Mic 60-5188

HAGANS, M.D., James Albert. SEQUENTIAL ANALYSIS IN MEDICAL RESEARCH.

The University of Oklahoma, Ph.D., 1960
Health Sciences, general

University Microfilms, Inc., Ann Arbor, Michigan

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

SEQUENTIAL ANALYSIS IN MEDICAL RESEARCH

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

JAMES ALBERT HAGANS, M.D.

Oklahoma City, Oklahoma

1960

SEQUENTIAL ANALYSIS IN MEDICAL RESEARCH

APPROVED BY

Earl R. Doering

Kent J. Worley

John C. Briden

W. A. Schottstaedt

DISSERTATION COMMITTEE

ACKNOWLEDGMENT

I am indebted to Doctor Carl R. Doering, my senior advisor, for his guidance and untiring help in the formulation of the ideas presented in this dissertation, and to Doctor Kirk T. Mosley, Doctor William W. Schottstaedt, Doctor Joseph B. Goldsmith, and Doctor Stewart G. Wolf for their encouragement and faith in the necessity to solve some of the problems of the medical researcher as he is faced with the analysis and interpretation of his experimental data. Doctor John C. Brixey, Doctor Carl E. Marshall, and Doctor Robert D. Morrison were ever ready to lend a word of encouragement and repeatedly displayed unique broadness in their understanding of the person with a relatively weak mathematical and strong biological and medical research background, besides guiding me through the basic theory and methods of modern biological statistics. I am deeply grateful to all of these men, not only for their excellent guidance and formal instruction, but also for their encouragement and for the examples they so aptly set.

Edward N. Brandt, Jr., M. S., helped me considerably by providing formal tutoring in the principles of analytic geometry and the differential and integral calculus.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

SEQUENTIAL ANALYSIS IN MEDICAL RESEARCH

CHAPTER I

INTRODUCTION

Sequential analysis has been found useful as a statistical technique for the interpretation of various biological and medical experiments. Originally developed by Wald, and extended by the Columbia University Statistical Research group, it serves as an efficient statistical tool for quality control in industry (1,2). Oakland, Fisher, Bross and Armitage were among the earlier workers who apparently perceived its potential and actual value to the area of biological and medical science (3,4,5,7). Burington and May described the mathematical theory and development in as elementary and lucid a way as is apparently possible (6). During the next five years multiple reports describing its uses in biology and medicine have appeared (8-23,29).

The present dissertation concerns itself primarily with the adaptation of sequential techniques of statistical analysis to the area of medical research. The field of medical experimentation has certain problems and peculiarities which, if not entirely unique to medicine, at least are sufficiently deviant from the usual problems of quality control in industry to be considered worthy of separate consideration. As a generalization, the major problems may be stated to be in the realm of

incomplete data, nonnormally distributed data and marked limitations as to the number of experimental subjects available.

Incomplete data faces the medical researcher and manifests itself in several ways. First, there is often no large backlog of readily available data, collected on subjects and in a situation comparable to that of the planned experimentation, which can be used as a "standard" or "control." In other words, at the onset of the experiment the researcher often does not have a satisfactory estimate of μ_x (the true mean) or $\sigma_{\bar{x}}$ (the true standard deviation) for the standard or control population. He may, of course, make an educated guess as to these quantities. Those familiar with the field will recognize only too well the limitations of this method. He may also accept data from another laboratory in another section of the country or world for his standard or control purposes. This, too, leaves much to be desired. He often has no good notion as to the nature of the distribution of his data, or he may know with fair certainty that they are not from a normal distribution. Further, he is often at a loss to say exactly what magnitude of change from the standard or control μ_x would be biologically or medically desirable. In brief, he can answer only partially, if at all, the major questions such as: Are the data from a normal distribution? What are the values of μ_x and $\sigma_{\bar{x}}$ of the standard or control group? In exactly what magnitude of change from the μ_x of the standard or control group are you interested? Yet, in order to employ the more commonly described methods of sequential analysis these pertinent questions must be answered. This alone has served to cause many medical researchers to avoid the use of currently available models of sequential analysis. Yet the techniques of sequential analysis have much

to offer the scientific experimenter in medical research. Particularly is this true with regard to reducing the average number of observations needed to reach a given decision at a given level of confidence.

Were it possible for the medical biostatistician to say to the medical research worker: "There is available an experimental plan which does not require that the samples come from a normal distribution, does not require that one give an estimate of μ_x and $\sigma_{\bar{x}}$ of the standard or control group before one begins the experimentation, and does not require that one set in advance the magnitude of the significant difference for which one is searching, and yet will permit one to draw inferences with reasonably small numbers of observations and with approximately correct probabilities at whatever level of confidence less than .05 one wishes" -- then the medical scientist should indeed be interested in such a technique. Basically, it is the purpose of this dissertation to describe the development of such a technique and illustrate its usefulness by various empirical sampling methods applied to various known distributions. There is a certain gratifying confidence gained by applying a technique to situations where the answers are known in advance and in repeatedly obtaining the correct answer by that technique in these various situations. Considerable empirical evidence is then at hand to support the contention of the usefulness of the technique in the solution of experimental problems.

Several points perhaps need clarification before proceeding further. It is not to be construed from the above that the medical researcher will play no part in the determination of the size of the difference from the standard or control μ_x in which he is interested.

He need not, however, make this decision in advance of the experimentation. The decision is to remain a joint one between the medical biostatistician and the medical researcher and is to be reached during the course of the experimentation, making use of the knowledge gained concerning μ_x in the standard or control group and $\sigma_{\bar{x}}$ in the standard or control and the treatment groups. The desirability of collecting data on the standard or control group simultaneously, at random and under the same experimental conditions as the data on the treatment group are collected, is obvious.

In order to prepare the reader for the subject matter of this dissertation, the following outline is presented:

Chapter II. An example of the application of a classical model of sequential analysis in medical research: measured data, normally distributed, one-sided alternative. Venous blood glucose data are used.

Chapter III. Consideration is given to the problem of data not normally distributed. Consideration is given to a method for obtaining evidence that means of size n from such nonnormal distributions will be approximately normally distributed. Consideration is given also to the comparison of means of appropriate size n from the treatment group to those from the standard or control group by the method of sequential analysis.

Chapter IV. Consideration is given to the two-sided alternative in medical research. Here, the need exists for knowing not only if the observations in the experimental group exceed the limits set about the control μ_x as in Wald's classical two-sided alternative model, but also for knowing if the experimental group mean is higher, if it is lower, or

if it is approximately the same as the standard or control mean.

1. Armitage was apparently the first to recognize this need in medical research.

2. An example is taken from the more recent literature.

3. Empirical evidence is presented that the method described in 2 above may lead one to frequent improbable conclusions.

4. Empirical evidence is presented and suggests that the model of Armitage is a more desirable method for solution of such two-sided alternative problems.

Chapter V. An approximation technique of sequential analysis is described incorporating the ideas in Chapters III and IV above. In essence: the two-sided alternative in medical (biological) research where one has not an extensive past experience with data collection in the current area of planned research, does not know whether or not the data are from a normal distribution, and would accept several different values of the difference (δ) from the standard or control mean as biologically significant. In short, the experimenter and biostatistician must work together and estimate μ_x and $\sigma_{\bar{x}}$ of the standard or control group and decide upon a satisfactory difference for which to test (taking into account the knowledge gained of the variation present in their collected data as the experiment progresses).

1. Empirical evidence is presented that such a model of analysis should be useful, using samples from known distributions (normal, rectangular, logarithmic).

2. Further empirical evidence is presented that such a model of analysis should be useful, using actual medical research data previously

collected and already analyzed by more standard statistical techniques using larger numbers of observations than those required by the described model of sequential analysis.

Chapter VI. Summary and Conclusions.

Bibliography.

Appendix. Tables and figures.

Brief Description of Material and Methods

Three types of distribution have been selected for empirical sampling: (1) an approximately normal distribution of size 100 established and used by Snedecor (27); (2) an approximately rectangular distribution of size 100 obtained by using the fourth place numbers from a five place table of Common Logarithms from Arkin and Colton (28); (3) an approximately logarithmic distribution of size 100 obtained by recording the antilogarithms of the numbers in the approximately normal distribution from Snedecor in (1) above, transformed first by dividing each number in Snedecor's distribution by 100. Table 1, Table 2 and Table 3, respectively, present these three distributions.

Data from five separate completed medical experiments previously analyzed by more standard statistical procedures are also used for empirical illustrations.

Table 4 presents the results of the 24-hour excretion of urine Na (in mEq. per liter) from subjects randomly given either a placebo or two injections of Mercuhydrin (1.0 cc. at 8:00 a.m. and 1.0 cc. at 4:00 p.m.) (30).

Table 5 presents the results of the 24-hour excretion of urine

K (in mEq. per liter) from subjects randomly given either a placebo or a dose of 250 mg. of Diamax four times daily by mouth (30).

Table 6 presents the change in gastric free HCl observed in subjects after they were randomly given a dose per Levine tube of a solution containing a placebo or 2.5 mg. of Reserpine (16).

Table 7 presents the results of a serum cholesterol experiment reported by Drolette and Hegsted (23).

Tables 8 and 9 present the results of the blood glucose experiments given to the author by personal communication from Doctor Kelly M. West (24).

CHAPTER II

AN EXAMPLE OF THE APPLICATION OF A CLASSICAL MODEL OF SEQUENTIAL ANALYSIS IN MEDICAL RESEARCH

An example of the use of a classical model of sequential analysis for the interpretation of a medical experiment follows. Wald's standard notation is used: m_0 and $\sigma_{\bar{x}}^2$ (the control group mean and variance), m_1 (the mean desired in the treatment group to be considered significantly different from the control group mean), α and β (the probability risks of wrong decisions), a and b, as well as h_0 and h_1 (the intercepts), S (the slope) and $\bar{n}_{m_0}$ and $\bar{n}_{m_1}$ (the average sample numbers) (1).

From previous experiments on several hundred healthy young male subjects between the ages of 20 and 30 years studied in their laboratory, the investigators knew that the average change in blood sugar observed three hours after administration of a placebo and following a 12-hour fast was -3.4 mg. per cent $\pm$ 7.41 mg. per cent (modified Symogi method). They wished to evaluate the hypoglycemic effects of a new oral agent (chlorpropamide) in healthy young male subjects, and they wanted to answer the question: Does the agent cause a drop of -15 mg. per cent or more in the blood glucose level three hours after oral administration of 1.0 gram of chlorpropamide, or is it accompanied by a change of -3.4 mg. per cent or less? It was decided that the risk for making an

erroneous decision in either direction should be 1 in 100. With this much information, and before starting the experiment, one may construct the sequential graph upon which the analysis can be performed as follows:

m_0 = mean of control group = -3.4

σ_x^2 = variance of control group = 54.93; σ_x = 7.41

m_1 = arbitrarily selected mean considered by the experimenters to be a satisfactory indication of a desirable effect from the drug = -15 mg. per cent.

α = probability of being wrong about m_0 (here α = .01).

β = probability of being wrong about m_1 (here β = .01).

$a = \log_e (1-\beta)/\alpha = 4.595$

$b = \log_e (1-\alpha)/\beta = 4.595$

Upper Intercept = $h_1 = (a\sigma^2)/(m_1-m_0) = (4.595 \times 54.93)/[-15 - (-3.4)]$

$= 252.40/-11.6 = -21.76$

Lower Intercept = $h_0 = (-b\sigma^2)/(m_1-m_0) = (-4.595 \times 54.93)/[-15 - (-3.4)]$

$= +21.76$

(Note: here $|h_1| = |h_0| = |21.76|$)

Slope = $S = [-3.4 + (-15)]/2 = -18.4/2 = -9.2$

$\bar{n}_{m_1} = [(1-\beta)a - \beta b] / [(m_1-m_0)^2/2\sigma^2]$ = average number of observations required to accept the mean of the experimental group (m_e) $\leq m_1$ if this is true.

$\bar{n}_{m_0} = [(1-\alpha)b - \alpha a] / [(m_1-m_0)^2/2\sigma^2]$ = average number of observations required to accept the mean of the experimental group (m_e) $\geq m_0$ if this is true.

(Note: here $\bar{n}_{m_1} = \bar{n}_{m_0}$ = approximately 4.)

The equation for the upper line of the sequential graph is: $h_1 + Sn$, and for the lower line is: $h_0 + Sn$. By substituting in various convenient values of n (i.e., as here n = 0, n = 5, n = 10) one obtains three points for the upper line and three points for the lower line as follows:

n =	Lower Line	Upper Line
0	21.76	-21.76
5	-24.24	-67.76
10	-70.24	-113.76

An ordinary sheet of graph paper is used and on the abscissa one indicates n (the number of observations) and on the ordinate one indicates the sum of X (here X = blood glucose three hours after chlorpropamide minus control blood glucose in mg. per cent for each subject tested). The results are shown on the graph presented in Figure 1.

The results of the first 10 chlorpropamide tests performed at random on healthy male subjects between 20 and 30 years of age are presented in Table 9.

The average change in blood sugar was -17.7 mg. per cent for the 10 subjects. The analysis on the sequential graph is made by plotting the figures in Column V of Table 9 which are -22, then -44, then -46, then -69, etc. until a decision is reached (i.e., until the upper or lower line is crossed). In this experiment, the result of the second observation led to crossing of the lower line and the experiment could have been terminated at that point and the decision reached that the mean of the experimental group (m_e) would be equivalent to -15 mg. per cent or more (i.e., accept $m_e \leq m_1$). Since more observations were already available in this

experiment, the analysis was continued by beginning again at zero and taking the third subject as if he were the first subject and plotting the sum of X from that point on until a decision was reached again (i.e., plot the figures in Column VI of Table 9 which are -2, -25, -51, -68, etc.). This is represented by the black triangle plots in Figure 1. This second trial led to the same decision on the third observation, namely, to accept that the mean of the experimental group is equivalent to -15 mg. per cent or more. For interest, one could run a third trial since there seem to be sufficient observations left unused. This time one would begin with the sixth observation as though it were the first, and beginning again at zero on the graph, plot the figures in Column VII which are -17, -37, -46, -59, and -82 until a decision is reached. This is represented by the star plots in Figure 1, and here it took five observations to reach the same decision. It is to be emphasized that the second and third trials were unnecessary and were performed here simply for interest because the observations were already available and to illustrate the consistency of the decision made.

The experimenter could have stopped his experiment here after the second test with chlorpropamide, or, if he were dissatisfied with this small group, he could have begun the analysis again and stopped the experiment after the third test (a total of five tests in all). If he were still dissatisfied, he could have continued the experiment, as he did in this case, and begun the analysis a third time and stopped after the fifth test (a total of 10 tests in all, now with three decisions all the same).

CHAPTER III

A METHOD FOR OBTAINING EVIDENCE THAT MEANS OF SAMPLES OF SIZE n WILL BE APPROXIMATELY NORMALLY DISTRIBUTED

Given the experimental situation where one wishes to compare a mean of a treatment or experimental group to a mean of a standard or control group, but has not extensive past data under similar experimental conditions available: it is suggested that by random procedure one obtain nine control and nine treatment observations. This will permit the obtaining of four means of size n = 2 from the control and four from the experimental group, three from each group of size n = 3 and two from each group of size n = 4. From pairs of these samples of size n (paired within either the treatment or the control group) one may proceed to estimate $\sigma_{\bar{x}}$ (the standard error of the mean of samples of size n) by three methods:

(1) Estimate of $\sigma_{\bar{x}}$ = Range/Constant (see reference 25).

(2) Estimate of $\sigma_{\bar{x}} = \left[\sum(\bar{x}_i - \bar{\bar{x}})^2\right]/(n-1)$

(3) Estimate of $\sigma_{\bar{x}} = \sqrt{S_{x_p}{}^2/n}$

One may then compare these three estimates of $\sigma_{\bar{x}}$ from a pair of the above samples to see if they reasonably agree or not (the pairing should not be between the control and the treatment groups, as here the two samples may not come from a population with the same μ_x, although presumably they are from a population with a common $\sigma_{\bar{x}}$). Ideally, one

should test this presumable homogeneity of variances by the appropriate F ratio. This has been done for all the data presented later and is summarized in Table 10. One should be willing to accept homogeneity unless p = .01 or less.

A detailed example of the above outlined procedure follows to illustrate step-by-step the methodology employed: one obtains from Table 1 (the Normal Distribution) two random samples (using a table of random numbers) with nine observations in each sample. The first sample will be called the control and the second sample the treatment group. The ultimate purpose is to determine if the two samples came from a population with a common μ_x or not. The immediate objective is to determine what size sample (i.e., n = ?) will yield means approximately normally distributed.

The results of the sampling were as follows:

Control Group	Treatment Group
36	27
33	36
34	14
32	42
27	17
36	23
25	26
27	11
11	11

Beginning with samples of size n = 2, one has immediately two sets of pairs from each group as follows:

From the Control Group:

Pair 1:	36	34
	33	32
Pair 2:	27	25
	36	27

From the Treatment Group:

Pair 1:	27	14
	36	42
Pair 2:	17	26
	23	11

Beginning with the first pair of the control group, one proceeds as follows:

36	34
33	32
69	66

(1) $\bar{X}_1 = 34.5$ $\bar{X}_2 = 33.00$

(2) Range: $\bar{X}_1 - \bar{X}_2 = 34.5 - 33 = 1.5$

(3) $CSS_1 = 2385.0 - 2380.5 = 4.5$ $CSS_2 = 2180.0 - 2178.0 = 2.0$

(4) $S^2_{x_{pooled}} = (4.5+2)/(1+1) = 6.5/2 = 3.25$

(5) $S^2_{\bar{x}} = \sum(\bar{x}_i - \bar{\bar{x}})^2/(n-1) = \left[\sum \bar{x}_i^2 - (\sum \bar{x}_i)^2/n\right]/(n-1) = (2279.25 - 2278.125)/1$

$= 1.125$

With (2), (4) and (5) above one may readily obtain three estimates of $\sigma_{\bar{x}}$ for means of samples of size n = 2. These estimates are:

(1) $S_{\bar{x}} = R/C = 1.5/1.13 = 1.327$ (see reference 25)

(2) $S_{\bar{x}} = \sqrt{S^2_{\bar{x}}} = \sqrt{1.125} = 1.060$

(3) $S_{\bar{x}} = \sqrt{S^2_{x_p}/n} = \sqrt{3.25/2} = \sqrt{1.125} = 1.060$

An approximate 95 per cent confidence interval is obtained on $S_{\bar{x}}$ as follows: SE of $S_{\bar{x}} = S_{\bar{x}}/\sqrt{2n} = 1.060/\sqrt{4}$ (under the assumption of normality).

Here, the approximate 95 per cent CI = $S_{\bar{x}} \pm 2\ S_{\bar{x}}/\sqrt{2n} = 1.060 \pm 2(1.060)/\sqrt{4}$

$= 1.060 \pm 1.060 = 0$ to 2.120.

The other two estimates of $S_{\bar{x}}$ (1.060 and 1.327) fall well within this range.

The next pair of samples from the control group is then tested similarly, then the next pair (which here comes from the treatment group), etc. The results of testing of all four pairs are summarized in Table 11 and one accepts that means of size n = 2 should be approximately normally distributed. The same procedure was applied to the rectangular distribution and the logarithmic distribution, and the results are also summarized in Table 11. If the three $S_{\bar{x}}$ estimated from each set of pairs of samples of size n = 2 do not agree, one should proceed to testing samples size n = 3. This occurred in the logarithmic distribution (Table 11). Should they not agree in the samples of size n = 3, one should proceed to samples size n = 4, etc. until agreement is reached in at least two sets of pairs of samples. This occasionally occurred in the examples of medical research data presented later.

Although no such data occurred in this presentation, it is possible that sometimes by inspection of the results of sampling of nine control and nine experimental subjects there will be internal evidence suggesting outliers or badly skewed distributions. In some instances the experimenter may know from previous work of his own or of others that the population from which his samples came is probably severely skewed. In these special instances, it is probably desirable to seek some normalizing transformation of the data (as taking logarithms, square or cube roots, etc.) before proceeding to the above described method. Every effort should be made to

normalize the data in a specific instance by whatever means are reasonably and permissibly available. In general, this should permit one to work with the means of small sample sizes (i.e., $n \leq 5$).

CHAPTER IV

CONSIDERATION OF THE TWO-SIDED ALTERNATIVE IN MEDICAL RESEARCH

The two-sided alternative is frequently of extreme interest to the medical researcher. Wald's original description of it was designed to meet the needs of quality control, and thus led to accepting a product which did not vary more than a specified amount in either direction (up or down) from the control or standard product, or to rejecting it as varying more than the desired amount. The sign of the variation (+ or -) from the control mean is thus in essence ignored (1). This model may have some applications in experimental medicine, but much more commonly the experimenter wishes to answer the question: Is the mean of the experimental group essentially the same as the mean of the control group or not, and if not, is the mean of the experimental group higher or is it lower?

Armitage was apparently the first to recognize and attempt to solve this need in the area of medical research. He described a model for binomially distributed data, which permitted answering the question: Is the percentage of successes in the treatment group essentially the same as in the standard group, or is it significantly greater, or significantly less? The same idea can of course be extended to the situation

with regard to the measured variable. The problem of the comparison of multiple treatment groups to a single standard or control arises, but will not be considered in this dissertation. It is the subject of a separate report to be made at a later date.

The most intriguing idea that comes immediately to mind is: instead of m_0 = mean of control, m_1 = desired mean of treatment group, why not let m_0 = desired mean of treatment group to be regarded as significantly lower than mean of control, and m_1 = desired mean of treatment group to be regarded as significantly higher than mean of control? This, in effect, allows the mean of the control to follow a center path between the two parallel lines of the sequential graph. It also doubles the usual δ (i.e., m_1-m_0) and thus affects both the slope (i.e., $(m_1+m_0)/2$) and the intercepts $\left[\text{i.e., } a\sigma^2/(m_1-m_0)\right]$ as well as the average number (i.e., $\left[(1-\alpha)b-\alpha a\right]/(m_1-m_0)^2/2\sigma^2$). Drolette and Hegsted have recently reported an application of this technique, in which they carried out multiple comparisons of several treatment groups against a single control (stating no special allowance for this latter was necessary) (23). Happily, because the δ they selected was small in regard to $\sigma_{\bar{x}}$, they apparently avoided trouble. Empirical evidence will be presented below to illustrate that this type of analysis contains a highly probable error and can lead one into difficulty, particularly if the δ selected leads to an average number less than 1 (which it will often do if δ is 2σ or more). In their application their average number was six or seven, and it probably should have been 25 or 26 (explanation later in chapter). The model of Armitage would have in essence let m_0 remain the mean of the control group, set m_1 as the desired mean of the

treatment group to have been considered significantly higher than m_0, and m_2 as the desired mean of the treatment group to be considered significantly lower than m_0.

Results of sampling and testing will now be cited to lend empirical evidence to support that Drolette and Hegsted's method is probably in error and that the method of Armitage is probably more desirable. If one samples from Snedecor's normal distribution where it is known that $\mu_x = 30$ and $\sigma_{\bar{x}} = 10$, one should expect not very often to cross the lines of the graph leading to the conclusion that the sample did not come from a distribution with $\mu_x = 30$.

First, using the method of Drolette and Hegsted, one has: $\mu_x = 30$ and $\sigma_{\bar{x}} = 10$; m_0 is arbitrarily selected as 10 and m_1 as 50, thus a δ approximately $2\sigma_{\bar{x}}$ on either side of μ_x (a not unreasonable difference to be searching for in medical experimental work). Also, α is arbitrarily selected as .01 and β as .01. One then has the necessary basic information for construction of the sequential graph:

$h_0 = -b\sigma^2/(m_1-m_0) = (-4.595 \times 100)/(50-10) = -459.5/40 = -11.487 \cong -11.5$

$h_1 = a\sigma^2/(m_1-m_0) = +11.5$

$S = (m_1+m_0)/2 = (50+10)/2 = 60/2 = 30.$

Points for plotting the graph may now be determined as follows:

n =	Upper Line	Lower Line
0	11.5	-11.5
5	161.5	138.5
10	311.5	288.5
20	611.5	588.5
40	1211.5	1188.5

$$\bar{n}_{m_0} = \bar{n}_{m_1} = \left[(1-\alpha)b-\alpha a\right]/(m_1-m_0)^2/2\sigma^2$$

$$= \left[(1-.01)4.595 - .01(4.595)\right]/(50-10)^2/\left[2(100)\right]$$

$$= (4.549 - .0459)/1600/200 = 4.503/8 = 0.562 \quad (\text{i.e., } < 1!)$$

Random samples from Snedecor's data follow:

	n =	X	ΣX	Upper Line	Lower Line
First sample:	1	36	36	41.5	18.5
	2	33	69	71.5	48.5
	3	34	103	101.5	78.5
Second sample:	1	3	3	41.5	18.5
Third sample:	1	26	26	41.5	18.5
	2	17	43	71.5	48.5
Fourth sample:	1	49	49	41.5	18.5
Fifth sample:	1	27	27	41.5	18.5
	2	30	57	71.5	48.5
	3	44	101	101.5	78.5
	4	41	142	131.5	108.5

Surely there are too many erroneous results. In five samples of size n = 1 to n = 4, five conclusions have been reached that the mean of the population from which the samples came was not equivalent to 30 (i.e., reject $m_e = 30$). Three times the conclusion was reached to accept it equivalent to or greater than m_1 (i.e., ≥ 50) and twice to accept it equivalent to or less than m_0 (i.e., ≤ 10)! Yet it is known that all of these decisions are highly improbable. Evidently, there is a highly probable error in this method, and it is more subtle and requires larger and larger samples to demonstrate when α is selected so that the average number is greater than 1. Thus the reason for the illustrative example presented.

According to the method of Armitage, an apparently more appropriate analysis of the above problem would have been:

Let $m_0 = 30$

$\sigma_{\bar{x}} = 10$

$m_1 = 50$

$m_2 = 10$

$\alpha = .01$

$\beta = .01$

$h_0 = -b\sigma^2/(m_1-m_0) = (-4.595 \times 100)/(50-30) = -459.5/20 = -22.975 \cong -23$

$h_1 = a\sigma^2/(m_1-m_0) \cong 23$

$S_1 = (m_1+m_0)/2 = (50+30)/2 = 80/2 = 40.$

For the lower portion of the graph, using m_2, the intercepts remain unchanged but:

$S_2 = (m_2+m_0)/2 = (10+30)/2 = 40/2 = 20.$

[Also: $S_2 = S_1 - (m_1-m_2)/2 = 40 - 40/2 = 20$].

Points for plotting the graph may now be determined as follows:

	m_1 and m_0		m_0 and m_2	
n =	Upper Line	Lower Line	Upper Line	Lower Line
0	23	-23	23	-23
5	223	177	123	77
10	423	377	223	177
etc.				

Here, $\bar{n}_{m_0} = \bar{n}_{m_1} = \bar{n}_{m_2} = 4.563/400/200 = 4.563/2 = 2.2815 \cong 2$ or 3.

The sequential analysis of the random samples from Snedecor's data follows:

				m_1 and m_0		m_0 and m_2	
	n =	X	ΣX	UL	LL	UL	LL
First sample:	1	36	36	63	17	43	-3
	2	33	69	103	57	63	17
	3	34	103	143	97	83	37
	4	3	106	183	137	103	57

Random samples from Snedecor's data continued:

				m_1 and m_0		m_0 and m_2	
	n =	X	ΣX	UL	LL	UL	LL
Second sample:	1	26	26	63	17	43	-3
	2	17	43	103	57	63	17
	3	49	92	143	97	83	37
	4	27	119	183	137	103	57
Third sample:	1	30	30	63	17	43	-3
	2	44	74	103	57	63	17
	3	41	115	143	97	83	37
	4	31	146	183	137	103	59
	5	33	179	223	177	123	77
	6	30	209	263	217	143	97

Here one repeatedly obtains the more probable answer, that is, that the samples came from a population which has a mean = 30, and not $\geq$ 50 or $\leq$ 10. The size sample for decision varied from n = 4 to n = 6. It is to be noted that it will always require at least one more than the average sample size ($\bar{n}_{m_0}$) to obtain the decision of no difference from the control mean, unless one extends the lower line of the upper portion of the graph and the upper line of the lower portion of the graph to the left past the point where they intersect. It is preferred not to do this, but rather to require a few extra observations to reach the decision of accept no significant difference from the control. This is an added precaution in order not to reach this decision prematurely. However, when the treatment group mean is more like 50 or 10 than 30, these decisions will be reached on the average with the average sample sizes ($\bar{n}_{m_1}$ or $\bar{n}_{m_2}$).

This can be readily demonstrated by adding 20 to each random observation obtained previously, then again, by subtracting 20 as follows:

		20 added		20 subtracted		m_1 and m_0		m_0 and m_2	
	n =	X	ΣX	X	ΣX	UL	LL	UL	LL
First sample:	1	56	56	16	16	63	17	43	-3
	2	53	109	13	29	103	57	63	17
	3			14	43			83	37
	4			-17	26			103	57
Second sample:	1	46	46	6	6	63	17	43	-3
	2	37	83	-3	3	103	57	63	17
	3	69	152			143	97		
Third sample:	1	47	47	7	7	63	17	43	-3
	2	50	97	10	17*	103	57	63	17*
	3	64	161	24	41	143	97	83	37
	4			21	62			103	57
	5			11	73			123	77

*Prefer not to consider on the line as a decision.

One repeatedly comes to a correct decision in a small number of observations (on the average (2+4+3+2+3+5)/6 = 19/6 = 3.166) which is a reasonable approximation to the estimated average sample size of 2 to 3.

Drolette and Hegsted described a method of analysis referring to Rosander's text, but such an analysis as they presented was not found (26). According to this method of Drolette and Hegsted, the following should be done:

Mean of control = 240

$$\sigma_{\bar{x}} = 48$$

$$m_0 = 215$$

$$m_1 = 265$$

$$\alpha = .02$$

$$\beta = .02$$

$$\bar{n}_{m_0} = \bar{n}_{m_1} = 6.44 \cong 6 \text{ or } 7.$$

If one alters this to:

Mean of control = 240

$\sigma_{\bar{x}} = 48$

$m_0 = 215$

$m_1 = 265$

$\alpha = .01$

$\beta = .01$

One obtains:

$h_0 = -b\sigma^2/(m_1-m_0) = -10586.88/(265-215) = -211.737 \cong -211.7$

$h_1 = +211.7$

$S = (265+215)/2 = 480/2 = 240$

$\bar{n}_{m_0} = \bar{n}_{m_1} = 4.563/2500/4608 = 4.563/.542 = 8.418 \cong 8$ or 9.

The analysis of their data by this model is presented in Table 12.

As pointed out previously, when $\bar{n}_{m_0}$ and $\bar{n}_{m_1}$ are greater than 1, the probable error is more difficult to illustrate empirically, as can be seen in Table 12. However, continued plotting, if additional observations were available, would probably lead to the conclusion that one should accept $m_e \leq m_0$ or $m_e \geq m_1$, either of which would be an improbable conclusion.

If one now applies the model of Armitage to these data, one proceeds as follows:

$m_0 = 240$ (accepted as μ_x of control population)

$\sigma_{\bar{x}} = 48$ (accepted as $\sigma_{\bar{x}}$ of control population)

$m_1 = 265$ (arbitrary, from Drolette and Hegsted)

$m_2 = 215$ (arbitrary, from Drolette and Hegsted)

$\alpha = .01$

$\beta = .01$

One then has:

$$h_0 = -b\sigma^2/(m_1-m_0) = -10,586.88/25 = -423.475 \cong -423.5$$

$$h_1 \cong 423.5$$

$$S_1 = (240+265)/2 = 505/2 = 252.5$$

$$S_2 = (215+240)/2 = 455/2 = 227.5$$

$$n_{m_0} = n_{m_1} = n_{m_2} = 4.563/625/4608 = 4.503/.135 = 33.355$$

This latter is reasonable, as the δ (m_1-m_0 or m_0-m_2) is approximately $1/2\,\sigma_{\bar{x}}$ from $\mu_{\bar{x}}$ -- a very short distance away. It should thus require a larger sample size to make such a decision, since δ is so small. The analysis of their data by this model is shown in Table 13.

By this model of analysis, not taking into account the problem of multiple comparisons in order to remain consistent with Drolette and Hegsted, the decisions are:

1. Ethiopians have a significantly lower cholesterol (decision reached at n = 5 at the .01 level).

2. Boys have a significantly lower cholesterol (decision reached at n = 12 at the .01 level).

3. No decision on adult men was made. One would need at least 34 or 35 observations to reach a decision of no significant difference from control, and there were only 24 observations available.

4. Cardiacs have a significantly higher cholesterol (decision reached at n = 15 at the .01 level).

If everything is recalculated using the model patterned after Armitage, but setting $\alpha = \beta = .02$ (as did Drolette and Hegsted), the average number for decision is now $25.76 \cong 25$ or 26. Essentially the same decisions as above are reached, and a summary chart follows, showing

the actual number of observations observed to be required for a decision in each case:

	Model of Drolette and Hegsted	Model of Armitage	
$\alpha = \beta =$	.02	.02	.01
Ethiopians	3	5	5
Boys	7	12	12
Cardiacs	9	15	15
Adult men	(decision impossible on sequential graph)	(a)	(b)

[a]Need at least 26 or 27 observations for a decision of no significant difference (i.e., accept $m_e = m_0$).

[b]Need at least 34 or 35 observations for a decision of no significant difference (i.e., accept $m_e = m_0$).

Although it took two more observations in Ethiopians and five more in boys and six more in cardiacs to reach similar decisions to those reached by Drolette and Hegsted, it is felt that the decisions reached here by this technique are more meaningful for the reasons already presented.

In summary, the model of sequential analysis for the two-sided-alternative problem in medical research presented here is preferred to the model of Wald because it permits a decision as to whether the mean of the treatment group is higher than, essentially the same as or lower than the mean of the control group, and it is preferred to the model of Drolette and Hegsted because one is less likely to reach improbable decisions of rejecting $m_e = m_0$ when in truth $m_e = m_0$; and one can, on the sequential graph, reach the decision $m_e = m_0$ and thus know when to stop sampling from a population whose mean does not differ from the control mean by a significant amount, without resorting to truncation techniques, or other complex procedures requiring additional calculations.

CHAPTER V

AN APPROXIMATION MODEL OF SEQUENTIAL ANALYSIS THAT SHOULD BE OF INTEREST TO THE MEDICAL RESEARCH WORKER

Making use of the principles discussed and outlined in the previous sections, it should be possible to develop an approximation model of sequential analysis for use in the following situations in medical experimentation:

1. There is not extensive past experience or there are not extensive data previously collected under similar experimental circumstances from which one could obtain satisfactory estimates of μ_x and $\sigma_{\bar{x}}$ of the control or standard group.

2. The experimenter does not know the distribution of data in the population from which his samples of data came.

3. The experimenter is not certain of an exact value of difference from the (unknown) standard or control mean that he would definitely accept as biologically (medically) significant. He would be willing, perhaps, to accept several different values of a significant difference (i.e., several different δ's). The experimenter is willing to decide upon the appropriate δ in conjunction with the advice of the statistician, taking into account the estimates of μ_x of the control group, and of $\sigma_{\bar{x}}$ in all of the experimental groups, that are obtained

from the experiment as it progresses.

4. The experimenter is not disturbed by operating at an approximate level of confidence (i.e., if he operates at the .01 level, it may actually on occasion be .02 or .03) because he will nearly always operate at levels at least presumably smaller than .05.

The Approximation Model of Sequential Analysis Applied to Experimental Sampling Data

The following three examples are offered as empirical evidence that such an approach should be useful to the medical researcher. The general sampling situation for each of the three examples will be the same.

Two random samples of size n = 9 each will be drawn from a known distribution making use of a table of random numbers. Means of the appropriate size will be compared by the sequential analysis, under the assumption of the approximate normality of the distribution of the means.

The first samples from the normal distribution are already available from Chapter III. They are:

Control Group	Treatment Group
36	27
33	36
34	14
32	42
27	17
36	23
25	26
27	11
11	35

It has already been determined that means of samples of size n = 2 should be approximately normally distributed (Chapter III). There are four samples of size 2 in the control group and four in the treatment group.

From these four pairs of samples in the control group, one estimates the mean of the control group as follows:

$$(36+33+34+32+27+36+25+27)/8 = 216/8 = 27.0.$$

This is the estimate of μ_x of the control group population. For the estimate of $\sigma_{\bar{x}}$ of the control and treatment populations (for it is assumed that $\sigma_{\bar{x}_{control}} = \sigma_{\bar{x}_{treatment}}$), one pools the estimates from each of the eight samples of size n = 2.

$$CSS_1 = 2385 - 2380.5 = 4.5$$

$$CSS_2 = 2180 - 2178 = 2.0$$

$$CSS_3 = 2025 - 1984.5 = 40.5$$

$$CSS_4 = 1354 - 1352 = 2.0$$

$$CSS_5 = 2025 - 1984.5 = 40.5$$

$$CSS_6 = 1960 - 1568 = 392.0$$

$$CSS_7 = 818 - 810 = 18.0$$

$$CSS_8 = 797 - 684.5 = 112.5$$

$$S^2_{x_{pooled}} = (4.5+2.0+40.5+2.0+40.5+392.0+18+112.5)/8 = 612/8 = 76.5,$$

$$S_{x_p} \tilde{=} 8.75$$

This is the estimate of $\sigma_{\bar{x}}$. The four means in the treatment group, from the four pairs of samples, are as follows:

1. $(27+36)/2 = 63/2 = 31.5$
2. $(14+42)/2 = 56/2 = 28.0$
3. $(17+23)/2 = 40/2 = 20.0$
4. $(26+11)/2 = 37/2 = 18.5.$

From $S^2_{x_{pooled}}$ one estimates $\sigma_{\bar{x}}$ as $\sqrt{S^2_{x_p}/n}$ for means of size n = 2:

$$S_{\bar{x}} = \sqrt{76.5/2} = \sqrt{38.25} = 6.185$$

With the estimate of μ_x for the control as 27 and the estimate of $\sigma_{\bar{x}}$ (mean of samples of size n = 2) as 6.185, one approaches the experimenter and asks if the lower and upper limits of significant change (i.e., m_2 and m_1) would be reasonable at $27 \pm 1.96\ (6.185) = 27 \pm 12.12 = 14.88$ (m_2) to 39.12 (m_1). A group that gained approximately 15 pounds instead of 27 pounds would be considered to have gained significantly less weight; and, one that gained 39 pounds instead of 27 pounds would be considered to have gained significantly more weight. These are considered biologically reasonable and δ has been set as $1.96\ S_{\bar{x}}$, which is a reasonable difference to look for in terms of what has been learned about the variation from the experimental data. This is equivalent to setting δ^2 as $(1.96)^2\ S_{\bar{x}}^2$: $\left[\delta^2 = 3.841 \times 38.25 = 146.918;\ \delta = 12.12\right]$. It is also equivalent to testing $27 \pm 1.96\ (8.75)$ in the original distribution of the variable. In the sequential model this fixes the denominator of the average sample size = 1.9205 (i.e., $\delta^2/2\sigma^2 = 3.841\sigma^2/2\sigma^2 = 1.9205$), thus the solution of $\left[(1-\alpha)b-\alpha a\right]/1.92 \cong$ the average sample size necessary for decision. With this model, if one selects $\alpha = \beta = k$, one has the average sample sizes recorded in the following table:

When k =	Average Sample Size is:
.02	1.986.....(1 or 2)
.01	2.27.....(2 or 3)
.001	3.59.....(3 or 4)

Arbitrarily, it is elected to perform an analysis at the .01 level of confidence here.

One thus has:

Estimate of $\mu_x = m_0 = 27$

Estimate of $\sigma_{\bar{x}} = 6.185$, $\sigma_{\bar{x}}^2 = 38.25$

$m_1 = 39.12$

$m_2 = 14.88$

$\alpha = \beta = .01$

$a = b = 4.595$[1]

$h_0 = -b\sigma_{\bar{x}}^2/(m_1-m_0) = (-4.595 \times 38.25)/12.12 = -175.76/12.12 = -14.50$

$h_1 = 14.50$

$S_1 = (m_0+m_1)/2 = (27+39.12)/2 = 66.12/2 = 33.06$

$S_2 = (m_2+m_0)/2 = (14.88+27)/2 = 41.88/2 = 20.94$

$\bar{n}_{m_0} = \bar{n}_{m_1} = \bar{n}_{m_2} = 2.27 \cong 2$ or 3.

The sequential analysis thus becomes:

	Treatment $\bar{X}$		m_1 and m_0		m_0 and m_2	
N =	$\bar{X}$	$\sum\bar{X}$	UL	LL	UL	LL
1	31.5	31.5	47.56	18.56	35.44	6.44
2	28.0	59.5	80.62	51.62	56.38	27.38
3	20.0	79.5	113.68	84.68	77.32	48.32

---

[1]Under the conditions of $\alpha > 0, \beta > 0$ and $\alpha + \beta < 1.0$, Wald has defined $a = \ln (1-\beta)/\alpha$ and $b = \ln (1-\alpha)/\beta$. Thus:

$\alpha e^a = 1 - \beta$ $\qquad$ $\beta e^b = 1 - \alpha$

$\alpha = e^{-a} - \beta e^{-a}$ $\qquad$ $\alpha = 1 - \beta e^b$

$e^{-a} - \beta e^{-a} = 1 - \beta e^b$

$\beta = (1-e^{-a})/[e^b(1-e^{-a}e^{-b})]$

$= e^{-b}[(1-e^{-a})/(1-e^{-a}e^{-b})]$

Therefore, β approaches e^{-b} as a or a + b increase.

Similarly, α approaches e^{-a} as b or a + b increase.

For the special case where $\alpha = \beta = \gamma$ and a = b = k, one may say γ approaches e^{-k} as k increases. The above approximations are quite satisfactory in practical applications, especially when α = .05 or less and β = .05 or less. An example of the special case, selecting a = b = 4, and referring to an exponential table yields $e^{-4} = 0.0183$, when the true probability obtained from $(e^b-1)/(e^ae^b-1)$ is $0.01798 \cong 0.0180$.

On plotting the third mean, one enters the accept as no different from the standard or control (i.e., accept $m_e = m_0 = 27$), (i.e., reject $m_e \leq 14.88$ and reject $m_e \geq 39.12$). Since in this instance the treatment or experimental group was a random sample from the same identical universe as was the control or standard group, it is reassuring that it was accepted as such by this approximation model of analysis.

Proceeding now to the data from the rectangular distribution from Table 2, the results of random sampling are:

Control Group	Treatment Group
7	8
2	5
2	8
6	6
8	4
1	1
4	3
2	3
3	4

Evidence has already been established that means of samples of size n = 2 should be approximately normally distributed. The situation is thus the same as in our first example, and the results are summarized as follows:

Estimate of μ_x of control population:

$$= (7+2+2+6+8+1+4+2)/8 = 4.0$$

Estimate of σ_x of control population:

$$S^2_{x_{pooled}} = (13.5+8.0+24.5+7.0+4.5+2.0+4.5+0)/8 = 64/8 = 8,$$

$$S_{x_p} = \sqrt{8} = 2.828 \cong 2.83$$

The four means in the treatment group are:

1. $(8+5)/2 = 13/2 = 6.5$
2. $(8+6)/2 = 14/2 = 7.0$

3. $(4+1)/2 = 5/2 = 2.5$

4. $(3+3)/2 = 6/2 = 3.0$

From $S^2_{x_p}$ one estimates $\sigma_{\bar{x}}$ as follows:

$$S_{\bar{x}} = \sqrt{S^2_p/n} = \sqrt{8/2} = \sqrt{4} = 2.0$$

To obtain m_1 and m_2, one takes $4.0 \pm 1.96\ (2) = 4 \pm 3.92 = 0.08\ (m_2)$ to $7.92\ (m_1)$.

It is thus going to be determined if the mean of the treatment or experimental population is the same as that of the control (i.e., $m_e = m_0$) or is .08 or less ($m_e \leq m_2$) or 7.92 or more ($m_e \geq m_1$). Here one cannot consult the biological or medical experimenter for the reasonableness of this test, as it is not meaningful data in that field. One thus accepts this δ as at least as reasonable as any other arbitrary δ selected, and more reasonable than many, for it has taken into account the variation inherent in the experimental data. The average sample sizes for various values of $\alpha = \beta = k$ here will be the same as in the previous example. Again, the analysis is arbitrarily conducted at the .01 level of confidence.

One thus has:

Estimate of $\mu_x = m_0 = 4$

Estimate of $\sigma_{\bar{x}} = 2.0$

$m_1 = 7.92$

$m_2 = 0.08$

$\alpha = \beta = .01$

$h_0 = -b\sigma^2_{\bar{x}}/(m_1-m_0) = (-4.595 \times 4)/3.92 = 4.688 \cong -4.7$

$h_1 \cong 4.7$

$S_1 = (7.92 \times 4)/2 = 11.92/2 = 5.96$

$S_2 = (.08+4.00)/2 = 4.08/2 = 2.04$

$\bar{n}_{m_0} = \bar{n}_{m_1} = \bar{n}_{m_2} = 2.27 \cong 2$ or 3.

The sequential analysis follows:

			m_1 and m_0		m_0 and m_2	
N =	$\bar{X}$	$\sum\bar{X}$	UL	LL	UL	LL
1	6.5	6.5	10.66	1.26	6.74	-2.66
2	7.0	13.5	16.62	7.22	8.78	-0.62
3	2.5	16.0	22.58	13.18	10.82	1.42
4	3.0	19.0	28.54	19.14	12.86	3.46

On plotting the fourth mean here one enters the accept $m_e = m_0$ zone and thus rejects $m_e \geq m_1$ (7.92) and $m_e \leq m_2$ (0.08). Since again the treatment or experimental group was a random sample from the same (originally rectangular) universe as was the control group, it is reassuring that it was accepted as such by this technique of sequential analysis.

Proceeding now to the data from the logarithmic distribution from Table 3, it has already been learned that one should work with samples of size n = 3 in order for the means to approximate normality. The procedure is much the same, but here there are fewer samples of size n in the original nine observations in the control and experimental group.

The results of the random sampling are as follows:

Control Group	Treatment Group
1.9955	2.0895
2.3445	1.2885
2.0415	1.5485
2.5115	2.1375
1.9495	1.9955
1.6215	2.6915
1.9955	1.9955
1.6985	1.4125
2.1875	1.0715

The estimate of $\mu_{\bar{x}}$ of the control population is:

$$(1.9955+2.3445+\ldots+\ldots2.1875)/9 = 18.3455/9 = 2.0383$$

The estimate of $\sigma_{\bar{x}}$ of the control population is:

$$S^2_{x_{pooled}} = (.0719+.4052+.3339+.4366+.2705+.1215)/(2+2+2+2+2+2)$$

$$= 1.6396/12 = 0.1366$$

$$S_{x_p} = \sqrt{.1366} = .3696$$

The three means in the treatment group are:

1. $4.9265/3 = 1.6421$
2. $6.8245/3 = 2.2748$
3. $4.4795/3 = 1.4931$

From $S^2_{x_p}$ one estimates $\sigma_{\bar{x}}$ as:

$$S_{\bar{x}} = \sqrt{S^2_{x_p}/3} = \sqrt{.1366/3} = \sqrt{.0455} = .213$$

To obtain m_1 and m_2, one takes $2.038 \pm 1.96\ (.213) = 2.038 \pm .417 =$ 1.621 (m_2) to 2.455 (m_1).

The test will now be made to determine if the mean of the treatment or experimental population is essentially the same as that of the control ($m_e = m_0$) or is 1.621 or less ($m_e \leq m_2$) or 2.455 or more ($m_e \geq m_1$). One now has:

Estimate of $\mu_{\bar{x}} = m_0 = 2.038$

Estimate of $\sigma_{\bar{x}} = S_{\bar{x}} = 0.213$

$m_1 = 2.455$

$m_2 = 1.621$

$\alpha = \beta = .01$

$$h_0 = -b\sigma^2_{\bar{x}}/(m_1-m_0) = \left[-4.595(.0455)\right]/.417 = -.2091/.417 = -0.5014$$

$$\cong -0.501$$

$$h_1 \cong 0.501$$

$S_1 = (2.455+2.038)/2 = 4.493/2 = 2.2465$

$S_2 = (2.038+1.621)/2 = 3.659/2 = 1.8295$

$\bar{n}_{m_0} = \bar{n}_{m_1} = \bar{n}_{m_2} = 2.27$

The sequential analysis is:

			m_1 and m_0		m_0 and m_2	
N =	$\bar{X}$	$\sum\bar{X}$	UL	LL	UL	LL
1	1.6421	1.6421	2.7479	1.7451	2.3309	1.3281
2	2.2748	3.9169	4.9944	3.9944	4.1604	3.1576
3	1.4931	5.4100	7.2409	6.2381	5.9899	4.9871
		Results of the necessary additional sampling:				
4	2.2281	7.6381	9.4874	8.4846	7.8194	6.8166
5	2.1985	9.8366	11.7339	10.7311	9.6489	8.6461
6			13.9804	12.9796	11.4784	10.4756

Note: Here, with the original three means in the treatment group no decision was reached (i.e., remained in the zone of indecision). It was thus necessary to obtain another random sample of size n = 3 for the treatment group, and it still remained in the zone of indecision. Another random sample of size n = 3 was obtained from the logarithmic distribution, and with this mean of the fifth sample of size n = 3 it entered the zone of accept $m_e = m_0$; thus one rejects $m_e \leq m_2$ and $m_e \geq m_1$. This is again reassuring as all these samples in the treatment group came from the same originally logarithmic distribution as did the control group sample.

The examples which have been presented above are not selected for their clarity or because they worked out well, but are the results of random sampling as it actually occurred. The implication is not intended that occasionally a wrong decision will not be made, for surely it will. It so happens none were obtained here. Wrong or improbable answers may result from a poor estimate of μ_x, or of $\sigma_{\bar{x}}$ or an unfortunate series of

apparently biased samples in the treatment group, thus a poor estimate of m_e. It has been stressed that although it is set that $\alpha = \beta = .01$, the probability levels are not exactly .01, but approximately so. Every effort has been made to obtain a good estimate of μ_x, $\sigma_{\bar{x}}$ and m_e, as well as to assure reasonable approximation to normality of the distribution of means to be compared. The assumption that the control and treatment samples came from populations with essentially homogeneous variances is also necessary and should be tested by the F ratio, because one pools the estimate of $\sigma_{\bar{x}}$ from both the control and the treatment group samples.

The Approximation Model of Sequential Analysis Applied to Actual Medical Research Data

Having demonstrated empirically by random sampling that this model of analysis is useful in testing random samples from a known population such as a normal, a rectangular and a logarithmic distribution, its use will be examined now in some actual experimental situations in medical research.

The example of Drolette and Hegsted has already been cited, but it is not quite the same, as here one had a good estimate of μ_x and σ_x before the experiment was begun. Also, the δ was selected in another more arbitrary way, not taking into account the variation in the data, but using other expert opinion methods. There are those who might challenge that a blood cholesterol of 265 was much different from one of 240, or that one of 215 was much different from one of 240. However, taking $240 \pm 1.96\ (48) = 240 \pm 94.08$ gives 145.92 for m_2 and 334.08 for m_1. There are few, if any, who would challenge the biological or medical significance of a cholesterol of 145.92 compared to 240, or one of 334.08

compared to 240 -- most would regard the former as hypocholesterolemic and the latter as hypercholesterolemic, and the 240 as essentially a "normal" value. One could reanalyze Drolette and Hegsted's data from the above point of view, and in addition ignore the previous work of others done in different laboratories at different times. To avoid the multiple comparison issue here, one might wish only to compare the adult men in the United States (as the control) to the Ethiopians. Assuming the samples are random (for they are used as such by Drolette and Hegsted) the procedures previously outlined in this dissertation will be followed, designating the adult men in the United States as the control group and the Ethiopians as the treatment group.

Control Group (Adult Men in U.S.A.)	Treatment Group (Ethiopians)
255	227
189	103
272	163
308	134
248	93
279	153
230	129
238	163
267	227

The pairs of samples of size n = 2 thus become:

From the Control Group:

Pair 1:	255	272
	189	308
Pair 2:	248	230
	278	238

From the Treatment Group:

Pair 1:	227	163
	103	134
Pair 2:	93	129
	153	163

The results of the testing of the cholesterol data to see if the means of samples of size n = 2 will be approximately normally distributed are presented in the following table:

	$S_{\bar{x}}$ from Range	$S_{\bar{x}}$ from Two Means	$S_{\bar{x}}$ from $S^2_{x_p}$
First sample pair:	60.18	48.08	26.58
		(approximate 95% CI = 0 to 53.16)	

It is necessary to proceed to the testing of means of samples of size n = 3:

First sample pair:	23.43	28.00	21.69
		(approximate 95% CI = 3.99 to 39.39)	
Second sample pair:	22.25	26.59	28.24
		(approximate 95% CI = 5.18 to 51.30)	

Therefore, accept means of sample size n = 3 to be approximately normally distributed.

The estimate of μ_x for the control group is:

$$(255+189+...+...238+267)/9 = 2286/9 = 254.0$$

The estimate of σ_x^2 for the control group is:

$$S^2_{x_{pooled}} = (3844.7+1800.7+7690.7+1880.7+758+4952)/(2+2+2+2+2+2)$$

$$= 20,846.8/12 = 1737.23$$

$$S_{x_p} = 41.68$$

The three means in the treatment group are:

1. 493/3 = 164.3
2. 380/3 = 126.7
3. 519/3 = 173.0

From $S^2_{x_p}$ one estimates $\sigma_{\bar{x}}$ as before:

$$S_{\bar{x}} = \sqrt{1737.23/3} = \sqrt{579.08} = 24.06$$

To obtain m_1 and m_2, one proceeds as follows:

$$254 \pm 1.96\ (24.06) = 254 \pm 47.16$$

One now has: 206.84 (m_2) to 301.16 (m_1).

It will be tested to determine if the mean of the treatment group is essentially the same as that of the control group ($m_e = m_0$) or more (i.e., $m_e \geq m_1 = 301.16$) or less (i.e., $m_e \leq m_2 = 206.84$), testing means of samples of size n = 3.

One thus has:

Estimate of $\mu_{\bar{x}} = m_0 = 254$

Estimate of $\sigma_{\bar{x}} = S_{\bar{x}} = 24.06$

$m_1 = 301.16$

$m_2 = 206.84$

$\alpha = \beta = .01$

$h_0 = -b\sigma_{\bar{x}}^2/(m_1-m_0) = (-4.595 \times 579.08)/47.16$

$= -2660.8726/47.16 = -56.422$

$h_1 = 56.42$

$S_1 = (301.16+254)/2 = 555.16/2 = 277.58$

$S_2 = (254+206.84)/2 = 460.84/2 = 230.42$

$\bar{n}_{m_0} = \bar{n}_{m_1} = \bar{n}_{m_2} = 2.27$

The sequential analysis with regard to Ethiopians is as follows:

			m_1 and m_0		m_0 and m_2	
N =	$\bar{X}$	$\sum\bar{X}$	UL	LL	UL	LL
1	164.3	164.3	334.00	221.16	286.84	174.0
2	126.7	291.0	611.58	498.74	517.26	404.42
3	173.0	464.0	889.16	776.32	747.68	634.84

It is seen at once that the Ethiopians have a significantly lower serum cholesterol than the adult men of the United States. As occurred here, when a decision is reached on the first plotted mean, it is felt to

be desirable to continue to plot at least two more means to insure that the decision is not altered (i.e., that one has not obtained a rare or freak random sample on the first trial in the treatment group).

For completeness, if one ignores the problems inherent in multiple comparisons as Drolette and Hegsted have done, one may compare the other groups as follows:

	Boys		Cardiacs		m_1 and m_0		m_0 and m_2	
N =	$\bar{X}$	$\sum\bar{X}$	$\bar{X}$	$\sum\bar{X}$	UL	LL	UL	LL
1	242.7	242.7	281.7	281.7	334.00	221.16	286.84	174.0
2	167.0	409.7	266.0	547.7	611.58	498.74	517.26	404.42
3	169.3	579.0	253.0	800.7	889.16	776.32	747.68	634.84
4			282.7	1083.4	1166.74	1053.9	978.10	865.26
5			295.3	1378.7	1444.32	1331.48	1208.52	1095.68
6					1721.90	1609.06	1438.94	1326.10
7					1999.48	1886.64	1669.36	1556.52

With this method of analysis, one may conclude that Ethiopians have a significantly lower serum cholesterol than do adult men of the United States (one mean of sample size n = 3 needed), that boys have a significantly lower serum cholesterol than do adult men (three means of sample size n = 3 needed), and no decision is as yet reached with the cardiac group (five means of sample size n = 3). One would need to continue sampling in this group, but from the trend indicated from the first five sample means, one may anticipate that in not too many more samples it will be demonstrated that the mean of the cardiac group will be significantly higher than that of the adult male controls of the United States. The decisions reached by this model of sequential analysis on these data are probably more meaningful to the clinician interested in serum cholesterol levels in various population groups than are those offered by Drolette and Hegsted.

Further examples will now be cited from data locally collected at the University of Oklahoma Medical Center. Data from four experiments are cited and reported in Tables 4, 5, 6 and 8 respectively. Because the fasting levels of the placebo and Reserpine groups differ somewhat in Table 6, the differences are taken as the treatment response values to be compared here.

Essentially the same procedure will be used in the following analyses, but it will not be described in the same detail. Only pertinent results will be listed. The control (placebo) and treatment group data presented have been analyzed by the more standard appropriate t-test, and in each instance the treatment group was found to differ significantly statistically from the control group (i.e., $p < .05$). One would hope that the conclusions from the approximation model of sequential analysis would be similar.

Data from Table 4:

Control Group	Treatment Group
217	251
144	355
148	311
178	228
141	229
161	256
131	292
163	296
139	252

Data from Table 5:

Control Group	Treatment Group
55	249
76	100
72	143
45	106
51	180
61	128
86	154
105	196
67	161

Data from Table 6:

Control Group	Treatment Group
- 4	+45
+ 4	+38
+ 3	+26
- 3	+30
+ 8	+68
- 6	+30
- 6	+34
+ 6	+53
-12	+19

Data from Table 8:

Control Group	Treatment Group
- 4	-22
- 1	-22
+ 1	- 2
-12	-23
- 2	-26
- 1	-17
- 2	-20
-14	- 9
- 9	-13

Table 14 presents the evidence suggesting that the means of samples of size n = 4 for the data in Table 4, size n = 3 for the data in Table 5, size n = 5 for the data in Table 6, and size n = 3 for the data in Table 8 should be approximately normally distributed.

For clarity, before proceeding to the sequential analysis, the data from Tables 4, 5 and 6 are presented again in light of the above.

Data from Table 4:

Control Group	Treatment Group	
217	251	224
144	355	332
148	311	141
178	228	333
141	229	280
161	256	223
131	292	293
163	296	248
	252	219
	212	
	242	

Data from Table 5:

Control Group	Treatment Group	
55	249	188
76	100	143
72	143	87
45	106	150
51	180	78
61	128	92
86	154	178
105	196	149
67	161	158
	173	117
	98	

Data from Table 6:

Control Group	Treatment Group	
- 4	+45	+56
+ 4	+38	+ 2
+ 3	+26	+12
- 3	+30	+24
+ 8	+68	+52
- 6	+30	
- 6	+34	
+ 6	+53	
-12	+19	
+ 6	+42	

Proceeding now to employ the approximation model of sequential analysis as described earlier in this section, one has from the data in Table 4:

Estimate of μ_x of control group = m_0 = (217+144+.....+131+163)/8

= 1283/8 = 160.4

Estimate of σ_x^2 of control group = $S_{x_p}^2$ = (3420.75+728+9974.75+3024.75)/ (3+3+3+3) = 17,148.25/12 = 1429.02

S_{x_p} = 37.8

The means in the treatment group are:

1. 1145/4 = 286.3
2. 1073/4 = 268.3

3. $930/4 = 232.5$

4. $1086/4 = 271.5$

5. $983/4 = 245.8$

From $S^2_{x_p}$ one obtains the estimate of $\sigma_{\bar{x}}$:

$$S_{\bar{x}} = \sqrt{1429.02/4} = \sqrt{357.26} = 18.90$$

To obtain m_1 and m_2, one takes:

$$160.4 \pm 1.96\ (18.9)$$

$$= 160.4 \pm 37.0$$

$$= 123.4\ (m_2) \text{ to } 197.4\ (m_1)$$

Setting $\alpha = \beta = .01$, one now has the necessary material to perform the analysis:

Estimate of $\mu_x = m_0 = 160.4$

Estimate of $\sigma_{\bar{x}} = S_{\bar{x}} = 18.9$

$$\alpha = \beta = .01$$

$$m_1 = 197.4$$

$$m_2 = 123.4$$

$$h_0 = -b\sigma_{\bar{x}}^2/(m_1-m_0) = \left[-4.595(357.26)\right]/37 = -1641.6/37 \cong -44.37$$

$$h_1 \cong +44.37$$

$$S_1 = (m_1+m_0)/2 = (160.4+197.4)/2 = 357.8/2 = 178.9$$

$$S_2 = (m_0+m_2)/2 = (123.4+160.4)/2 = 283.8/2 = 141.9$$

$$\bar{n}_{m_0} = \bar{n}_{m_1} = \bar{n}_{m_2} \cong 2.27$$

			m_1 and m_0		m_0 and m_2	
N =	$\bar{X}$	$\sum\bar{X}$	UL	LL	UL	LL
1	286.3	286.3	223.60	134.53	182.27	97.53
2	268.3	554.6	402.5	313.43	328.17	239.43
3	232.5	787.1	581.4	492.33	470.07	381.33

The decision is reached on plotting the first treatment mean and

reconfirmed by plotting the next two treatment means to accept $m_e \geq m_1$, thus reject $m_e = m_0$ and $m_e \leq m_2$.

From the data in Table 5:

Estimate of μ_x of control group = m_0 = (55+76+.....105+67)/9 = 618/9 = 68.7

Estimate of σ_x^2 of control group = $S_{x_p}^2$ =(248.7+130.7+722+1176.2+2888+1012.7)/ (2+2+2+2+2+2) = 16,764.1/12 = 1397.0

S_{x_p} = 37.38

The means in the treatment group are:

1. 492/3 = 164.0
2. 414/3 = 138.0
3. 511/3 = 170.3
4. 459/3 = 153.0
5. 380/3 = 126.7
6. 348/3 = 116.0
7. 424/3 = 141.3

From $S_{x_p}^2$ one obtains the estimate of $\sigma_{\bar{x}}$:

$S_{\bar{x}} = \sqrt{1397/3} = \sqrt{465.7} = 21.58$

To obtain m_1 and m_2 one takes:

$68.7 \pm 1.96\ (21.58)$

$= 68.7 \pm 42.3$

$= 26.4\ (m_2)$ to $111.0\ (m_1)$

Setting $\alpha = \beta = .01$, one now has the necessary material to perform the analysis:

Estimate of $\mu_x = m_0 = 68.7$

Estimate of $\sigma_{\bar{x}} = S_{\bar{x}} = 21.58$

$m_1 = 111.0$

$m_2 = 26.4$

$\alpha = \beta = .01$

$h_0 = \left[-4.595(465.7)\right]/42.3 = -2139.89/42.3 \cong -50.5$

$h_1 \cong +50.5$

$S_1 = (68.7+111)/2 = 179.7/2 = 89.85$

$S_2 = (26.4+68.7)/2 = 95.1/2 = 47.55$

$\bar{n}_{m_0} = \bar{n}_{m_1} = \bar{n}_{m_2} \cong 2.27$

			m_1 and m_0		m_0 and m_2	
N =	$\bar{X}$	$\sum\bar{X}$	UL	LL	UL	LL
1	164	164	140.45	39.25	98.15	- 3.05
2	138	302	230.30	129.10	145.70	+44.50
3	170.3	472.3	320.15	218.95	193.25	92.05

The decision is reached again on plotting the first treatment mean and confirmed by plotting the next two treatment means to accept $m_e \geq m_1$, thus reject $m_e = m_0$ and $m_e \leq m_2$.

From the data in Table 6:

Estimate of μ_x of control group = m_0 = (-4+4+3...-1-+6)/10 = (+27-31)/10

= -4/10 = -0.40

Estimate of $\sigma_{\bar{x}}^2$ of control group = $S_{x_p}^2$ =(101.2+259.2+1099.2+603.2)/(4+4+4+4)

= 2062.8/16 = 128.93

$S_{x_p} = 11.35$

The means in the treatment group are:

1. 207/5 = +41.4
2. 178/5 = +35.6
3. 146/5 = +29.2

From $S_{x_p}^2$ one obtains the estimate of $\sigma_{\bar{x}}$:

$S_{\bar{x}} = \sqrt{128.93/5} = \sqrt{25.79} = 5.08$

To obtain m_1 and m_2, one takes:

$$-0.40 \pm 1.96\ (5.08)$$

$$= -0.40 \pm 9.96$$

$$= -10.36\ (m_2)\ \text{to}\ +9.56\ (m_1)$$

Setting $\alpha = \beta = .01$, one now has the necessary material to perform the analysis:

Estimate of $\mu_x = m_0 = -0.40$

Estimate of $\sigma_{\bar{x}} = S_{\bar{x}} = 5.08$

$m_1 = +9.56$

$m_2 = -10.36$

$\alpha = \beta = .01$

$h_0 = \left[-4.595(25.79)\right]/9.96 = -118.505/9.96 \cong -11.89$

$h_1 \cong +11.89$

$S_1 = (-0.40+9.56)/2 = 9.46/2 = 4.58$

$S_2 = -10.36 + (-0.4) = -10.76/2 = -5.38$

$\bar{n}_{m_0} = \bar{n}_{m_1} = \bar{n}_{m_2} \cong 2.27$

			m_1 and m_0		m_0 and m_2	
N =	$\bar{X}$	$\Sigma\bar{X}$	UL	LL	UL	LL
1	+41.4	+ 41.4	+16.62	-7.16	+6.51	-17.27
2	+35.6	+ 77.0	+21.35	-2.43	+1.13	-22.65
3	+29.2	+106.2	+26.08	+2.30	-4.25	-28.03

The decision is reached on plotting the first treatment mean and confirmed by plotting the next two treatment means to accept $m_e \geq m_1$, thus reject $m_e = m_0$ and $m_e \leq m_2$.

From the data in Table 8:

Estimate of μ_x of control group = $m_0 = (-4-1+1.....-14-9)/9 = -44/9$

$= -4.89$

Estimate of $\sigma_{\bar{x}}$ of placebo control group:

$$S^2_{x_p} = (12.7+74+266.7+42+72.7+62)/(2+2+2+2+2+2) = 530.1/12 = 44.18$$

Estimate of $\sigma_{\bar{x}} = S_{\bar{x}} = \sqrt{44.18/3} = \sqrt{14.73} = 3.84$

The means in the treatment group are:

1. $-46/3 = -15.3$
2. $-66/3 = -22.0$
3. $-42/3 = -14.0$

To obtain m_1 and m_2, one takes:

$$-4.89 \pm 1.96\ (3.84)$$

$$= -12.57\ (m_2)\ \text{to}\ +2.79\ (m_1)$$

Setting $\alpha = \beta = .01$, one has:

Estimate of $\mu_x = m_0 = -4.89$

Estimate of $\sigma_{\bar{x}} = S_{\bar{x}} = 3.84$

$$m_1 = +2.79$$

$$m_2 = -12.57$$

$$\alpha = \beta = .01$$

$$h_0 = \left[-4.595(14.73)\right]/7.68 = -67.68/7.68 \cong -8.81$$

$$h_1 \cong 8.81$$

$$S_1 = \left[2.79+(-4.89)\right]/2 = -2.10/2 = -1.05$$

$$S_2 = \left[-4.89+(-12.57)\right]/2 = -17.46/2 = -8.73$$

$$\bar{n}_{m_0} = \bar{n}_{m_1} = \bar{n}_{m_2} \cong 2.27$$

			m_1 and m_0		m_0 and m_2	
N =	$\bar{X}$	$\sum\bar{X}$	UL	LL	UL	LL
1	-15.3	-15.3	+7.76	- 9.86	+ 0.08	-17.54
2	-22.0	-37.3	+6.71	70.91	- 8.65	-26.27
3			+5.06	-11.96	-17.38	-35.00

Upon plotting the second mean from the treatment group, the

decision is reached to accept m_e equivalent to or less than m_2, thus reject m_e equivalent to m_0 and m_e equivalent to or greater than m_1.

CHAPTER VI

SUMMARY AND CONCLUSIONS

The models of sequential analysis were originally developed for quality control purposes. Incomplete data and unknown or nonnormal distributions in medical research create problems which differ sufficiently from those of industrial quality control as to require special consideration.

A two-sided alternative approximation model of sequential analysis has been developed for measured data normally distributed as described by Armitage. It makes use of the approach to normality achieved by the means of random samples of appropriate size n (central limit theorem).

Empirical evidence of the model's usefulness has been presented by its application, not only to random samples from known distributions (normal, rectangular and logarithmic), but also to five sets of data from medical research.

There is need for development of a model of sequential analysis permitting correction for initial value effects and multiple comparisons.

BIBLIOGRAPHY

1. Statistical Research Group, Columbia University: Sequential Analysis of Statistical Data: Applications, Columbia University Press, New York, 1945.

2. Wald, A.: Sequential Analysis, John Wiley and Sons, New York, 1947.

3. Oakland, G. B.: An Application of Sequential Analysis to White Fish Sampling, Biometrics 6: 59, 1950.

4. Fisher, R. A.: Sequential Experimentation, Biometrics 8: 183, 1952.

5. Bross, I. B.: Sequential Medical Plans, Biometrics 8: 188, 1952.

6. Burington, R. S. and May, D. C.: Handbook of Probability and Statistics with Tables, Handbook Pub., Inc., Sandusky, Ohio, 1953.

7. Armitage, P.: Sequential Tests in Prophylactic and Therapeutic Trials, Quart. Journal of Medicine 23: 255, 1954.

8. Morris, R. S.: Sequential Sampling Technique for Spruce Bud Worm Egg Surveys, Canadian Journal of Zoology 32: 302, 1954.

9. Kilpatrick, G. S. and Oldham, P. D.: Calcium Chloride and Adrenalin as Bronchial Dilators Compared by Sequential Analysis, Brit. Med. J. 2: 1388, 1954.

10. Waters, W. E.: Sequential Sampling in Forest Insect Surveys, Journ. For. Si., I: 68, 1955.

11. Barthalomay, A. F.: The Sequential Probability Ratio Test, J. Clin. Invest. 35: 688, 1956.

12. Vallee, B. L.; Wacher, W. E. C.; Barthalomay, A. F. and Robin, E. D.: Zinc Metabolism in Hepatic Dysfunction I, New England J. Med. 255: 403, 1956.

13. Newton, D. R. L. and Tanner, J. M.: N-Acetyl-Paraamino-Phenol as an Analgesic: A Controlled Clinical Trial Using the Method of Sequential Analysis, Brit. Med. J. 2: 1096, 1956.

14. Snell, F. M. and Armitage, P.: Critical Comparison of Diamorphine and Pholcodine as Cough Suppressants by a New Method of Sequential Analysis, Lancet 1: 860, 1957.

15. Doering, C. R.; Hagans, J. A.; Ashley, F. W.; Clark, M. L. and Wolf, S.: Sequential Analysis in Therapeutic Research I, The J. of Laboratory and Clinical Medicine 50: 621, 1957.

16. Hagans, J. A.; Doering, C. R.; Clark, M. L.; Schneider, E. M. and Wolf, S.: Sequential Analysis in Therapeutic Research II, The J. of Laboratory and Clinical Medicine 50: 629, 1957.

17. Armitage, P.: Restricted Sequential Procedures, Biometrika, 44: 9, 1957.

18. Armitage, P.: Sequential Methods in Clinical Trials, Am. J. Pub. Health 48: 1395, 1958.

19. Bross, I. D.: Sequential Clinical Trials, J. Chron. Dis. 8: 349, 1958.

20. Billewicz, W. Z.: Some Practical Problems in a Sequential Medical Trial, Bull. Int. Inst. Stat. 36: 165, 1958.

21. Cass, L. J.; Frederik, W. S. and Barthalomay, A. F.: Clinical Evaluation of Ethoheptazine and Othoheptazine Combined with Aspirin, J. Am. Med. Assoc. 166: 1829, 1958.

22. Olmstead, J. G.: Sequential Analysis in the Assignment of Chromosomal Sex by Means of Nuclear Chromatin Patterns in Smears of Oral Mucosa, Am. J. Clin. Path. 32: 346, 1959.

23. Hegsted, D. M. and Drolette, M. E.: Sequential Analysis in Nutrition Studies, The Amer. J. of Clin. Nutrition 8: 112, 1959.

24. West, K. M., University of Oklahoma School of Medicine, Oklahoma City, Oklahoma: Personal Communication.

25. Pearson, E. S. and Hartley, H. O.: Biometrika Tables for Statisticians, 1: 165, 1956.

26. Rosander, A. C.: Elementary Principles of Statistics, D. Van Nostrand Co., Inc., New York, N. Y., 1957.

27. Snedecor, G. W.: Statistical Methods, The Iowa State College Press, Ames, Iowa, 1956.

28. Arkin, H. and Colton, R. R.: Tables for Statisticians, Barnes and Noble, Inc., New York, N. Y., 1953.

29. Klett, James C. and Lasky, Julian: A Clinical Trial of Five Phenothiazines Using Sequential Analysis, Report #2, Central Neuropsychiatric Research Laboratory, Perry Point, Maryland, February, 1960.

30. Mock, D. C., Kyriakopolous, A. and Clark, M. L., University of Oklahoma School of Medicine, Oklahoma City, Oklahoma: Personal Communication.

APPENDIX

TABLE 1

AN APPROXIMATELY NORMAL DISTRIBUTION[a]

Number	Result	Number	Result	Number	Result	Number	Result
1	3	26	24	51	30	76	37
2	7	27	24	52	30	77	37
3	11	28	24	53	30	78	38
4	12	29	25	54	30	79	38
5	13	30	25	55	30	80	39
6	14	31	25	56	31	81	39
7	15	32	26	57	31	82	39
8	16	33	26	58	31	83	40
9	17	34	26	59	31	84	40
10	17	35	26	60	32	85	41
11	18	36	27	61	32	86	41
12	18	37	27	62	33	87	41
13	18	38	27	63	33	88	42
14	19	39	28	64	33	89	42
15	19	40	28	65	33	90	42
16	19	41	28	66	33	91	43
17	20	42	29	67	34	92	43
18	20	43	29	68	34	93	44
19	21	44	29	69	34	94	45
20	21	45	29	70	35	95	46
21	21	46	30	71	35	96	47
22	22	47	30	72	35	97	48
23	22	48	30	73	36	98	49
24	23	49	30	74	36	99	53
25	23	50	30	75	36	100	57

[a]From Snedecor (27).

TABLE 2

AN APPROXIMATELY RECTANGULAR DISTRIBUTION[a]

Number	Result	Number	Result	Number	Result	Number	Result
1	0	26	1	51	4	76	9
2	4	27	5	52	8	77	2
3	8	28	9	53	2	78	5
4	3	29	2	54	6	79	9
5	7	30	6	55	0	80	2
6	0	31	4	56	1	81	3
7	4	32	7	57	5	82	7
8	8	33	1	58	9	83	1
9	2	34	4	59	3	84	6
10	7	35	8	60	6	85	0
11	4	36	2	61	3	86	1
12	8	37	5	62	7	87	6
13	2	38	8	63	0	88	0
14	6	39	2	64	4	89	4
15	0	40	5	65	8	90	8
16	2	41	1	66	9	91	4
17	6	42	6	67	3	92	8
18	9	43	0	68	7	93	2
19	3	44	4	69	0	94	6
20	7	45	8	70	4	95	0
21	4	46	1	71	1	96	0
22	8	47	5	72	5	97	4
23	2	48	9	73	8	98	8
24	5	49	3	74	2	99	2
25	9	50	7	75	5	100	6

Result	Observed f	Theoretical f
0	11	10
1	8	10
2	13	10
3	7	10
4	13	10
5	9	10
6	10	10
7	8	10
8	13	10
9	8	10

[a]From Arkin and Colton: Five place tables of common logarithms (28).

TABLE 3

AN APPROXIMATELY LOGARITHMIC DISTRIBUTION[a]

Number	Result	Number	Result	Number	Result	Number	Result
1	1.0715	26	1.7375	51	1.9955	76	2.3445
2	1.1745	27	1.7375	52	1.9955	77	2.3445
3	1.2885	28	1.7375	53	1.9955	78	2.3985
4	1.3185	29	1.7785	54	1.9955	79	2.3985
5	1.3485	30	1.7785	55	1.9955	80	2.4545
6	1.3805	31	1.7785	56	2.0415	81	2.4545
7	1.4125	32	1.8195	57	2.0415	82	2.4545
8	1.4455	33	1.8195	58	2.0415	83	2.5115
9	1.4795	34	1.8195	59	2.0415	84	2.5115
10	1.4795	35	1.8195	60	2.0895	85	2.5705
11	1.5135	36	1.8625	61	2.0895	86	2.5705
12	1.5135	37	1.8625	62	2.1375	87	2.5705
13	1.5135	38	1.8625	63	2.1375	88	2.6305
14	1.5485	39	1.9055	64	2.1375	89	2.6305
15	1.5485	40	1.9055	65	2.1375	90	2.6305
16	1.5485	41	1.9055	66	2.1375	91	2.6915
17	1.6845	42	1.9495	67	2.1875	92	2.6915
18	1.5845	43	1.9495	68	2.1875	93	2.7545
19	1.6215	44	1.9495	69	2.1875	94	2.8185
20	1.6215	45	1.9495	70	2.2385	95	2.8835
21	1.6215	46	1.9955	71	2.2385	96	2.9515
22	1.6595	47	1.9955	72	2.2385	97	3.0195
23	1.6595	48	1.9955	73	2.2905	98	3.0905
24	1.6985	49	1.9955	74	2.2905	99	3.3785
25	1.6985	50	1.9955	75	2.2905	100	3.7155

[a]Transformed from Snedecor's data (Table 1).

TABLE 4

TWENTY-FOUR HOUR EXCRETION OF URINE NA (MEQ/LITER) FROM DIURETIC EXPERIMENT[a]

After Placebo		After Mercuhydrin	
Subject Number	Result	Subject Number	Result
1	217	1	251
2	144	2	355
3	148	3	311
4	178	4	228
5	141	5	229
6	161	6	256
7	131	7	292
8	163	8	296
9	139	9	252
10	160	10	212
11	187	11	242
12	206	12	224
13	163	13	332
14	158	14	141
15	142	15	333
16	200	16	280
17	192	17	223
18	171	18	293
		19	248
		20	219

[a]See Reference 30.

TABLE 5

TWENTY-FOUR HOUR EXCRETION OF URINE K (MEQ/LITER) FROM DIURETIC EXPERIMENT[a]

After Placebo		After Diamox	
Subject Number	Result	Subject Number	Result
1	55	1	249
2	76	2	100
3	72	3	143
4	45	4	106
5	51	5	180
6	51	6	128
7	86	7	154
8	105	8	196
9	67	9	161
10	85	10	173
11	68	11	98
12	86	12	188
13	88	13	143
14	77	14	87
15	59	15	150
16	82	16	78
17	98	17	92
18	71	18	178
19	84	19	149
20	81	20	158
21	94	21	117
		22	135

[a]See Reference 30.

TABLE 6

CHANGE IN GASTRIC FREE HCl FROM RESERPINE EXPERIMENT[a]

After Administration of: Placebo				Reserpine			
Subject Number	Before	After	Change	Subject Number	Before	After	Change
1	40	36	- 4	1	15	60	+45
2	52	56	+ 4	2	17	55	+38
3	55	58	+	3	13	39	+26
4	73	70	- 3	4	15	45	+30
5	34	42	+ 8	5	11	79	+68
6	17	11	- 6	6	42	72	+30
7	48	42	- 6	7	42	76	+34
8	51	57	+ 6	8	50	103	+53
9	51	39	-12	9	50	69	+19
10	60	66	+ 6	10	25	67	+42
11	8	11	+ 3	11	35	91	+56
12	2	16	+14	12	52	54	+ 2
13	5	0	- 5	13	27	39	+12
14	10	39	+29	14	68	92	+24
15	25	26	+ 1	15	18	70	+52
16	42	42	0	16	18	60	+42
17	42	53	+11	17	50	41	- 9
18	7	28	+21				
19	10	34	+24				
20	45	49	+ 4				
21	45	51	+ 6				

[a]Those subjects whose before (i.e., fasting) value was zero are omitted. From previously published data (16).

TABLE 7

SERUM CHOLESTEROL DATA[a]

Ethiopians		Boys		Adult Men		Cardiacs	
Number	Result	Number	Result	Number	Result	Number	Result
1	227	1	187	1	255	1	258
2	103	2	196	2	189	2	338
3	163	3	345	3	272	3	249
4	134	4	145	4	308	4	298
5	93	5	199	5	248	5	307
6	153	6	157	6	279	6	193
7	129	7	179	7	230	7	236
8	163	8	126	8	238	8	248
9	227	9	203	9	267	9	275
10	122	10	185	10	169	10	355
11	155	11	171	11	222	11	214
		12	137	12	206	12	279
		13	122	13	254	13	179
				14	247	14	428
				15	183	15	424
				16	228	16	389
				17	230		
				18	187		
				19	340		
				20	160		
				21	246		
				22	245		
				23	195		
				24	193		

[a]See Reference 23.

TABLE 8

CHANGE IN FASTING VENOUS BLOOD GLUCOSE IN MG. PER CENT THREE HOURS AFTER ORAL MEDICATION

Subject Number	Placebo	1.0 Gram Chlorpropamide
1	- 4	-22
2	- 1	-22
3	+ 1	- 2
4	-12	-23
5	- 2	-26
6	- 1	-17
7	- 2	-20
8	-14	- 9
9	- 9	-13

[a]See Reference 24.

TABLE 9

TRIALS USING THE RESULTS OF THE CHLORPROPAMIDE-BLOOD GLUCOSE DATA[a]

I Subject Number	II Fasting Blood Glucose mg%	III Blood Glucose 3 hours after 1.0 gram of Chlorpropamide mg%	IV X = (III-II) mg%	V First Trial ΣX mg%	VI Second Trial ΣX mg%	VII Third Trial ΣX mg%
1	72	50	-22	1 = - 22	-0-	-0-
2	77	55	-22	2 = - 44	-0-	-0-
3	73	71	- 2	3 = - 46	1 = - 2	-0-
4	82	59	-23	4 = - 69	2 = - 25	-0-
5	85	59	-26	5 = - 95	3 = - 51	-0-
6	73	56	-17	6 = -112	4 = - 68	1 = -17
7	72	52	-20	7 = -132	5 = - 88	2 = -37
8	69	60	- 9	8 = -141	6 = - 97	3 = -46
9	75	62	-13	9 = -154	7 = -110	4 = -59
10	82	59	-23	10 = -177	8 = -133	5 = -82

[a]See Reference 24.

TABLE 10

SUMMARY OF F TESTS FOR HOMOGENEITY OF S^2_{POOLED} IN CONTROL AND TREATMENT GROUPS

Source of Samples	$S^2_{p_{control}}$	$S^2_{p_{treatment}}$	df	F	p
Normal Distribution	12.25	140.75	(4,4)	11.49	b
Rectangular Distribution	13.25	2.75	(4,4)	4.82	a
Logarithmic Distribution	0.135	0.138	(6,6)	1.02	a
Cholesterol Data	2222.68	1265.12	(6,6)	1.76	a
Urine Na Data	699.79	2166.58	(6,6)	3.10	a
Urine K Data	183.57	846.15	(6,6)	4.61	b
Gastric Juice Free HCl Data	45.03	212.85	(8,8)	4.73	b
Blood Glucose Data	58.90	29.45	(6,6)	2.00	a

a = $p > .05$ (not significant).
b = $p < .05$ and $> .01$ (borderline significance).
c = $p < .01$ (significant). Significant heterogeneity did not occur.

TABLE 11

RESULTS OF TESTING FOR APPROPRIATE SAMPLE SIZE (n)
ON DATA FROM TABLES 1, 2 AND 3

		$S_{\bar{x}}$ from Range	$S_{\bar{x}}$ from Two Means	$S_{\bar{x}}$ from $S^2_{x_p}$
Normal Distribution:				
n = 2	First pair	1.327	1.060	1.060
		(Approximate 95% Confidence Interval = 0 to 2.120)		
	Second pair	4.867	3.889	3.260
		(Approximate 95% Confidence Interval = 0 to 6.520)		
	Third pair	3.097	2.500	10.397
		(Approximate 95% Confidence Interval = 0 to 20.794)		
	Fourth pair	1.327	0.935	6.134
		(Approximate 95% Confidence Interval = 0 to 12.268)		
∴ accept means (S.S. n = 2) to be approximately normally distributed.				
Rectangular Distribution:				
n = 2	First pair	0.442	0.354	2.318
		(Approximate 95% Confidence Interval = 0 to 4.636)		
	Second pair	1.327	1.606	2.806
		(Approximate 95% Confidence Interval = 0 to 5.612)		
	Third pair	0.442	0.354	1.225
		(Approximate 95% Confidence Interval = 0 to 2.450)		
	Fourth pair	0.442	.125	1.060
		(Approximate 95% Confidence Interval = 0 to 2.12)		
∴ accept means (S.S. n = 2) to be approximately normally distributed.				
Logarithmic Distribution:				
n = 2	First pair	0.813	0.2055	0.0160
		(Approximate 95% Confidence Interval = 0 to 0.32)		
n = 3	First pair	0.088	0.0704	0.1993
		(Approximate 95% Confidence Interval = 0.0365 to 0.3621)		
	Second pair	0.132	0.1057	0.2534
		(Approximate 95% Confidence Interval = 0.0474 to 0.4594)		
∴ accept means (S.S. n = 3) to be approximately normally distributed.				

TABLE 12

ONE ANALYSIS OF THE CHOLESTEROL DATA WITH REGARD TO ADULT MEN

n =	X	ΣX	UL	LL
1	255	255	451.7	28.3
2	189	444	691.7	268.3
3	272	716	931.7	508.3
4	308	1024	1171.7	748.3
5	248	1272	1411.7	988.3
6	279	1551	1651.7	1228.3
7	230	1781	1891.7	1468.3
8	238	2019	2131.7	1708.3
9	267	2286	2371.7	1948.3
10	269	2555	2611.7	2188.3
11	222	2761	2851.7	2428.3
12	206	2967	3091.7	2668.3
13	254	3221	3331.7	2908.3
14	247	3468	3571.7	3148.3
15	183	3651	3811.7	3388.3
16	228	3879	4051.7	3628.3
17	230	4109	4291.7	3868.3
18	187	4296	4531.7	4108.3
19	340	4636	4771.7	4348.3
20	160	4796	5011.7	4588.3
21	246	5042	5251.7	4828.3
22	245	5287	5491.7	5068.3
23	195	5482	5731.7	5308.3
24	193	5675	5971.7	5548.3

TABLE 13

RESULTS AND SEQUENTIAL ANALYSIS OF THE CHOLESTEROL DATA REPORTED BY DROLETTE AND HEGSTED, IGNORING THE MULTIPLE COMPARISONS PROBLEM

n =	Ethiopians	Boys	Men	Cardiacs	m_1 and m_0		m_0 and m_2	
					UL	LL	UL	LL
1	227	187	255	258	676	-171.0	651.0	-196.0
2	330	383	444	596	928.5	81.5	878.5	31.5
3	393	728	716	845	1181.0	334.0	1106.0	259.0
4	527	873	1024	1143	1433.5	586.5	1333.5	486.5
5	620	1072	1272	1450	1686.0	839.0	.	714.0
6		1229	1551	1643	1938.5	1091.5	.	941.5
7		1408	1781	1879	2191.0	1344.0	.	1169.0
8		1534	2019	2127	2443.5	1596.5	.	1396.5
9		1737	2282	2402	2696.0	1849.0	.	1624.0
10		1922	2555	2757	2948.5	2101.5	.	1851.5
11		2093	2761	2971	3201.0	2354.0	.	2079.0
12		2230	2967	3250	3453.5	.	.	2306.5
13			3221	3429	3706.0	.	.	2534.0
14			3468	3857	3958.5	.	.	2761.5
15			3651	4281	4211.0	.	.	2989.0
16			3879		4463.0	.	.	3216.5
17			4109		4716.0	.	.	3444.0
18			4296		4968.5	.	.	3671.5
19			4636		5221.0	.	.	3899.0
20			4796		5473.5	.	.	4126.5
21			5042		5726.0	.	.	4354.0
22			5287		5978.5	.	.	4581.5
23			5482		6231.0	.	.	4807.0
24			5675		6483.5	5636.5	5883.5	5036.5
25					.	.	.	.
26					.	.	.	.
27					.	.	.	.
28					.	.	.	.
29					.	.	.	.
30					.	.	.	.
31					.	.	.	.
32					.	.	.	.
33					.	.	.	.
34					.	.	.	.
35					9008.5	8161.5	8158.5	7311.5

TABLE 14

RESULTS OF TESTING FOR APPROPRIATE SAMPLE SIZE (n)
ON DATA FROM TABLES 4, 5, 6 AND 8

		$S_{\bar{x}}$ from Range	$S_{\bar{x}}$ from Two Means	$S_{\bar{x}}$ from $S^2_{x_p}$
Table 4:				
n = 4	First pair	11.07	16.12	13.15
	(Approximate 95% Confidence Interval on $S_{\bar{x}}$: 3.85 to 22.45)			
	Second pair	8.79	12.79	28.50
	(Approximate 95% Confidence Interval on $S_{\bar{x}}$: 8.36 to 48.64)			
∴ accept means (S.S. n = 4) to be approximately normally distributed.				
Table 5:				
n = 3	First pair	9.11	10.18	5.62
	(Approximate 95% Confidence Interval on $S_{\bar{x}}$: 1.04 to 10.28)			
	Second pair	15.38	18.38	34.94
	(Approximate 95% Confidence Interval on $S_{\bar{x}}$: 6.42 to 63.46)			
∴ accept means (S.S. n = 3) to be approximately normally distributed.				
Table 6:				
n = 5	First pair	1.80	2.12	2.58
	(Approximate 95% Confidence Interval on $S_{\bar{x}}$: 0.98 to 4.18)			
	Second pair	2.48	4.09	6.62
	(Approximate 95% Confidence Interval on $S_{\bar{x}}$: 0.76 to 12.48)			
∴ accept means (S.S. n = 5) to be approximately normally distributed.				
Table 8:				
n = 3	First pair	1.01	1.20	2.69
	(Approximate 95% Confidence Interval on $S_{\bar{x}}$: 0.49 to 4.89			
	Second pair	3.96	4.74	5.07
	(Approximate 95% Confidence Interval on $S_{\bar{x}}$: 0.93 to 9.21)			
∴ accept means (S.S. n = 3) to be approximately normally distributed.				

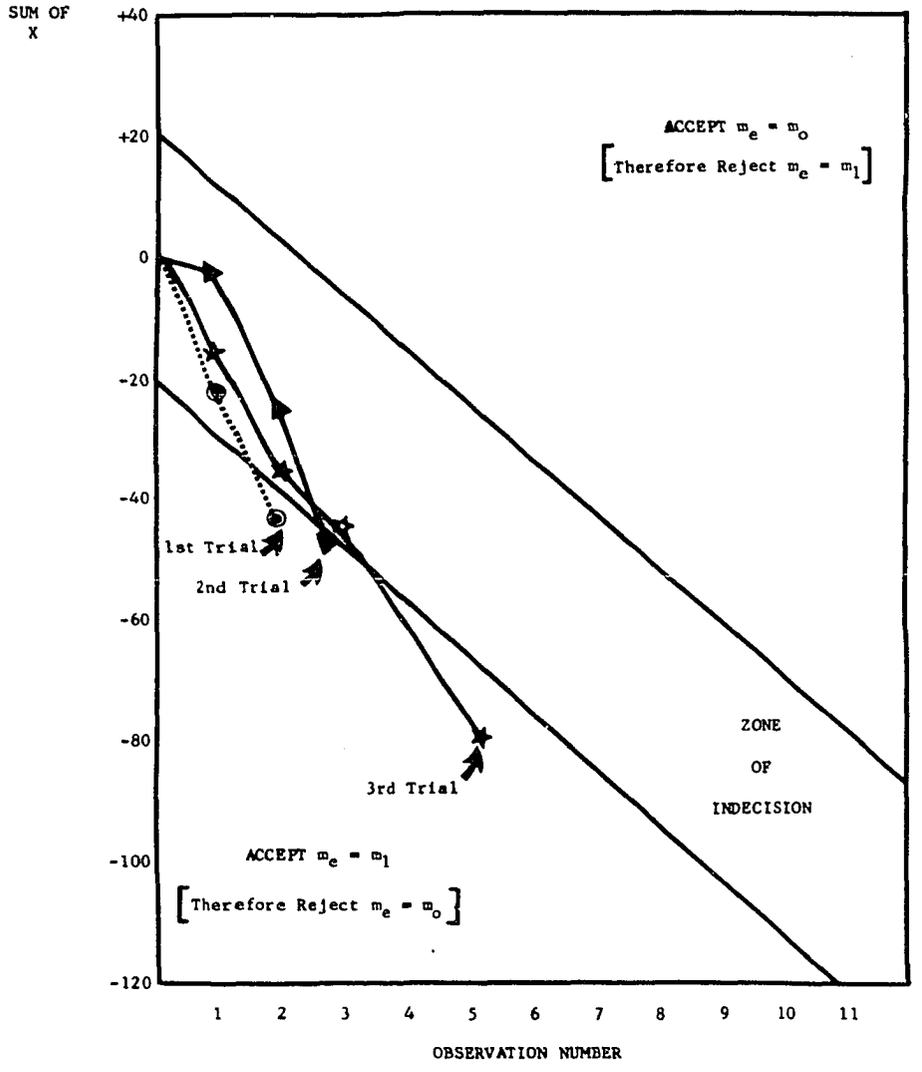

Figure 1. The sequential graph presenting the results of the three consecutive trials with chlorpropamide. The abscissa represents n (the number of observations) and the ordinate represents the sum of X (where X = blood glucose three hours after chlorpropamide minus the fasting blood glucose in mg% for each subject tested).