03, 3, tr_AV03,    non_cent );

    non_cent = (1/2) / var_e03 * lam_db;
    IF non_cent > 70   THEN  non_cent = 70;
   *szc003  as per PROC GLM;
    szc003g= 1 - PROBF( mz003/sn_sig03, 1, tr_AV03                );
    szn003 = 1 - PROBF( fz003,          1, 2,          non_cent );
    szn003g= 1 - PROBF( mz003/sn_sig03, 1, tr_AV03,    non_cent );


   *Error 04;
    non_cent = (1/2) / var_e04 * lam_dt;
    IF non_cent > 70   THEN  non_cent = 70;
    stn104 = 1 - PROBF( ft104, 3, 2, non_cent );

    non_cent = (1/2) / var_e04 * lam_db;
    IF non_cent > 70   THEN  non_cent = 70;
    szn004 = 1 - PROBF( fz004, 1, 2, non_cent );


   *Error 05;
    non_cent = (1/2) / var_e05 * lam_dt;
    IF non_cent > 70   THEN  non_cent = 70;
    stn105 = 1 - PROBF( ft105, 3, 2, non_cent );

    non_cent = (1/2) / var_e05 * lam_db;
    IF non_cent > 70   THEN  non_cent = 70;
    szn005 = 1 - PROBF( fz005, 1, 2, non_cent );
```

```
*Error 07;
 non_cent = (1/2) / var_e07 * lam_dt;
 IF non_cent > 70    THEN  non_cent = 70;
 stn107 = 1 - PROBF( ft107, 3, 2, non_cent );

 non_cent = (1/2) / var_e07 * lam_db;
 IF non_cent > 70    THEN  non_cent = 70;
 szn007 = 1 - PROBF( fz007, 1, 2, non_cent );


*Error 46;       *Test different df in F, and in ANOVA;
 non_cent = (1/2) / var_e46 * lam_dt;
 IF non_cent > 70    THEN  non_cent = 70;
*stc046  as per PROC GLM;
 stc046g= 1 - PROBF( mt046/sn_sig46, 3, tr_AV46           );
*stc146  as per PROC GLM;
 stc146g= 1 - PROBF( mt146/sn_sig46, 3, tr_AV46           );
 stn146 = 1 - PROBF( ft146         , 3, 2,      non_cent );
 stn146g= 1 - PROBF( mt146/sn_sig46, 3, tr_AV46,  non_cent );

 non_cent = (1/2) / var_e46 * lam_db;
 IF non_cent > 70    THEN  non_cent = 70;
*szc046  as per PROC GLM;
 szc046g= 1 - PROBF( mz046/sn_sig46, 1, tr_AV46           );
 szn046 = 1 - PROBF( fz046,          1, 2,      non_cent );
 szn046g= 1 - PROBF( mz046/sn_sig46, 1, tr_AV46,  non_cent );


*Error 55;
 non_cent = (1/2) / var_e55 * lam_dt;
 IF non_cent > 70    THEN  non_cent = 70;
*stc055  as per PROC GLM;
 stc055g= 1 - PROBF( mt055/sn_sig55, 3, tr_AV55           );
*stc155  as per PROC GLM;
 stc155g= 1 - PROBF( mt155/sn_sig55, 3, tr_AV55           );
 stn155 = 1 - PROBF( ft155         , 3, 2,      non_cent );
 stn155g= 1 - PROBF( mt155/sn_sig55, 3, tr_AV55,  non_cent );

 non_cent = (1/2) / var_e55 * lam_db;
 IF non_cent > 70    THEN  non_cent = 70;
*szc055  as per PROC GLM;
 szc055g= 1 - PROBF( mz055/sn_sig55, 1, tr_AV55           );
 szn055 = 1 - PROBF( fz055,          1, 2,      non_cent );
 szn055g= 1 - PROBF( mz055/sn_sig55, 1, tr_AV55,  non_cent );

*   Debug Print options follow: ;
    %db(&dbps4,
        TITLE1 "Data set sasglm2.fratios with noncent osls (dbps4)";
      * PROC PRINT DATA = sasglm2.fratios  (OBS=26);
      );


*****************************************************************************;
* Step Four:
    Names of osls.
        Only e.u. and normal errors have singular-normal based osls.
        The ranks are named the same, but with a prefix of r.

        stcO##  for treats,        normal, central F, treatments O
        stcO##g for treats, sing-normal, central F, treatments O

        stc1##  for treats,        normal, central F, treatments ne O
        stc1##g for treats, sing-normal, central F, treatments ne O

        stn1##  for treats,        normal, noncentral F, treatments O
        stn1##g for treats, sing-normal, noncentral F, treatments O
```

```
        szcO##  for covars,         normal, central F,  treats O or ne O
      szcO##g for covars, sing-normal, central F, treats O or ne O

        sznO##  for covars,         normal, central F, treats O or ne O
        sznO##g for covars, sing-normal, central F, treats O or ne O



        Straight 45 degree line indicates actual f distribution.;

PROC RANK  DATA = sasglm2.fratios   OUT=sasglm2.fratios
   TIES = MEAN   FRACTION;
   VAR
        stc002    stc102    stn102      szc002    szn002
        stc002g   stc102g   stn102g     szc002g   szn002g

        stc003    stc103    stn103      szc003    szn003
        stc003g   stc103g   stn103g     szc003g   szn003g

        stc004    stc104    stn104      szc004    szn004
        stc005    stc105    stn105      szc005    szn005
        stc007    stc107    stn107      szc007    szn007

        stc046    stc146    stn146      szc046    szn046
        stc046g   stc146g   stn146g     szc046g   szn046g

        stc055    stc155    stn155      szc055    szn055
        stc055g   stc155g   stn155g     szc055g   szn055g
            :
   RANKS
        rstc002   rstc102   rstn102     rszc002   rszn002
        rstc002g  rstc102g  rstn102g    rszc002g  rszn002g

        rstc003   rstc103   rstn103     rszc003   rszn003
        rstc003g  rstc103g  rstn103g    rszc003g  rszn003g

        rstc004   rstc104   rstn104     rszc004   rszn004
        rstc005   rstc105   rstn105     rszc005   rszn005
        rstc007   rstc107   rstn107     rszc007   rszn007

        rstc046   rstc146   rstn146     rszc046   rszn046
        rstc046g  rstc146g  rstn146g    rszc046g  rszn046g

        rstc055   rstc155   rstn155     rszc055   rszn055
        rstc055g  rstc155g  rstn155g    rszc055g  rszn055g
            :
*   Debug Print options follow: ;
    %db(&dbps5,
        TITLE1 "Data set sasglm2.fratios final      (dbps5)";
        PROC PRINT DATA = sasglm2.fratios  (OBS=26);
    );


DATA sasglm2.fratios;    SET  sasglm2.fratios;
   LABEL
        rstc002  = "Rank cent-OSL t|rest, treat eq O"
        rstc002g = "Rank cent-OSL t|rest, treat eq O,sig-nor"

        rstc003  = "Rank cent-OSL t|rest, treat eq O"
        rstc003g = "Rank cent-OSL t|rest, treat eq O,sig-nor"

        rstc004  = "Rank cent-OSL t|rest, treat eq O"
        rstc005  = "Rank cent-OSL t|rest, treat eq O"
        rstc007  = "Rank cent-OSL t|rest, treat eq O"
```

```
rstc046   = "Rank cent-OSL t|rest, treat eq 0"
rstc046g  = "Rank cent-OSL t|rest, treat eq 0,sig-nor"

rstc055   = "Rank cent-OSL t|rest, treat eq 0"
rstc055g  = "Rank cent-OSL t|rest, treat eq 0,sig-nor"


rstc102   = "Rank cent-OSL t|rest, treat ne 0"
rstc102g  = "Rank cent-OSL t|rest, treat ne 0,sig-nor"

rstc103   = "Rank cent-OSL t|rest, treat ne 0"
rstc103g  = "Rank cent-OSL t|rest, treat ne 0,sig-nor"

rstc104   = "Rank cent-OSL t|rest, treat ne 0"
rstc105   = "Rank cent-OSL t|rest, treat ne 0"
rstc107   = "Rank cent-OSL t|rest, treat ne 0"

rstc146   = "Rank cent-OSL t|rest, treat ne 0"
rstc146g  = "Rank cent-OSL t|rest, treat ne 0,sig-nor"

rstc155   = "Rank cent-OSL t|rest, treat ne 0"
rstc155g  = "Rank cent-OSL t|rest, treat ne 0,sig-nor"


rstn102   = "Rank nonc-OSL t|rest, treat ne 0"
rstn102g  = "Rank nonc-OSL t|rest, treat ne 0,sig-nor"

rstn103   = "Rank nonc-OSL t|rest, treat ne 0"
rstn103g  = "Rank nonc-OSL t|rest, treat ne 0,sig-nor"

rstn104   = "Rank nonc-OSL t|rest, treat ne 0"
rstn105   = "Rank nonc-OSL t|rest, treat ne 0"
rstn107   = "Rank nonc-OSL t|rest, treat ne 0"

rstn146   = "Rank nonc-OSL t|rest, treat ne 0"
rstn146g  = "Rank nonc-OSL t|rest, treat ne 0,sig-nor"

rstn155   = "Rank nonc-OSL t|rest, treat ne 0"
rstn155g  = "Rank nonc-OSL t|rest, treat ne 0,sig-nor"


rszc002   = "Rank cent-OSL z|rest, covar ne 0"
rszc002g  = "Rank cent-OSL z|rest, covar ne 0,sig-nor"

rszc003   = "Rank cent-OSL z|rest, covar ne 0"
rszc003g  = "Rank cent-OSL z|rest, covar ne 0,sig-nor"

rszc004   = "Rank cent-OSL z|rest, covar ne 0"
rszc005   = "Rank cent-OSL z|rest, covar ne 0"
rszc007   = "Rank cent-OSL z|rest, covar ne 0"

rszc046   = "Rank cent-OSL z|rest, covar ne 0"
rszc046g  = "Rank cent-OSL z|rest, covar ne 0,sig-nor"

rszc055   = "Rank cent-OSL z|rest, covar ne 0"
rszc055g  = "Rank cent-OSL z|rest, covar ne 0,sig-nor"


rszn002   = "Rank nonc-OSL z|rest, covar ne 0"
rszn002g  = "Rank nonc-OSL z|rest, covar ne 0,sig-nor"

rszn003   = "Rank nonc-OSL z|rest, covar ne 0"
rszn003g  = "Rank nonc-OSL z|rest, covar ne 0,sig-nor"

rszn004   = "Rank nonc-OSL z|rest, covar ne 0"
rszn005   = "Rank nonc-OSL z|rest, covar ne 0"
```

```
        rszn007  = "Rank nonc-OSL z|rest, covar ne 0"

        rszn046  = "Rank nonc-OSL z|rest, covar ne 0"
        rszn046g = "Rank nonc-OSL z|rest, covar ne 0,sig-nor"

        rszn055  = "Rank nonc-OSL z|rest, covar ne 0"
        rszn055g = "Rank nonc-OSL z|rest, covar ne 0,sig-nor"
    ;

 */
 /*
********************************************************************;
* Step Five:
    Analysize observed significance for test of f-distribution.;


PROC PLOT  DATA = sasglm2.fratios;
   TITLE1 "          Compare Three Types of Error          ";
   TITLE2 "u = experimental unit error                     ";
   TITLE3 "i = experimental unit-treatment interaction error";
   TITLE4 "n = normally distributed random error           ";

  ********** Treatments   Equal Block Variances *************;
  *Non-singular normal values for eu and normal error;
   *Obs Sig Level for treatments eq zero;
   PLOT  rstc002 * stc002 = 'u'
         rstc004 * stc004 = 'i'
         rstc046 * stc046 = 'n'
         / VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
           VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
           VREVERSE  HREVERSE·  HPOS = 80    OVERLAY ;

   PLOT  rstc002 * stc002 = 'u'
         rstc004 * stc004 = 'i'
         rstc046 * stc046 = 'n'
         / VAXIS = 0 TO .2 BY .05     HAXIS = 0 TO .2 BY .05
           VREF  = 0 .05 .10 .15 .2   HREF  = 0 .05 .10 .15 .2
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


   *Obs Sig Level for treatments ne zero;
   PLOT  rstni102 * stni102 = 'u'
         rstni104 * stni104 = 'i'
         rstni146 * stni146 = 'n'
         / VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
           VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


   PLOT  rstni102 * stni102 = 'u'
         rstni104 * stni104 = 'i'
         rstni146 * stni146 = 'n'
         / VAXIS = 0 TO .2 BY .05     HAXIS = 0 TO .2 BY .05
           VREF  = 0 .05 .10 .15 .2   HREF  = 0 .05 .10 .15 .2
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


   *Singular normal values for eu and normal error;
   *Obs Sig Level for treatments eq zero;
   PLOT  rstc002g* stc002g= 'u'
         rstc004 * stc004 = 'i'
         rstc046g* stc046g= 'n'
         / VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
           VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;
```

```
  PLOT   rstc002g* stc002g= 'u'
         rstc004 * stc004 = '1'
         rstc046g* stc046g= 'n'
         / VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
             VREF  = 0 .05 .10 .15 .2   HREF  = 0 .05 .10 .15 .2
             VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


  *Obs Sig Level for treatments ne zero;
  PLOT   rstn102g* stn102g= 'u'
         rstn104 * stn104 = '1'
         rstn146g* stn146g= 'n'
         / VAXIS = 0 TO 1 BY .2        HAXIS = 0 TO 1 BY .2
             VREF  = 0 .2 .4 .6 .8 1.   HREF  = 0 .2 .4 .6 .8 1.
             VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


  PLOT   rstn102g* stn102g= 'u'
         rstn104 * stn104 = '1'
         rstn146g* stn146g= 'n'
         / VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
             VREF  = 0 .05 .10 .15 .2   HREF  = 0 .05 .10 .15 .2
             VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;



  *********** Covariate   Equal Block Variances **************;
  *Non-singular normal values for eu and normal error;
   *Obs Sig Level for (always nonzero) covariate coefficient  ;
   PLOT  rszn002 * szn002 = 'u'
         rszn004 * szn004 = '1'
         rszn046 * szn046 = 'n'
         / VAXIS = 0 TO 1 BY .2        HAXIS = 0 TO 1 BY .2
             VREF  = 0 .2 .4 .6 .8 1.   HREF  = 0 .2 .4 .6 .8 1.
             VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


  PLOT   rszn002 * szn002 = 'u'
         rszn004 * szn004 = '1'
         rszn046 * szn046 = 'n'
         / VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
             VREF  = 0 .05 .10 .15 .2   HREF  = 0 .05 .10 .15 .2
             VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


  *Singular normal values for eu and normal error;
   *Obs Sig Level for (always nonzero) covariate coefficient  ;
   PLOT  rszn002g* szn002g= 'u'
         rszn004 * szn004 = '1'
         rszn046g* szn046g= 'n'
         / VAXIS = 0 TO 1 BY .2        HAXIS = 0 TO 1 BY .2
             VREF  = 0 .2 .4 .6 .8 1.   HREF  = 0 .2 .4 .6 .8 1.
             VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


  PLOT   rszn002g* szn002g= 'u'
         rszn004 * szn004 = '1'
         rszn046g* szn046g= 'n'
         / VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
             VREF  = 0 .05 .10 .15 .2   HREF  = 0 .05 .10 .15 .2
             VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;



PROC PLOT  DATA = sasglm2.fratios;
  TITLE1 "Effect of Unequal Within-block Variances";
```

```
   TITLE2 "s = same block variance   d = different block variances";

********************* Treatments *****************************;
*Only experimental unit error   treatments zero &  nonzero;
*Non-singular normal;
 PLOT   rstc002 * stc002 = 's'
        rstc003 * stc003 = 'd'
        /  VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
           VREF  = 0 .2 .4 .6.8 1.   HREF  = 0 .2 .4 .6 .8 1.
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

 PLOT   rstc002 * stc002 = 's'
        rstc003 * stc003 = 'd'
        /  VAXIS = 0 TO .2 BY .05    HAXIS = 0 TO .2 BY .05
           VREF  = 0 .05 .10 .15 .2  HREF  = 0 .05 .10 .15 .2
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


 PLOT   rstn102 * stn102 = 's'
        rstn103 * stn103 = 'd'
        /  VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
           VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

 PLOT   rstn102 * stn102 = 's'
        rstn103 * stn103 = 'd'
        /  VAXIS = 0 TO .2 BY .05    HAXIS = 0 TO .2 BY .05
           VREF  = 0 .05 .10 .15 .2  HREF  = 0 .05 .10 .15 .2
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

*Singular normal;
 PLOT   rstc002g* stc002g= 's'
        rstc003g* stc003g= 'd'
        /  VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
           VREF  = 0 .2 .4 .6.8 1.   HREF  = 0 .2 .4 .6 .8 1.
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

 PLOT   rstc002g* stc002g= 's'
        rstc003g* stc003g= 'd'
        /  VAXIS = 0 TO .2 BY .05    HAXIS = 0 TO .2 BY .05
           VREF  = 0 .05 .10 .15 .2  HREF  = 0 .05 .10 .15 .2
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


 PLOT   rstn102g* stn102g= 's'
        rstn103g* stn103g= 'd'
        /  VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
           VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

 PLOT   rstn102g* stn102g= 's'
        rstn103g* stn103g= 'd'
        /  VAXIS = 0 TO .2 BY .05    HAXIS = 0 TO .2 BY .05
           VREF  = 0 .05 .10 .15 .2  HREF  = 0 .05 .10 .15 .2
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


*Only experimental unit-treat inteact error  treats zero &  nonzero;
*Non-singular normal  no adjustments imply the singular normal dist.;
 PLOT   rstc004 * stc004 = 's'
        rstc007 * stc007 = 'd'
        /  VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
           VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
           VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

 PLOT   rstc004 * stc004 = 's'
```

```
            rstc007 * stc007 = 'd'
            /  VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
               VREF  = 0 .05 .10 .15 .2    HREF  = 0 .05 .10 .15 .2
               VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


   PLOT  rstn104 * stn104 = 's'
         rstn107 * stn107 = 'd'
         /  VAXIS = 0 TO 1 BY .2       HAXIS = 0 TO 1 BY .2
            VREF  = 0 .2 .4 .6 .8 1.   HREF  = 0 .2 .4 .6 .8 1.
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

   PLOT  rstn104 * stn104 = 's'
         rstn107 * stn107 = 'd'
         /  VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
            VREF  = 0 .05 .10 .15 .2    HREF  = 0 .05 .10 .15 .2
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


*Only normally distributed random error  treats zero &  nonzero;
*Non-singular normal;
 PLOT  rstc046 * stc046 = 's'
       rstc055 * stc055 = 'd'
       /  VAXIS = 0 TO 1 BY .2       HAXIS = 0 TO 1 BY .2
          VREF  = 0 .2 .4 .6 .8 1.   HREF = 0 .2 .4 .6 .8 1.
          VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

   PLOT  rstc046 * stc046 = 's'
         rstc055 * stc055 = 'd'
         /  VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
            VREF  = 0 .05 .10 .15 .2    HREF  = 0 .05 .10 .15 .2
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


   PLOT  rstn146 * stn146 = 's'
         rstn155 * stn155 = 'd'
         /  VAXIS = 0 TO 1 BY .2       HAXIS = 0 TO 1 BY .2
            VREF  = 0 .2 .4 .6 .8 1.   HREF  = 0 .2 .4 .6 .8 1.
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

   PLOT  rstn146 * stn146 = 's'
         rstn155 * stn155 = 'd'
         /  VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
            VREF  = 0 .05 .10 .15 .2    HREF  = 0 .05 .10 .15 .2
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

*Singular normal;
 PLOT  rstc046g* stc046g= 's'
       rstc055g* stc055g= 'd'
       /  VAXIS = 0 TO 1 BY .2       HAXIS = 0 TO 1 BY .2
          VREF  = 0 .2 .4 .6 .8 1.   HREF  = 0 .2 .4 .6 .8 1.
          VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

   PLOT  rstc046g* stc046g= 's'
         rstc055g* stc055g= 'd'
         /  VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
            VREF  = 0 .05 .10 .15 .2    HREF  = 0 .05 .10 .15 .2
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;


   PLOT  rstn146g* stn146g= 's'
         rstn155g* stn155g= 'd'
         /  VAXIS = 0 TO 1 BY .2       HAXIS = 0 TO 1 BY .2
            VREF  = 0 .2 .4 .6 .8 1.   HREF  = 0 .2 .4 .6 .8 1.
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;
```

```
   PLOT   rstn146g* stn146g= 's'
          rstn155g* stn155g= 'd'
          /  VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
             VREF  = 0 .05 .10 .15 .2    HREF  = 0 .05 .10 .15 .2
             VREVERSE  HREVERSE    HPOS = 80    OVERLAY ;

******************** Covariates *********************************;
*Only experimental unit error  for covariate;
*Non-singular normal;
  PLOT   rszn002 * szn002 = 's'
         rszn003 * szn003 = 'd'
         /  VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
            VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

  PLOT   rszn002 * szn002 = 's'
         rszn003 * szn003 = 'd'
         /  VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
            VREF  = 0 .05 .10 .15 .2    HREF  = 0 .05 .10 .15 .2
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;
*Singular normal;
  PLOT   rszn002g* szn002g= 's'
         rszn003g* szn003g= 'd'
         /  VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
            VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

  PLOT   rszn002g* szn002g= 's'
         rszn003g* szn003g= 'd'
         /  VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
            VREF  = 0 .05 .10 .15 .2    HREF  = 0 .05 .10 .15 .2
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

*Only experimental unit-treat inteact error  treats zero &  nonzero;
*Non-singular normal;
  PLOT   rszn004 * szn004 = 's'
         rszn007 * szn007 = 'd'
         /  VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
            VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

  PLOT   rszn004 * szn004 = 's'
         rszn007 * szn007 = 'd'
         /  VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
            VREF  = 0 .05 .10 .15 .2    HREF  = 0 .05 .10 .15 .2
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

*Only normally distributed random error  treats zero &  nonzero;
*Non-singular normal;
  PLOT   rszn046 * szn046 = 's'
         rszn055 * szn055 = 'd'
         /  VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
            VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;

  PLOT   rszn046 * szn046 = 's'
         rszn055 * szn055 = 'd'
         /  VAXIS = 0 TO .2 BY .05      HAXIS = 0 TO .2 BY .05
            VREF  = 0 .05 .10 .15 .2    HREF  = 0 .05 .10 .15 .2
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;
*Singular normal;
  PLOT   rszn046g* szn046g= 's'
         rszn055g* szn055g= 'd'
         /  VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
            VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
            VREVERSE  HREVERSE   HPOS = 80    OVERLAY ;
```

```
   PLOT   rszn046g* szn046g= 's'
          rszn055g* szn055g= 'd'
           / VAXIS = O TO .2 BY .05     HAXIS = O TO .2 BY .05
             VREF  = O .05 .10 .15 .2   HREF  = O .05 .10 .15 .2
             VREVERSE  HREVERSE    HPOS = 80    OVERLAY ;

*/
/*

PROC SORT  DATA = sasglm2.fratios;  BY  g_group;
PROC PLOT  DATA = sasglm2.fratios
           UNIFORM  HPERCENT=33 33 33    VPERCENT=50 50;
  BY g_group;

  TITLE1 "Observed Significance Levels for Treatments and Covariates";
  TITLE2 "u = experimental unit error                    ";
  TITLE3 "i = experimental unit-treatment interaction error";
  TITLE4 "n = normally distributed random error          ";
  PLOT   stc002 * szc002 = 'u'
         stc004 * szc004 = 'i'
         stc046 * szc046 = 'n' / OVERLAY  HAXIS = O to .2 by .05;

  PLOT   stc102 * szc002 = 'u'
         stc104 * szc004 = 'i'
         stc146 * szc046 = 'n' / OVERLAY  HAXIS = O to .2 by .05;

  PLOT   stn102 * szc002 = 'u'
         stn104 * szc004 = 'i'
         stn146 * szc046 = 'n' / OVERLAY  HAXIS = O to .2 by .05;


  PLOT   stc002g* szc002g= 'u'
         stc004 * szc004 = 'i'
         stc046g* szc046g= 'n' / OVERLAY  HAXIS = O to .2 by .05;

  PLOT   stc102g* szc002g= 'u'
         stc104 * szc004 = 'i'
         stc146g* szc046g= 'n' / OVERLAY  HAXIS = O to .2 by .05;

  PLOT   stn102g* szc002g= 'u'           .
         stn104 * szc004 = 'i'
         stn146g* szc046g= 'n' / OVERLAY  HAXIS = O to .2 by .05;

  */
PROC SORT  DATA = sasglm2.fratios;  BY lam_db;
PROC PLOT  DATA = sasglm2.fratios
           UNIFORM  HPERCENT=33 33 33    VPERCENT=50 50;
  BY lam_db;

  TITLE1 "Actual Upper Tail Areas for Treatments and Covariates";
  TITLE2 "Treatments are Zero in Top Row, Nonzero in Bottom Row     ";
  TITLE3 "Left to Right: experimental unit error,   experimental-unit";
  TITLE4 "  treatment interaction error,  normal random error       ";
  PLOT   stc002 * szn002
         stc004 * szn004
         stc046 * szn046

         stn102 * szn002
         stn104 * szn004
         stn146 * szn046
         / HAXIS = O .1 .25 .5 .75 1.0  VAXIS = O .1 .25 .5  .75  1.0
           HREF  =   .1 .25 .5 .75      VREF  =   .1 .25.5  .75
           BOX
           ;
```

```
//
   PLOT    stc002 * szc002 = 'u'
           stc004 * szc004 = '1'
           stc046 * szc046 = 'n'

           stc102 * szc002 = 'u'
           stc104 * szc004 = '1'
           stc146 * szc046 = 'n'
           /   HAXIS = 0 TO .5 BY .25    VAXIS = 0 TO 1 BY .25
               HREF  = 0 TO .5 BY .25    VREF  = 0 TO 1 BY .25
           ;
   *Central vs non-central osls;
   PLOT   stc102 * stn102 = 'u'
          stc104 * stn104 = '1'
          stc146 * stn146 = 'n'
          / VAXIS = 0 TO 1 BY .2      HAXIS = 0 TO 1 BY .2
            VREF  = 0 .2 .4 .6 .8 1.  HREF  = 0 .2 .4 .6 .8 1.
            VREVERSE  HREVERSE   HPOS = 80   OVERLAY ;


   PLOT   stc102 * stn102 = 'u'
          stc104 * stn104 = '1'
          stc146 * stn146 = 'n'
          / VAXIS = 0 TO .2 BY .05     HAXIS = 0 TO .2 BY .05
            VREF  = 0 .05 .10 .15 .2   HREF  = 0 .05 .10 .15 .2
            VREVERSE  HREVERSE   HPOS = 80   OVERLAY ;
```

APPENDIX G

OTHER TOPICS

Dissertation work began in the Summer of 1983 under Dr. William Stewart. The topic was calibration with a stratified sample from a finite population. The initial task was a bibliography covering 1920 to 1984 with some 350 entries. This topic was abandoned in the Fall of 1984 as Dr. Stewart, by then employed in the private sector, was unable to continue as advisor.

The second topic was the derivation shown on the following two pages. After obtaining the result, a literature search showed that the bulk of the proof had been published prior to the 1940's and that the method of moments had been replaced by k-statistics. Further effort with moments appeared unlikely to contribute to the discipline. This work took about a year and was abandoned in the Fall of 1985. Dr. Folks presented the result at the 1985 annual ISI meeting in Amsterdam.

The next topic was to use k-statistics to obtain approximate estimators of model equation parameters in the analysis of covariance, much as Cook (1951B) did for the parameter of simple linear regression. For this, one needs to express higher order, multivariate cumulants κ, in terms of multivariate moments μ. Harvey [1972] published recursive relationships which will generate the formula equating any order of a multivariate κ or μ to the other. His formulas have been programed in the symbolic manipulation language MUSIMP. Following the results of the second topic are two pages of output from this program. They show formula equating each of the bivariate cumulants of order eight, $\kappa_{8,0}$, $\kappa_{1,7}$, etc., to the bivariate moments $\mu_{8,0}$, $\mu_{7,1}$, etc. In the formula, for example, (6 U51 U01^2) indicates $6*\mu_{5,1}*(\mu_{0,1})^2$; multiplication is indicated by a blank and exponentiation by a caret. Here, $\mu_{5,1}$ is the bivariate, (X_1, X_2), expectation of $(X_1)^5*(X_2)$. Learning this programming language, writing the program, and financing the same took two years.

The next step would have been to derive formula equating the sample cumulants k to the finite population cumulants κ. The methods of Cook [1951A] appear difficult to program and would need to be modified for finite populations. Bell polynomials, tensors, and polykeys were examined. The first two appear to have

been derived; see Sugihara and Murota, [Utilitas Mathematics, Vol. 22 (1982), p. 265-291], and McCullagh [1987].   The difficulty of programming unfamiliar operators, and the uncertainty of the best method to use, reluctantly led to abandoning this topic in late 1987.

   Mid-1988 saw results on one unit error, and Cox's ratio.   Mid-1989 saw the general randomization model for experimental unit-treatment interaction error and an essay on the early understanding of the error term. Mid-1990 saw the proof of the theorem on projectors and a procedure when both errors are present.   The linkage of the noncentrality parameters proved difficult to unravel, and was solved only in 1992.   Organizing and writing the work, and, as always, financing it have been the major tasks since.

Regression Models for Finite Populations

J. Leroy Folks and Robert S. Wilson

Department of Statistics, Oklahoma State University, Stillwater, Oklahoma, 74078, USA

Sommaire

Données Y populations définies de dimension N à chaque valeur X, des simples échantillons choisis au hasard de dimensions n sont pris de chaque population Y. La ligne des enchantillons en écarts carrés a un écart à la moyenne pour la ligne de la population en écarts carrés. Les variances de la courbe et de l'interception des échantillons sont estimées par écart à la moyenne et les variances de ces derniers estimateurs sont données.

1 The population regression model

. For each of L distinct X values we have a finite population of Ys of size N and have fitted a least squares line to these LN (X,Y) pairs. Denote the intercept and slope by α and β so that

$$Y_{hj} = \alpha + \beta X_h + e_{hj},$$

$$h = 1, 2, \ldots, L$$

$$j = 1, 2, \ldots, N$$

2 Estimation of α and β

Assume that at each X_h we take a simple random sample of n from the N Ys, denoted by $\{y_{hj} \mid j = 1, 2, \ldots, n\}$. Fit a least squares line to these Ln (X,y) pairs and denote the intercept and slope by a and b. It follows from Jönrop and Rennermalm (1976) that

$$E(a) = \alpha \text{ and } E(b) = \beta.$$

3 Estimation of variances

The variances of a and b have forms analogous to the usual expressions.

$$V(a) = \frac{N-n}{nN} \sum_h \left\{ \frac{1}{L} - \frac{(X_h - \bar{X})\bar{X}}{\Sigma(X_h - \bar{X})^2} \right\}^2 s_h^2$$

and

$$V(b) = \frac{N-n}{nN} \frac{\Sigma(X_h - \bar{X})^2 s_h^2}{\left[\Sigma(X_h - \bar{X})^2\right]^2}$$

where

$$s_h^2 = \frac{1}{N-1} \sum_j (Y_{hj} - \bar{Y}_h)^2.$$

With the homogeneity condition that $s_h^2 = s^2$, $h = 1, 2, \ldots, L$,

$$V(a) = \frac{N-n}{nN} s^2 \left\{ \frac{1}{L} + \frac{\bar{X}^2}{\Sigma(X_h - \bar{X})^2} \right\}$$

and

$$V(b) = \frac{N-n}{nN} \frac{s^2}{\Sigma(X_h - \bar{X})^2}.$$

Of course, V(a) and V(b) are estimated unbiasedly by replacing s_h^2 or s^2 by

unbiased estimates s_h^2 or s^2

where
$$s_h^2 = \frac{1}{n-1} \sum_{j=1}^{n} (y_{hj} - \bar{y}_h)^2$$

and
$$s^2 = \frac{1}{L(n-1)} \sum_h \sum_j (y_{hj} - \bar{y}_h)^2.$$

Finally, the variances of the estimated variances are

$$V\{\hat{v}(a)\} = \frac{(N-n)^2}{n^2 N^3 (n-1)(N-1)} \sum_h \left\{ \frac{1}{L} - \frac{(X_h - \bar{X})\bar{X}}{\sum(X_h - \bar{X})^2} \right\}^2 Q_h$$

and
$$V\{\hat{v}(b)\} = \frac{(N-n)^2}{n^2 N^3 (n-1)(N-1)} \frac{\sum(X_h - \bar{X})^4 Q_h}{\{\sum(X_h - \bar{X})^2\}^4}$$

where

$$Q_h = \sum_j Y_{hj}^4 (N-1)(n-1)(N-n) - 4 \sum_{i \neq j} Y_{hi}^3 Y_{hj} (n-1)(N-n)$$

$$+ \sum_{i \neq j} Y_{hi}^2 Y_{hj}^2 \left[\frac{1}{N-1}\right] \left[-3n(n-1) - N^2(n-3) + N(n^2 - 3)\right]$$

$$+ \sum_{i \neq j \neq k} Y_{hi}^2 Y_{hj} Y_{hk} \left[\frac{4}{(N-1)(N-2)}\right] \left[3n(n-1) + N^2(2n-3) - N(2n^2-3)\right]$$

$$+ \sum_{i \neq j \neq k \neq \ell} Y_{hi} Y_{hj} Y_{hk} Y_{h\ell} \left[\frac{-2}{(N-1)(N-2)(N-3)}\right] \times$$

$$\left[3n(n-1) + N^2(2n-3) - N(2n^2 - 3)\right] .$$

BIBLIOGRAPHY

Jönrup, H. and Rennermalm, B. (1976). Regression analysis in samples from finite populations. _Scand. J. Statist._ **3**: 33-36.

Order of moments is: 8
 Variable order: v1:8   v2:0 $K_{8,0} =$

(U80) + (-35 U40^2) + (-630 U20^4) + (-5040 U10^8) + (-8 U70 U10) +
(-28 U60 U20) + (56 U60 U10^2) + (-56 U50 U30) + (-336 U50 U10^3) + (
420 U40 U20^2) + (1680 U40 U10^4) + (560 U30^2 U20) + (-1680 U30^2 U10^2
) + (-6720 U30 U10^5) + (10080 U20^3 U10^2) + (-25200 U20^2 U10^4) + (
20160 U20 U10^6) + (336 U50 U20 U10) + (560 U40 U30 U10) + (-2520 U40 U20
 U10^2) + (-5040 U30 U20^2 U10) + (13440 U30 U20 U10^3)

  Variable order: v1:7   v2:1 $K_{7,1} =$

(U71) + (-1 U70 U01) + (-7 U61 U10) + (-7 U60 U11) + (-21 U51 U20)
+ (42 U51 U10^2) + (-21 U50 U21) + (-35 U41 U30) + (-210 U41 U10^3) +
(-35 U40 U31) + (210 U31 U20^2) + (840 U31 U10^4) + (140 U30^2 U11) +
(-2520 U21 U10^5) + (-630 U20^3 U11) + (5040 U11 U10^6) + (-5040 U10^7
U01) + (14 U60 U10 U01) + (42 U50 U20 U01) + (84 U50 U11 U10) + (-126
U50 U10^2 U01) + (210 U41 U20 U10) + (70 U40 U30 U01) + (210 U40 U21 U10
) + (210 U40 U20 U11) + (-630 U40 U11 U10^2) + (840 U40 U10^3 U01) + (
280 U31 U30 U10) + (-1260 U31 U20 U10^2) + (-420 U30^2 U10 U01) + (420
U30 U21 U20) + (-1260 U30 U21 U10^2) + (-630 U30 U20^2 U01) + (3360 U30
U11 U10^3) + (-4200 U30 U10^4 U01) + (-1890 U21 U20^2 U10) + (5040 U21
U20 U10^3) + (2520 U20^3 U10 U01) + (7560 U20^2 U11 U10^2) + (-12600 U20^
2 U10^3 U01) + (-12600 U20 U11 U10^4) + (15120 U20 U10^5 U01) + (-630 U40
 U20 U10 U01) + (-2520 U30 U20 U11 U10) + (5040 U30 U20 U10^2 U01)

  Variable order: v1:6   v2:2 $K_{6,2}$

(U62) + (-20 U31^2) + (-2 U61 U01) + (-1 U60 U02) + (2 U60 U01^2) +
(-6 U52 U10) + (-12 U51 U11) + (-6 U50 U12) + (-15 U42 U20) + (30 U42
U10^2) + (-30 U41 U21) + (-15 U40 U22) + (60 U40 U11^2) + (-20 U32 U30
) + (-120 U32 U10^3) + (20 U30^2 U02) + (-60 U30^2 U01^2) + (90 U22 U20^
2) + (360 U22 U10^4) + (180 U21^2 U20) + (-540 U21^2 U10^2) + (-90 U20^
3 U02) + (360 U20^3 U01^2) + (-540 U20^2 U11^2) + (-720 U12 U10^5) + (
-3600 U11^2 U10^4) + (720 U10^6 U02) + (-5040 U10^6 U01^2) + (24 U51 U10
U01) + (24 U50 U11 U01) + (12 U50 U10 U02) + (-36 U50 U10 U01^2) + (60
U41 U20 U01) + (120 U41 U11 U10) + (-180 U41 U10^2 U01) + (60 U40 U21 U01
) + (30 U40 U20 U02) + (-90 U40 U20 U01^2) + (60 U40 U12 U10) + (-90
U40 U10^2 U02) + (360 U40 U10^2 U01^2) + (120 U32 U20 U10) + (80 U31 U30
U01) + (240 U31 U21 U10) + (240 U31 U20 U11) + (-720 U31 U11 U10^2) + (
960 U31 U10^3 U01) + (120 U30 U22 U10) + (240 U30 U21 U11) + (120 U30 U20
U12) + (-360 U30 U12 U10^2) + (-720 U30 U11^2 U10) + (480 U30 U10^3 U02
) + (-2400 U30 U10^3 U01^2) + (-540 U22 U20 U10^2) + (-540 U21 U20^2 U01
) + (2880 U21 U11 U10^3) + (-3600 U21 U10^4 U01) + (-540 U20^2 U12 U10)
+ (1080 U20^2 U10^2 U02) + (-5400 U20^2 U10^2 U01^2) + (1440 U20 U12 U10^3
) + (4320 U20 U11^2 U10^2) + (-1800 U20 U10^4 U02) + (10800 U20 U10^4 U01
^2) + (8640 U11 U10^5 U01) + (-360 U40 U11 U10 U01) + (-720 U31 U20 U10
U01) + (-720 U30 U21 U10 U01) + (-720 U30 U20 U11 U01) + (-360 U30 U20
U10 U02) + (1440 U30 U20 U10 U01^2) + (2880 U30 U11 U10^2 U01) + (-2160
U21 U20 U11 U10 ` + (4320 U21 U20 U10^2 U01) + (4320 U20^2 U11 U10 U01) +
(-14400 U20 U11 U10^3 U01)

Variable order: v1:5   v2:3 $K_{5,3} =$

(U53) + (-3 U52 U01) + (-3 U51 U02) + (6 U51 U01^2) + (-1 U50 U03)
+ (-6 U50 U01^3) + (-5 U43 U10) + (-15 U42 U11) + (-15 U41 U12) + (-5
U40 U13) + (-10 U33 U20) + (20 U33 U10^2) + (-30 U32 U21) + (-30 U31
U22) + (120 U31 U11^2) + (-10 U30 U23) + (-60 U23 U10^3) + (180 U21^2
U11) + (30 U20^2 U13) + (-360 U20 U11^3) + (120 U13 U10^4) + (1440 U11^
3 U10^2) + (-120 U10^5 U03) + (-5040 U10^5 U01^3) + (6 U50 U02 U01) + (
30 U42 U10 U01) + (60 U41 U11 U01) + (30 U41 U10 U02) + (-90 U41 U10 U01^
2) + (30 U40 U12 U01) + (30 U40 U11 U02) + (-90 U40 U11 U01^2) + (10
U40 U10 U03) + (120 U40 U10 U01^3) + (60 U32 U20 U01) + (120 U32 U11 U10
) + (-180 U32 U10^2 U01) + (120 U31 U21 U01) + (60 U31 U20 U02) + (-180
U31 U20 U01^2) + (120 U31 U12 U10) + (-180 U31 U10^2 U02) + (720 U31 U10^
2 U01^2) + (60 U30 U22 U01) + (60 U30 U21 U02) + (-180 U30 U21 U01^2) +
(20 U30 U20 U03) + (240 U30 U20 U01^3) + (40 U30 U13 U10) + (120 U30 U12
U11) + (-360 U30 U11^2 U01) + (-60 U30 U10^2 U03) + (-1200 U30 U10^2 U01
^3) + (60 U23 U20 U10) + (180 U22 U21 U10) + (180 U22 U20 U11) + (-540
U22 U11 U10^2) + (720 U22 U10^3 U01) + (-540 U21^2 U10 U01) + (180 U21
U20 U12) + (-540 U21 U12 U10^2) + (-1080 U21 U11^2 U10) + (720 U21 U10^3
U02) + (-3600 U21 U10^3 U01^2) + (-270 U20^2 U12 U01) + (-270 U20^2 U11
U02) + (1080 U20^2 U11 U01^2) + (-90 U20^2 U10 U03) + (-1800 U20^2 U10
U01^3) + (-180 U20 U13 U10^2) + (240 U20 U10^3 U03) + (7200 U20 U10^3 U01
^3) + (1440 U12 U11 U10^3) + (-1800 U12 U10^4 U01) + (-7200 U11^2 U10^3
U01) + (-1800 U11 U10^4 U02) + (10800 U11 U10^4 U01^2) + (2160 U10^5 U02
U01) + (-90 U40 U10 U02 U01) + (-720 U31 U11 U10 U01) + (-180 U30 U20 U02
U01) + (-360 U30 U12 U10 U01) + (-360 U30 U11 U10 U02) + (1440 U30 U11
U10 U01^2) + (720 U30 U10^2 U02 U01) + (-540 U22 U20 U10 U01) + (-1080
U21 U20 U11 U01) + (-540 U21 U20 U10 U02) + (2160 U21 U20 U10 U01^2) + (
4320 U21 U11 U10^2 U01) + (1080 U20^2 U10 U02 U01) + (-1080 U20 U12 U11 U10
) + (2160 U20 U12 U10^2 U01) + (4320 U20 U11^2 U10 U01) + (2160 U20 U11
U10^2 U02) + (-10800 U20 U11 U10^2 U01^2) + (-3600 U20 U10^3 U02 U01)

Variable order: v1:4   v2:4 $K_{4,4} =$

(U44) + (-18 U22^2) + (-144 U11^4) + (-4 U43 U01) + (-6 U42 U02) +
(12 U42 U01^2) + (-4 U41 U03) + (-24 U41 U01^3) + (-1 U40 U04) + (6
U40 U02^2) + (24 U40 U01^4) + (-4 U34 U10) + (-16 U33 U11) + (-24 U32
U12) + (-16 U31 U13) + (-4 U30 U14) + (-6 U24 U20) + (12 U24 U10^2) +
(-24 U23 U21) + (144 U22 U11^2) + (72 U21^2 U02) + (-216 U21^2 U01^2)
+ (6 U20^2 U04) + (-54 U20^2 U02^2) + (-360 U20^2 U01^4) + (72 U20 U12^2
) + (-24 U14 U10^3) + (-216 U12^2 U10^2) + (24 U10^4 U04) + (-360 U10^4
U02^2) + (-5040 U10^4 U01^4) + (24 U41 U02 U01) + (8 U40 U03 U01) + (
-36 U40 U02 U01^2) + (32 U33 U10 U01) + (96 U32 U11 U01) + (48 U32 U10
U02) + (-144 U32 U10 U01^2) + (96 U31 U12 U01) + (96 U31 U11 U02) + (
-288 U31 U11 U01^2) + (32 U31 U10 U03) + (384 U31 U10 U01^3) + (32 U30
U13 U01) + (48 U30 U12 U02) + (-144 U30 U12 U01^2) + (32 U30 U11 U03) +
(384 U30 U11 U01^3) + (8 U30 U10 U04) + (-72 U30 U10 U02^2) + (-480 U30
U10 U01^4) + (48 U23 U20 U01) + (96 U23 U11 U10) + (-144 U23 U10^2 U01)
+ (144 U22 U21 U01) + (72 U22 U20 U02) + (-216 U22 U20 U01^2) + (144 U22
U12 U10) + (-216 U22 U10^2 U02) + (864 U22 U10^2 U01^2) + (48 U21 U20
U03) + (576 U21 U20 U01^3) + (96 U21 U13 U10) + (288 U21 U12 U11) + (
-864 U21 U11^2 U01) + (-144 U21 U10^2 U03) + (-2880 U21 U10^2 U01^3) + (
-72 U20^2 U03 U01) + (432 U20^2 U02 U01^2) + (24 U20 U14 U10) + (96 U20
U13 U11) + (-432 U20 U11^2 U02) + (1728 U20 U11^2 U01^2) + (-36 U20 U10^2
U04) + (432 U20 U10^2 U02^2) + (4320 U20 U10^2 U01^4) + (-288 U13 U11
U10^2) + (384 U13 U10^3 U01) + (-864 U12 U11^2 U10) + (576 U12 U10^3 U02
) + (-2880 U12 U10^3 U01^2) + (2304 U11^3 U10 U01) + (1728 U11^2 U10^2 U02
) + (-8640 U11^2 U10^2 U01^2) + (384 U11 U10^3 U03) + (11520 U11 U10^3
U01^3) + (-480 U10^4 U03 U01) + (4320 U10^4 U02 U01^2) + (-288 U31 U10
U02 U01) + (-288 U30 U11 U02 U01) + (-96 U30 U10 U03 U01) + (576 U30 U10
U02 U01^2) + (-864 U22 U11 U10 U01) + (-432 U21 U20 U02 U01) + (-864 U21
U12 U10 U01) + (-864 U21 U11 U10 U02) + (3456 U21 U11 U10 U01^2) + (1728
U21 U10^2 U02 U01) + (-288 U20 U13 U10 U01) + (-864 U20 U12 U11 U01) + (
-432 U20 U12 U10 U02) + (1728 U20 U12 U10 U01^2) + (-288 U20 U11 U10 U03)
+ (-5760 U20 U11 U10 U01^3) + (576 U20 U10^2 U03 U01) + (-4320 U20 U10^2
U02 U01^2) + (3456 U12 U11 U10^2 U01) + (-5760 U11 U10^3 U02 U01) + (3456
U20 U11 U10 U02 U01)

ERRATA

INDEX OF TERMS
The page number of the definition is underscored.

VITA

Robert S. Wilson

Candidate for the Degree of

Doctor of Science

Dissertation: THE RANDOMIZATION MODEL FOR THE ANALYSIS
OF COVARIANCE

Major Field: Statistics

Biographical:
    Personal Data: Born in Lake Charles, Louisiana,
        September 15, 1948, son of Donald Edward
        Wilson, Sr. and Nell Courtney Wilson.

    Education: Graduated from East Jefferson Public
        High School in June, 1966; received Bachelor of
        Arts degree in Sociology from Grinnell College,
        Grinnell, Iowa, in May, 1970, with the junior
        year on leave to the University of Singapore,
        Republic of Singapore. Received Master's of
        Science Degree in Statistics from Iowa State
        University, Ames, Iowa in May, 1980; received
        Doctor of Philosophy degree in Statistics from
        Oklahoma State University in May of 1993.

    Professional Data Analysis: For Bama Pie, Ltd. of
        Tulsa Oklahoma, Statistician, process improve-
        ment and taste analysis, from January 1987 to
        November 1989; for the Food and Nutrition
        Department, Iowa State University, Ames, Iowa,
        body fat measurements, January to May 1980; for
        the National Fire Prevention and Control Admin-
        istration, U. S. Dept. of Commerce, Washington,
        D.C., survey of national fire losses, the summer
        of 1978; for the Development Research Center of
        the World Bank, Washington, D. C., household
        consumption data for a national input-output
        matrix, January 1976 to May 1976, January and
        the summer of 1977; for the Brookings Institu-
        tion, Washington, D.C., analysis of the federal
        revenue sharing program, from January to
        November 1975; for the Bureau of Social Science
        Research, Washington, D. C., survey data anal-
        ysis, cleaning, and coding, from January 1973
        to December 1974; for the Manpower Administra-
        tion of the U. S. Dept. of Labor, Washington,
        D. C., program participation data analysis,
        from July 1971 to June 1972.

Professional Instruction: Graduate level Mathemati-
     cal Statistics and undergraduate level Business
     Calculus for the University of Central Oklahoma,
     Edmond, Oklahoma, from January to May 1992;
     Lotus 1-2-3 computer programming for the Indian
     Meridian Area Vocational-Technical School, from
     January to April 1991, Stillwater, Oklahoma;
     introductory and advanced undergraduate statis-
     tical courses for the University of Oklahoma of
     Stillwater, Oklahoma, from August 1981 to May
     1985; graduate level linear programming, statis-
     tics, and computer programming for the Facultad
     del la Universidad Autonoma de Chihuahua,
     Ciencias Agricolas, in Delicias, Chihuahua,
     Republic of Mexico, in the Summer of 1982
     (medium of instruction: Spanish). Teaching
     Assistant for Oklahoma State University in
     Statistics in the summer of 1985; for Iowa
     State University in the summer of 1980.

Professional Consulting: Moore Business Forms,
     doming and creasing in pads of paper, December
     1992. At Oklahoma State University: Marilyn
     Waters, Home Economics Department, mail survey
     of restaurant storage of pecans, from November
     to January 1992; Tamara Walkingstick, Forestry
     Dept., mail survey of smallholders of forested
     lands, from August 1991 to January 1992; Hamid
     Pourjalali, Economics Dept., economic modeling
     of large corporations, September 1991; H.
     Matthews, Agricultural Economics Dept., input-
     output matrix analyzing rural development of a
     U.S. Corps of Engineering dam, February to May
     1991; Dr. Sue Hong, Dept. of Home Economics,
     school leaving of teenaged mothers (SIPP data
     base), April 1991; Ellen Dickson, Dept. of
     Education, mail survey of school management
     style and effectiveness, June to August 1990;
     Sylvia Gaiko, Hotel and Restaurant Management,
     taste tests of Oklahoma wheat pasta, July 1990;
     Myoung Lee, student satisfaction with life at
     OSU, March 1987; Jeffery D. Fair, dimethyl
     sulfoxide as a topical therapy, December, 1986;
     Roy Stubbs, test of teaching style and student
     aptitude, February, 1986; Michael E. Akpan,
     survey of Nigerian teachers and students,
     January 1986; Ken Trosper, tank outlet pressure
     and flow rate, January 1986; Randy Currie,
     Agronomy Department, herbicide seed kill and
     seed life in soil, December 1985.

Professional Organizations: American Statistical
     Association, presentations at the 1988, 1989,
     and 1990 annual meetings.