INVERSE GAUSSIAN REGRESSION MODELS

By

MAMMO WOLDIE

Bachelor of Science
Haile Selassie I University
Addis Ababa, Ethiopia
1970

Master of Science
Western Michigan University
Kalamazoo, Michigan
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
July, 1984

INVERSE GAUSSIAN REGRESSION MODELS

Thesis Approved:

J. Gerry Folks
Thesis Adviser

Ronald W. McNew

W. H. Stewart

Wayll E. Ray

Norman N. Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

I am very grateful for the help, advice and encouragement of many friends and relatives during my study. I would appreciate it if they recognize how difficult a task it is for me to list their names. However, there are some who deserve special mention.

I wish to express my sincere appreciation to Dr. J. Leroy Folks for serving as my major adviser, for suggesting the problem and for his valuable guidance.

I also thank Dr. Ronald W. McNew, Dr. William H. Stewart and Dr. Daryll E. Ray for their willingness to serve in my committee.

Finally, I dedicate this thesis to my mother for her unforgettable love and care.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Lack of development of different statistical methodology in other distributions has increased peoples' dependence on the normal distribution. One distribution that could serve as an alternative in some areas is the inverse Gaussian distribution.

It was Schrödinger and Smoluchowski, both in 1915, who obtained the probability distribution of the first passage time with positive drift. A special case of this distribution is also given by Wald (1947). Tweedie (1945, 1956) derived many of the important statistical properties of this distribution. He also pointed out that the cumulant generating function (cgf) of the first passage time distribution and that of the normal are inversely related.

Two random variables X and Y with cgf $L_X(t)$ and $L_Y(t)$ are inversely related if for all t values which belong to the domain of both cgf then $L_X(t) = aL(t)$ and $L_Y(t) = bL^{-1}(t)$ where a and b are constants and $L(L^{-1}(t)) = t$. It is Tweedie who proposed the name Inverse Gaussian (IG) for the first passage time distribution. There are also some analogies between the two distributions.

A stochastic process $W(t)$, $t \geq 0$, is said to be a Wiener process if

i) $W(t)$ has independent increments; i.e. for any $t_0<t_1<t_2<t_3$ then $W(t_1) - W(t_0)$ and $W(t_3) - W(t_2)$ are independent.

ii) $W(t_1) - W(t_0) \sim N(\gamma(t_1 - t_0), \sigma^2(t_1 - t_0))$, where $\gamma > 0$. Suppose

$W(0) = 0$ and T is the time required for $W(t)$ to reach the value a for the first time. Then the pdf of T is

$$f_T(t) = \frac{a}{\sigma\sqrt{2\pi t^3}} \exp\{-\frac{(a-\gamma t)^2}{2\sigma^2 t}\}, \qquad t > 0, \tag{1.1}$$

$\gamma > 0$ and $\sigma > 0$. If we let $\mu = \frac{a}{\gamma}$ and $\lambda = \frac{a^2}{\sigma^2}$ in 1.1, we get the standard expression of the pdf of the IG distribution, denoted by $I(\mu,\lambda)$. Thus

$$f_T(t) = \sqrt{\frac{\lambda}{2\pi t^3}} \exp\{-\frac{\lambda(t-\mu)^2}{2\mu^2 t}\}, \qquad t > 0,\ \mu > 0 \text{ and } \lambda > 0. \tag{1.2}$$

Some of the applications include determining the amount of time a particle of the injected substance remains in the blood which is summarized in Folks and Chhikara (unpublished monograph), determining the amount of time in emptying a dam until the release stops for the first time (Hasofer, 1964), describing the demand of frequently purchased low cost consumer products (Banerjee and Bhattacharyya, 1976), describing wind speed data (Bardsley, 1980).

This thesis consists of seven other chapters. Chapter II is a review of the basic properties of the inverse Gaussian distribution in general and those of the regression models in particular. Chapter III presents in detail new results for the zero intercept linear regression models with fixed λ and different λ while Chapter IV contains similar and asymptotic results for the nonzero intercept linear regression models. The general nonlinear regression model and its special case are discussed in Chapter V. Minimal sufficient statistics are presented in Chapter VI. The trials of these models on real and simulated data are given in Chapter VII. Then an outline of the thesis is presented in Chapter VIII.

As much as possible the following notation is followed in this thesis. Suppose θ is an unknown parameter. Then $\tilde{\theta}$ stands for the maximum likelihood estimator of θ, θ^* for the root of the likelihood equation and $\hat{\theta}$ for other estimators such as ordinary and modified least squares estimators.

CHAPTER II

REVIEW OF THE BASIC PROPERTIES OF THE INVERSE GAUSSIAN DISTRIBUTION

A. Basic Facts

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from $I(\mu,\lambda)$. The pdf of X_i is

$$f_{X_i}(x_i;\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi x_i^3}} \exp\left\{\frac{-\lambda(x_i-\mu)^2}{2\mu^2 x_i}\right\}, \quad x_i > 0, \ \mu > 0 \text{ and } \lambda > 0.$$

Tweedie (1957) gave the expression for the rth moment as

$$E(X^r) = \mu^r \sum_{s=0}^{r-1} \frac{(r-1+s)!\mu^s}{r!(r-1-s)!(2\lambda)^s}.$$

Thus, $E(X) = \mu$ and $Var(X) = \mu^3/\lambda$. He showed that the relationship between the positive and negative moments is

$$E(X^{-k}) = E(X^{k+1})/\mu^{2k+1}.$$

Hence, $E(\frac{1}{X}) = \frac{1}{\mu} + \frac{1}{\lambda}$ and $Var(\frac{1}{X}) = \frac{1}{\lambda}(\frac{1}{\mu} + \frac{2}{\lambda})$. He also obtained the characteristic function of X to be

$$\exp\{\frac{\lambda}{\mu}\{1 - (1 - (\frac{2i\mu^2 t}{\lambda})^{1/2}\}\}.$$

Chhikara and Folks (1974) showed that the pdf of

$$Y = \frac{\sqrt{\lambda}\,(X-\mu)}{\mu\sqrt{X}}, \qquad X > 0, \text{ is}$$

$$g(y;\frac{\lambda}{\mu}) = (1 - \frac{y}{\sqrt{y^2+4\lambda/\mu}}) \frac{1}{\sqrt{2\pi}} \exp(\frac{-y^2}{2}), \quad y \in R. \qquad 2.1$$

For a fixed μ, if $\lambda \to \infty$ then the pdf of Y approaches the standard normal distribution. They also overcome the need for a separate table to compute probabilities using IG distribution by expressing the cumulative distribution function of Y, where the pdf of Y is given in 2.1, in terms of the cumulative standard normal distribution, ϕ, by

$$G(y) = \phi(y) + \exp(\frac{2\lambda}{\mu})\phi(-\sqrt{y^2 + (4\lambda/\mu)}), \quad y \in R.$$

Zigangirov (1962) and Shuster (1968) independently obtained the same result. If we let $W = Y^2$ then W is a chi-square with one d.f. (Shuster, 1968). Moreover,

1. if X is IG with parameters μ and λ then for $t > 0$, tX is IG with parameters $t\mu$ and $t\lambda$.

2. a necessary and sufficient condition for $\sum_{i=1}^{n} X_i$ to be IG, where $X_1, X_2, \ldots, X_n$ are independent IG variables with parameters μ_i and λ_i, is that $\lambda_i/\mu_i^2 = k$ for all i. If this is so, then $\sum_{i=1}^{n} X_i$ is IG with parameters $\sum_{i=1}^{n} \mu_i$ and $k(\sum_{i=1}^{n} \mu_i)^2$.

B. Sampling Distribution

The likelihood function is

$$L(\mu,\lambda;\underset{\sim}{x}) = (\frac{\lambda}{2\pi})^{n/2} \prod_{i=1}^{n} x_i^{-3/2} \exp\{-\frac{\lambda}{2} \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{x_i\mu^2}\}.$$

Grödinger (1915) derived that the MLE of μ and λ are $\bar{X}$ and $\tilde{\lambda}$ where $\tilde{\lambda} = n/\sum_{i=1}^{n}(1/X_i - 1/\bar{X})$, respectively. Tweedie (1957) showed that (1) $\bar{X}$ is IG with parameters μ and $n\lambda$, (2) $n\lambda/\tilde{\lambda}$ is a chi-square with n-1 d.f. where $\tilde{\lambda} = n/\sum_{i=1}^{n}(1/X_i - 1/\bar{X})$, (3) $\bar{X}$ and $\tilde{\lambda}$ are independent and (4) $(\bar{X},$

$\sum_{i=1}^{n}(1/X_i - 1/\bar{X}))$ is a complete sufficient statistic for (μ,λ). Note however that the completeness property of the IG is shown by Wasan (1968).

C. Regression Models

Tweedie (1957) was the first to pave the way for the development of regression analysis by introducing what is known as the "Analysis of Reciprocals" (AOR). Davis (1977), continuing along this line, discussed the following three models:

1. $Y_i \sim I(\beta x_i,\lambda)$,
2. $Y_i \sim I(\beta x_i,\lambda_i)$ where $\frac{\beta^2 x_i^2}{\lambda_i} = k$ for all i,
3. $Y_i \sim I(\alpha+\beta x_i,\lambda)$.

Her results are summarized in paragraphs 1a, 1b and 2.

1. Simple Linear Regression Model - Zero Intercept

   a. Common λ

   The model is $Y_i = \beta x_i + e_i$, $E(e_i) = 0$, where $Y_i \sim I(\beta x_i,\lambda)$ $i = 1,\ldots,n$, $\beta > 0$, $x_i > 0$, $\lambda > 0$ and the Y's are independent.

   i. The MLE are

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} \frac{Y_i}{x_i^2}}{\sum_{i=1}^{n} \frac{1}{x_i}} \quad \text{and} \quad \tilde{\lambda} = \frac{n}{\sum_{i=1}^{n} (\frac{1}{Y_i} - \frac{1}{\tilde{\beta} x_i})}$$

   ii. $\tilde{\beta}$ is IG with parameters β and $\lambda \sum_{i=1}^{n} \frac{1}{x_i}$. Thus $Var(\tilde{\beta}) = \frac{\beta^3}{\lambda \Sigma \frac{1}{x}}$.

iii. $\dfrac{\lambda(\Sigma\frac{1}{x})(\tilde{\beta}-\beta)^2}{\beta^2\tilde{\beta}}$ and $\dfrac{n\lambda}{\tilde{\lambda}}$ are both chi-square with one and n-1 d.f., respectively.

iv. $\tilde{\beta}$ and $\tilde{\lambda}$ are independent.

v. $(\tilde{\beta},\tilde{\lambda})$ is a complete sufficient statistic for (β,λ).

vi. $\tilde{\beta}$ and $\dfrac{(n-3)\tilde{\lambda}}{n}$ are UMVUE of β and λ, respectively.

vii. To test H_0: $\beta=\beta_0$ against H_a: $\beta\neq\beta_0$ we can use the following statistic

$$\frac{(n-1)\tilde{\lambda}(\Sigma\frac{1}{x})(\tilde{\beta}-\beta_0)^2}{n\beta_0^2\tilde{\beta}} \qquad 2.2$$

which has an F distribution with one and n-1 d.f. Thus we will reject H_0 if the given statistic exceeds $F_{1,n-1,1-\alpha}$.

One can also construct confidence regions on β based on the statistic given in 2.2. Thus a $100(1-\alpha)\%$ confidence interval on β is (L,U) where

$$L = \frac{\tilde{\beta}}{1+\sqrt{\dfrac{F_{1,n-1,1-\alpha}(\tilde{\beta}\sum_{i=1}^{n}\frac{1}{Y_i}-\sum_{i=1}^{n}\frac{1}{x_i})}{(n-1)\sum_{i=1}^{n}\frac{1}{x_i}}}}$$

$$U = \frac{\tilde{\beta}}{1-\sqrt{\dfrac{F_{1,n-1,1-\alpha}(\tilde{\beta}\sum_{i=1}^{n}\frac{1}{Y_i}-\sum_{i=1}^{n}\frac{1}{x_i})}{(n-1)\sum_{i=1}^{n}\frac{1}{x_i}}}}$$

provided that the expression

$$(n-1)\sum_{i=1}^{n}\frac{1}{x_i} - F_{1,n-1,1-\alpha}\left(\tilde{\beta}\sum_{i=1}^{n}\frac{1}{Y_i} - \sum_{i=1}^{n}\frac{1}{x_i}\right)$$

is positive. If this expression is negative then a $100(1-\alpha)\%$ interval estimate of β is (L,∞).

b. Different λ

The model is $Y_i=\beta x_i+e_i$, $E(e_i)=0$, where $Y_i \sim I(\beta x_i,\lambda_i)$ such that $\frac{\beta^2 x_i^2}{\lambda_i} = k$ for all $i=1,\ldots,n$; $x_i > 0$, $\lambda_i > 0$ and the Y's are independent.

i. The MLE are $\tilde{\beta} = \frac{\bar{Y}}{\bar{x}}$

$$\tilde{k} = \frac{\bar{Y}^2}{n\bar{x}^2 \sum_{i=1}^{n}\left(\frac{x_i^2}{Y_i} - \frac{\bar{x}^2}{\bar{Y}}\right)}$$

$$\tilde{\lambda}_i = \frac{nx_i^2}{\sum_{i=1}^{n}\left(\frac{x_i^2}{Y_i} - \frac{\bar{x}}{\bar{Y}}\right)}$$

ii. $\tilde{\beta}$ is IG with parameters β and $\frac{n\beta^2\bar{x}}{k}$.

iii. $\frac{n\bar{x}(\tilde{\beta}-\beta)^2}{k\tilde{\beta}}$ and $\frac{n\lambda_i}{\tilde{\lambda}_i}$ are independent chi-squares, with one and n-1 d.f., respectively.

iv. $(\tilde{\beta},\tilde{\lambda}_1,\ldots,\tilde{\lambda}_n)$ is a complete sufficient statistic for $(\beta,\lambda_1,\ldots,\lambda_n)$.

v. $\tilde{\beta}$ and $\frac{(n-3)}{n}\tilde{\lambda}_i$ are UMVUE for β and λ_i, respectively.

vi. When testing $\beta=\beta_0$ against $\beta\neq\beta_0$ one can use the follow-

ing statistic

$$\frac{(n-1)n\bar{x}(\tilde{\beta}-\beta_0)^2}{\beta_0^2\tilde{\beta}\sum_{i=1}^{n}\left(\frac{x_i^2}{Y_i}-\frac{\bar{x}^2}{\bar{Y}}\right)}$$

which is an F with one and n-1 d.f. Hence one should reject H_0 if the given statistic is greater than $F_{1,n-1,1-\alpha}$.

Based on the given statistic a $100(1-\alpha)\%$ CI on β is (L,U) where

$$L = \frac{\tilde{\beta}}{1+\sqrt{\dfrac{F_{1,n-1,1-\alpha}(\tilde{\beta}\sum_i \frac{x_i^2}{Y_i}-n\bar{x})}{(n-1)n\bar{x}}}}$$

and

$$U = \frac{\tilde{\beta}}{1-\sqrt{\dfrac{F_{1,n-1,1-\alpha}(\tilde{\beta}\sum_i \frac{x_i^2}{Y_i}-n\bar{x})}{(n-1)n\bar{x}}}}$$

provided that $(n-1)n\bar{x} - F_{1,n-1,1-\alpha}(\tilde{\beta}\sum_i \frac{x_i^2}{Y_i} - n\bar{x}) > 0$. If however, $(n-1)n\bar{x} - F_{1,n-1,1-\alpha}(\tilde{\beta}\sum_i \frac{x_i^2}{Y_i} - n\bar{x}) < 0$ then the interval of β is (L,∞).

2. Simple Linear Regression Model with Intercept

The model is $Y_i = \alpha + \beta x_i + e_i$, $E(e_i) = 0$ where $Y_i \sim I(\alpha + \beta x_i, \lambda)$, $i = 1,\ldots,n$; $\alpha + \beta x_i > 0$, $\lambda > 0$ and the Y's are independent. Davis (1977) did not find closed expressions for the MLE of α and β. However, she did find unbiased estimators. Some of

these unbiased estimators of β are

$$\frac{\sum_{i=1}^{n} \frac{Y_i - \bar{Y}}{x_i^2}}{\sum_{i=1}^{n} \frac{x_i - \bar{x}}{x_i^2}} \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^{n} \left(\frac{Y_i - \bar{Y}}{x_i - \bar{x}}\right)$$

with $\bar{Y} - \hat{\beta}\bar{x}$ being an unbiased estimator of α.

3. Simple Nonlinear Regression Model

Folks and Chhikara in an unpublished monograph assumed the model $\frac{1}{Y_i} = \alpha + \beta x_i + e_{\lambda i}$, where $E(e_{\lambda i}) = \frac{1}{\lambda}$,

$$\mathrm{Var}(e_{\lambda i}) = \frac{\alpha + \beta x_i}{\lambda} + \frac{2}{\lambda^2} \quad \text{and} \quad Y_i \sim I\left(\frac{1}{\alpha + \beta x_i}, \lambda\right)$$

$i = 1,\ldots,n$; $\alpha + \beta x_i > 0$, $x_i > 0$, $\lambda > 0$ and the Y's are independent. They obtained that the roots of the likelihood are

$$\beta^* = \frac{n \sum_{i=1}^{n} (x_i - \bar{x})(Y_i - \bar{Y})}{\left(\sum_{i=1}^{n} x_i Y_i\right)^2 - \left(\sum_{i=1}^{n} Y_i\right)\left(\sum_{i=1}^{n} x_i^2 Y_i\right)}$$

$$\alpha^* = \frac{1}{\bar{Y}} - \beta^* \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} Y_i}$$

$$\text{and} \quad \frac{1}{\lambda^*} = \frac{1}{n} \sum_{i=1}^{n} \frac{((\alpha^* + \beta^* x_i) Y_i - 1)^2}{Y_i} .$$

They also showed that $(\alpha^*, \beta^*, \lambda^*)$ is a complete sufficient statistic for (α, β, λ). However it will be shown later that these

are not in general MLE.

4. Classification Model

Fries and Bhattacharyya (1983) assumed the following model

$$Y_{ijk} \sim I(\theta_{ij},\lambda) \qquad \begin{array}{l} i = 1,\ldots,I \\ j = 1,\ldots,J \\ k = 1,\ldots,N \end{array}$$

where the Y_{ijk}'s are independent and $\frac{1}{\theta_{ij}} = \mu + \alpha_i + \beta_j$. They obtained the following results.

a. The root of the likelihood equation is $\underset{\sim}{\phi}^* = IJM^{-1}\underset{\sim}{e}$, where

$$\underset{\sim}{\phi}' = (\mu,\alpha_1,\ldots,\alpha_{I-1},\beta_1,\ldots,\beta_{J-1})$$

$$M = X'DX$$

$$D = \mathrm{diag}(\bar{y}_{11},\bar{y}_{12},\ldots,\bar{y}_{IJ})$$

$X' = (\underset{\sim}{x}_{11},\underset{\sim}{x}_{12},\ldots,\underset{\sim}{x}_{IJ})$, each $\underset{\sim}{x}_{ij}$ consisting of -1s,

0s and 1s such that $\mu + \alpha_i + \beta_j = \underset{\sim}{\phi}' \underset{\sim}{x}_{ij}$

$$\underset{\sim}{e}' = (1,0,\ldots,0).$$

They showed that $\sqrt{n}(\underset{\sim}{\phi}^* - \underset{\sim}{\phi})$ is asymptotically $N_{I+J-1}(0, \frac{1}{\lambda}(X'\theta X)^{-1})$ and $\sqrt{n}(\frac{1}{\lambda^*} - \frac{1}{\lambda})$ is also asymptotically $N(0,\frac{2}{\lambda^2 IJ})$, where $\theta = \mathrm{diag}(\theta_{11},\theta_{12},\ldots,\theta_{IJ})$. In addition $\underset{\sim}{\phi}^*$ and λ^* are asymptotically independent. However, it will be shown later that $\underset{\sim}{\phi}^*$ is not in general MLE.

b. Modified Least Squares Estimator

Although a closed expression exists for the root of the likelihood it is not possible to find its expected value

or its variance. However using a modified version of the least squares approach they obtained unbiased estimators and their variances.

The model is $\underset{\sim}{s} = X\underset{\sim}{\eta} + \underset{\sim}{\varepsilon}$, $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$ where $\underset{\sim}{s} = (s_{11}, s_{12}, \ldots, s_{IJ})'$, s_{ij} denoting $\frac{1}{\bar{y}_{ij}}$ and $\underset{\sim}{\eta} = (\mu + \frac{1}{n\lambda}, \alpha_1, \ldots, \alpha_{I-1}, \beta_1, \ldots, \beta_{J-1})'$. Thus

$$\hat{\mu} = \bar{s}_{..} - \frac{1}{n\hat{\lambda}}$$

$$\hat{\alpha}_i = \bar{s}_{i.} - \bar{s}_{..}$$

$$\hat{\beta}_j = \bar{s}_{.j} - \bar{s}_{..}, \qquad \text{where}$$

$$\frac{1}{\hat{\lambda}} = \frac{1}{IJ(n-1)} \sum_{ijk} \left(\frac{1}{y_{ijk}} - \frac{1}{\bar{y}_{ij}}\right).$$

They proved that $\sqrt{n}(\hat{\underset{\sim}{\eta}} - \underset{\sim}{\eta})$ is asymptotically $N_{I+J-1}(\underset{\sim}{0}, \frac{1}{\lambda}G'\theta^{-1}G)$ and $\sqrt{n}(\frac{1}{\hat{\lambda}} - \frac{1}{\lambda})$ is also asymptotically $N(0, \frac{2}{\lambda^2 IJ})$ where $G = X(X'X)^{-1}$. Moreover, $\hat{\underset{\sim}{\phi}}$ and $\hat{\lambda}$ are asymptotically independent.

CHAPTER III

SIMPLE REGRESSION MODELS WITH ZERO INTERCEPT

A. Common λ

$Y_i \sim I(\beta x_i, \lambda)$, $i = 1,\ldots,n$; $\beta > 0$, $x_i > 0$, $\lambda > 0$ and the Y's are indedent.

1. a. Kósei Iwase and Noriaki Setó (1983) showed that the UMVU Estimator of μ^3/λ where $X \sim IG(\mu,\lambda)$ is

$(\bar{x}^3 v)/(n-1)F(1,1.5;(n+1)/2;-\bar{x}v/n)$ where $v = \sum_{i=1}^{n}(1/x_i - 1/\bar{x})$,

and $F(1,1.5;\frac{n+1}{2};-\frac{\bar{x}v}{n})$ is

$$\frac{2\Gamma(\frac{n+1}{2})}{\sqrt{\pi}\,\Gamma(\frac{n-2}{2})} \cdot \int_0^1 \frac{\sqrt{t}(1-t)^{\frac{n-4}{2}}}{(1+t\frac{\bar{x}v}{n})}\,dt.$$

The problem of obtaining a value for F can be overcome using the following recursion formula 15.2.2(1).

$$c(c-1)(z-1)F(a,b;c-1;z) + c(c-1-(2c-a-b-1)z)F(a,b;c-1;z) +$$

$$(c-a)(c-b)zF(a,b;c+1;z) = 0.$$

If we let $a = 1$, $b = 3/2$, $c = (n-1)/2$ and $z = -\bar{x}v/n$, then for $n \geq 5$

$$F(1,3/2;(n+1)/2;z) = 1/[(n-3)(n-4)z][(n-1)(n-3)(1-z)$$

$F(1,3/2;(n-3)/2;z) - (n-1)((n-3) - (2n-9)z)F(1,3/2;(n-1)/2,z))]$.

However, we need to know the values of F for $n=0$, 1, 2, and 3. For $n=0$ use $(z+1)/(z-1)^2$, $n=1$ use $(1-z)^{-3/2}$, $n=2$ use $(1-z)^{-1}$ and $n=3$ use $[1/(1+\sqrt{1-z})\sqrt{1-z}]$. Thus using $n=0$ and 2 we get the values for $n=4$. Using $n=2$ and 4 we can obtain for $n=6$ etc. We can similarly obtain for odd values of n.

b. Based on the above result, the UMVU Estimator of $Var(Y_i) = \beta^3 x_i^3/\lambda$ is

$$\tilde{\beta}^3 x_i^3 \frac{n}{\tilde{\lambda}(n-1)} \qquad F(1,1.5;(n+1)/2;-\tilde{\beta}x_i/\tilde{\lambda})$$

where $\tilde{\beta}$ and $\tilde{\lambda}$ are MLEs.

2. Power

For λ known to test the hypothesis H_0: $\beta = \beta_0$ against the alternative hypothesis H_a: $\beta \neq \beta_0$ we use the statistic

$$W = \frac{\lambda(\Sigma\frac{1}{x})(\tilde{\beta} - \beta_0)^2}{\beta_0^2\tilde{\beta}} .$$

Thus we reject H_0 if $W > \chi^2_{1,\alpha}$. To determine the power of the test we need to find the distribution of W.

a. Let $z = \sqrt{\lambda(\Sigma 1/x)}(\tilde{\beta}-\beta_0)/\beta_0\sqrt{\tilde{\beta}}$. Our aim is to find the pdf of z. Thus,

$$\sqrt{\tilde{\beta}} = \frac{\beta_0 z + \sqrt{\beta_0^2 z^2 + 4\lambda(\Sigma 1/x)\beta_0}}{2\sqrt{\lambda\Sigma 1/x}}$$

and

$$\frac{d\tilde{\beta}}{dz} = \frac{\beta_0^2[z^2 + \sqrt{z^2 + 4\lambda(\Sigma 1/x)/\beta_0}]^2}{2\lambda(\Sigma 1/x)\sqrt{z^2 + 4\lambda(\Sigma 1/x)/\beta_0}}$$

Since $\tilde{\beta}$ is IG with parameters β and $\lambda\Sigma\frac{1}{x}$ then the pdf of z is

$$g_Z(z) = \frac{1}{\sqrt{2\pi}}\left(1 - \frac{z}{\sqrt{z^2 + 4\lambda(\Sigma 1/x)/\beta_0}}\right)$$

$$\exp[-\left(1 + \frac{\beta_0^2}{\beta^2}\right)\frac{z^2}{4} - \frac{\lambda(\Sigma 1/x)}{2\beta_0}\left(1 - \frac{\beta_0}{\beta}\right)^2 + \left(1 - \frac{\beta_0}{\beta^2}\right)\frac{z\sqrt{z^2 + 4\lambda(\Sigma 1/x)/\beta_0}}{4}] ,$$

$z \in R$.

b. Let $W = z^2 = \dfrac{\lambda(\Sigma_i \frac{1}{x_i})(\tilde{\beta} - \beta_0)^2}{\beta_0^2\tilde{\beta}}$. Our aim is to find the pdf of W. Hence

$$f_W(w) = \frac{1}{2\sqrt{w}}[g_Z(\sqrt{w}) + g_Z(-\sqrt{w})]$$

$$= \frac{1}{2\sqrt{2\pi w}} \exp[-\left(1 + \frac{\beta_0^2}{\beta^2}\right)\frac{w}{4} - \frac{\lambda(\Sigma_i \frac{1}{x_i})}{2\beta_0}\left(1 - \frac{\beta_0}{\beta}\right)^2]$$

$$\{\left(1 - \frac{w}{c(w,\frac{\lambda}{\beta_0})}\right) \exp[\left(1 - \frac{\beta_0^2}{\beta^2}\right)\frac{c(w,\frac{\lambda}{\beta_0})}{4}] +$$

$$\left(1 + \frac{w}{c(w,\frac{\lambda}{\beta_0})}\right) \exp[\left(\frac{\beta_0^2}{\beta^2} - 1\right)\frac{c(w,\frac{\lambda}{\beta_0})}{4}]\} \qquad 3.1$$

where $w > 0$ and $c(w,\frac{\lambda}{\beta_0}) = \sqrt{w^2 + \frac{4\lambda(\sum_i \frac{1}{x_i})w}{\beta_0}}$. This pdf is quite similar to that of a noncentral chi-square. When $\beta = \beta_0$ the distribution of W is a chi-square with one d.f.

Let p denote the power of the test. Thus the power function

$$p(\beta) = \text{Prob}(W > \chi^2_{1,\alpha}) = \int_{\chi^2_{1,\alpha}}^{\infty} f_\beta(w)\, dw.$$

For λ unknown, the statistic used for testing $\beta = \beta_0$ against $\beta \neq \beta_0$ is

$$F = \frac{(n-1)\tilde{\lambda}(\Sigma\frac{1}{x})(\tilde{\beta} - \beta_0)^2}{\tilde{\beta}\beta_0^2 n}$$

We will reject H_0 if $F > F_{1,n-1,1-\alpha}$. To evaluate the power of the test we will first find the distribution of F. Let

$$F = \frac{(n-1)W}{Y}$$

and $U = Y$ where pdf of W is given in 3.1 and U is a χ^2_{n-1}. Thus

$$W = \frac{FU}{n-1}$$

and

$$Y = U.$$

The Jacobian of the transformation is $\frac{U}{n-1}$. Suppose we denote

the joint pdf of (W,Y) by $f_{W,Y}(w,y)$. Thus the joint pdf of (F,U)

$$g_{F,U}(f,u) = f_{W,Y}(\frac{fu}{n-1},u)\frac{u}{n-1}, \qquad f > 0,\ u > 0.$$

Integrating out u we will get the marginal of F,

$$g_F(f) = \frac{1}{\Gamma(\frac{n-1}{2})2^{\frac{n+2}{2}}\sqrt{(n-1)\pi f}} \exp[-\frac{\lambda(\Sigma\frac{1}{x})}{2\beta_0}(1-\frac{\beta_0}{\beta})^2]\int_0^\infty u^{\frac{n}{2}-1}$$

$$\{(1-\frac{u}{d(u,\frac{\lambda}{\beta_0})}) \exp[-[(1+\frac{\beta_0^2}{\beta^2})\frac{f}{2(n-1)}+1]\frac{u}{2} + (1-\frac{\beta_0^2}{\beta^2})$$

$$\frac{fd(u,\frac{\lambda}{\beta_0})}{4}] + (1+\frac{u}{d(u,\frac{\lambda}{\beta_0})}) \exp[-[(1+\frac{\beta_0^2}{\beta^2})\frac{f}{2(n-1)}+1]\frac{u}{2}$$

$$+ (1-\frac{\beta_0^2}{\beta^2})\frac{fd(u,\frac{\lambda}{\beta_0})}{4}]\}du$$

where

$$d(u,\frac{\lambda}{\beta_0}) = \sqrt{u^2+\frac{4\lambda(\Sigma\frac{1}{x})(n-1)u}{\beta_0 f}}$$

Let $z = \dfrac{u((\beta^2+\beta_0^2)f+2\beta^2(n-1))}{4\beta^2(n-1)}$

$$\frac{du}{dz} = [\frac{dz}{du}]^{-1} = \frac{4\beta^2(n-1)}{((\beta^2+\beta_0^2)f+2\beta^2(n-1))}$$

$$g_F(f) = \frac{\exp[-\frac{\lambda(\Sigma\frac{1}{x})}{2\beta_0}(1-\frac{\beta_0}{\beta})^2]}{2\Gamma(\frac{n-1}{2})\sqrt{(n-1)\pi f}((1+\frac{\beta_0^2}{\beta^2})\frac{f}{2}+n-1)^{\frac{n}{2}}}$$

$$\int_0^\infty z^{\frac{n}{2}-1} \exp(-z)\{(1-\frac{z}{d(z,\beta)}) \exp[(1-\frac{\beta_0^2}{\beta^2})\frac{fd(z,\beta)}{(1+\frac{\beta_0^2}{\beta^2})f+2(n-1)}]$$

$$+ (1+\frac{z}{d(z,\beta)}) \exp[(\frac{\beta_0^2}{\beta^2}-1)\frac{fd(z,\beta)}{(1+\frac{\beta_0^2}{\beta^2})f+2(n-1)}]\}dz$$

where $f > 0$ and

$$d(z,\beta) = \sqrt{z^2 + \frac{\lambda(\Sigma\frac{1}{x})}{\beta_0 f}((1+\frac{\beta_0^2}{\beta^2})+2(n-1))z} .$$

Although it is not possible to find a closed form for the pdf of F, it is quite analogous to a noncentral F.

The power of the test is

$$p_\beta(\beta) = \text{Prob}(F > F_{1,n-2,1-\alpha}) = \int_{F_{1,n-2,1-\alpha}}^\infty g_F(f)\ df.$$

3. Confidence Interval for $E(Y_i)$

We want to obtain an interval estimate for the mean of Y_i.

Case (a) λ known: We know that

$$\frac{\lambda(\Sigma\frac{1}{x})(\tilde{\beta}-\beta)^2}{\tilde{\beta}\beta^2}$$

is a chi-square with one d.f. Solving the inequality

$$\frac{\lambda(\Sigma\frac{1}{x})(\tilde{\beta}-\beta)^2}{\tilde{\beta}\beta^2} \leq \chi^2_{1,1-\alpha}$$

for β we get an interval estimate of β depending upon whether the coefficient of β^2 is positive or negative. Using this interval estimate of β a $100(1-\alpha)\%$ CI on βx_i is (L,U), where

$$L = \tilde{\beta}x_i\left(\frac{1}{1-\frac{\chi_1^2\tilde{\beta}}{\lambda\Sigma\frac{1}{x}}} - \frac{\sqrt{\lambda(\Sigma\frac{1}{x})\chi_1^2\tilde{\beta}}}{\lambda\Sigma\frac{1}{x}-\chi_1^2\tilde{\beta}}\right)$$

$$U = \tilde{\beta}x_i\left(\frac{1}{1-\frac{\chi_1^2\tilde{\beta}}{\lambda\Sigma\frac{1}{x}}} + \frac{\sqrt{\lambda(\Sigma\frac{1}{x})\chi_1^2\tilde{\beta}}}{\lambda\Sigma\frac{1}{x}-\chi_1^2\tilde{\beta}}\right), \qquad \chi_1^2 = \chi^2_{1,1-\alpha},$$

provided that the coefficient of β^2 is positive. If the coefficient of β^2 is negative then the interval is (L,∞). If L is negative then the interval estimate of βx_i is $(0,U)$ and $(0,\infty)$ provided that the coefficient of β^2 is positive and negative, respectively.

Case (b) λ unknown: Based on the results given by Davis (1977) the following CI on βx_i is constructed. Solving the inequality

$$\frac{(n-1)\tilde{\lambda}(\Sigma\frac{1}{x})(\tilde{\beta}-\beta)^2}{\tilde{\beta}\beta^2 n} \leq F_{1,n-1,1-\alpha}$$

one can obtain an interval estimate of β depending upon whether the coefficient of β^2 is positive or negative. Based on this interval estimate of β, a $100(1-\alpha)\%$ CI on βx_i is (L,U), where

$$L = \frac{\tilde{\beta}x_i}{1+\sqrt{\dfrac{F(\tilde{\beta}\sum_i \frac{1}{y_i} - \sum_i \frac{1}{x_i})}{(n-1)\sum_i \frac{1}{x_i}}}}, \qquad U = \frac{\tilde{\beta}x_i}{1-\sqrt{\dfrac{F(\tilde{\beta}\sum_i \frac{1}{y_i} - \sum \frac{1}{x_i})}{(n-1)\sum_i \frac{1}{x_i}}}},$$

$F = F_{1,n-1,1-\alpha}$, provided that the coefficient of β^2 is positive. If the coefficient of β^2 is negative then the interval is (L,∞).

4. Prediction Interval

Suppose in addition to the n independent Y's we have a future observation Y from an IG distribution with parameters βx_* and λ. The exponent term of their joint distribution

$$\lambda \sum_{i=1}^{n} \frac{(y_i - \beta x_i)^2}{\beta^2 x_i^2 y_i} + \frac{\lambda(y-\beta x_*)^2}{\beta^2 x_*^2 y}$$

which is a chi-square with $n+1$ d.f. can be decomposed into

$$\lambda \sum_{i=1}^{n} (\frac{1}{y_i} - \frac{1}{\tilde{y}_i}) + \frac{\lambda(\Sigma\frac{1}{x})(\tilde{\beta}-\beta)^2}{\beta^2\tilde{\beta}} + \frac{\lambda(y-\beta x_*)^2}{\beta^2 x_*^2 y}$$

with $\tilde{y}_i$ being an estimate of $E(Y_i)$, which are independent chi-squares with d.f. n-1, one and one, respectively. By combining the last two terms we get

$$\lambda \sum_{i=1}^{n} \left(\frac{1}{y_i} - \frac{1}{\tilde{y}_i}\right) + \frac{\lambda \sum_{i=1}^{n} \frac{1}{x_i}(\tilde{\beta}x_* - y)^2}{\tilde{\beta}y(y + (\sum_{i=1}^{n} \frac{1}{x_i})x_*^2\tilde{\beta})} +$$

$$\frac{\lambda(y + (\Sigma\frac{1}{x})x_*^2\tilde{\beta} - \beta x_*(x_*(\Sigma\frac{1}{x}) + 1))^2}{\beta^2 x_*^2(y + (\Sigma\frac{1}{x})x_*^2\tilde{\beta})} .$$

We want to determine the distribution of

$$\frac{\lambda(\Sigma\frac{1}{x})(\tilde{\beta}x_* - y)^2}{\tilde{\beta}y(y + (\Sigma\frac{1}{x})x_*^2\tilde{\beta})}.$$

We know that

$$\tilde{\beta} \sim I(\beta, \lambda\Sigma \frac{1}{x_i}) \quad , \quad Y \sim I(\beta x_*, \lambda)$$

and they are assumed to be independent. Thus

$$\frac{\tilde{\beta}}{\Sigma\frac{1}{x_i}} \sim I(\frac{\beta}{\Sigma\frac{1}{x_i}}, \lambda).$$

Let $W = (\Sigma\frac{1}{x})x_*^2\tilde{\beta}$ which is IG with parameters $\beta(\Sigma\frac{1}{x})x_*^2$ and $\lambda(\Sigma\frac{1}{x})^2x_*^2$. The joint distribution of (W,Y) is

$$f_{W,Y}(w,y) = \frac{\lambda(\Sigma\frac{1}{x})x_*}{2\pi w^{3/2}\ y^{3/2}} \exp\{-\frac{\lambda}{2}[\frac{(y - \beta x_*)^2}{\beta^2 x_*^2 y} + \frac{(w - \beta(\Sigma\frac{1}{x})x_*^2)^2}{\beta^2(\Sigma\frac{1}{x})^2 x_*^2 w}]\}.$$

Let $V = Y + W$ which is IG with parameters

$$\beta x_*(1 + (\Sigma\tfrac{1}{x})x_*) \qquad \text{and} \qquad \lambda(1 + (\Sigma\tfrac{1}{x})x_*)^2.$$

Let's now obtain the conditional distribution of Y given $V = y + w$

$$f_{Y|V}(y|v = y + w) = \frac{f_{Y,W}(y,w)}{g_V(v)}$$

where the denominator is the pdf of $Y + W$.

$$= \sqrt{\frac{\lambda(y+w)^3}{2\pi(1 + (\Sigma\frac{1}{x})x_*)^2 y^3 w^3}} \quad \exp\{-\frac{\lambda}{2}[\frac{1}{y} + \frac{(\Sigma\frac{1}{x})^2 x_*^2}{w} - \frac{(1 + (\Sigma\frac{1}{x})x_*)^2}{y+w}]\}$$

$0 < y < v = y + w$. To find the conditional distribution of

$$\lambda(\frac{1}{y} + \frac{(\Sigma\frac{1}{x})^2 x_*^2}{w} - \frac{(1 + (\Sigma\frac{1}{x})x_*)^2}{y+w})$$

given $V = y + w$ we will use

$$E[\exp(\lambda t(\frac{1}{y} + \frac{(\Sigma\frac{1}{x})x_*^2}{w} - \frac{(1 + (\Sigma\frac{1}{x})x_*)^2}{y+w}))\,|\,V = y + w]$$

which is equal to

$$\int_0^{y+w} \exp[\lambda t(\frac{1}{y} + \frac{(\Sigma\frac{1}{x})^2 x_*^2}{w} - \frac{(1 + (\Sigma\frac{1}{x})x_*)^2}{y+w})]$$

$$\sqrt{\frac{\lambda(y+w)^3}{2\pi(1 + \Sigma\frac{1}{x})^2 y^3 w^3}} \exp\{-\frac{\lambda}{2}(\frac{1}{y} + \frac{(\Sigma\frac{1}{x})^2 x_*^2}{w} - \frac{(1 + (\Sigma\frac{1}{x})x_*)^2}{y+w})\}dy$$

$$= (1 - 2t)^{-\frac{1}{2}} \int_0^{y+w} f_{Y|V}(y|V = y + w)dy = (1 - 2t)^{-\frac{1}{2}}.$$

Since this is the same for all values we have conditioned on then

$$\frac{\lambda(\Sigma\frac{1}{x})(\tilde{\beta}x_* - y)^2}{\tilde{\beta}y(y + \tilde{\beta}x_*^2(\Sigma\frac{1}{x}))}$$

is a chi-square with one d.f. Let us obtain an interval estimate of β.

Case (a) λ known: Solving the inequality

$$\frac{\lambda(\Sigma\frac{1}{x})(\tilde{\beta}x_* - y)^2}{\tilde{\beta}y(y + \tilde{\beta}x_*^2(\Sigma\frac{1}{x}))} \leq \chi^2_{1,1-\alpha} \quad \text{for } y,$$

a $100(1-\alpha)\%$ PI on Y is (L,U) where

$$L = \frac{\tilde{\beta}x_*}{2}\left(-(\Sigma\frac{1}{x})x_* + \frac{\lambda(\Sigma\frac{1}{x})(2+(\Sigma\frac{1}{x})x_*) - \sqrt{(\Sigma\frac{1}{x})X_1^2\tilde{\beta}[(\Sigma\frac{1}{x})x_*(\tilde{\beta}x_*X_1^2+4\lambda)+4\lambda]}}{\lambda\Sigma\frac{1}{x} - X_1^2\tilde{\beta}}\right)$$

$$U = \frac{\tilde{\beta}x_*}{2}\left(-(\Sigma\frac{1}{x})x_* + \frac{\lambda(\Sigma\frac{1}{x})(2+(\Sigma\frac{1}{x})x_*) + \sqrt{(\Sigma\frac{1}{x})X_1^2\tilde{\beta}[(\Sigma\frac{1}{x})x_*(\tilde{\beta}x_*X_1^2+4\lambda)+4\lambda]}}{\lambda\Sigma\frac{1}{x} - X_1^2\tilde{\beta}}\right)$$

where $X_1^2 = \chi^2_{1,1-\alpha}$, provided that the parabola opens upward. If the parabola opens downward then the interval is (L,∞).

Case (b) λ unknown: The ratio of

$$\frac{\lambda(\sum_i \frac{1}{x_i})(\tilde{\beta}x_* - Y)^2}{\tilde{\beta}Y(Y + \tilde{\beta}(\sum_i \frac{1}{x_i})x_*^2)} \quad \text{with} \quad D = \frac{\lambda}{n-1}\sum_{i=1}^{n}\left(\frac{1}{Y_i} - \frac{1}{\tilde{Y}_i}\right)$$

is an F with one and n-1 d.f. Solving the inequality

$$\frac{\left(\sum_{i=1}^{n} \frac{1}{x_i}\right)(\tilde{\beta}x_* - y)^2}{\tilde{\beta}Dy\left(y + \tilde{\beta}\left(\Sigma\frac{1}{x}\right)x_*^2\right)} \leq F_{1,n-1,1-\alpha} \quad \text{for } y,$$

if the coefficient of y^2 is positive, then a 100(1-α)% PI on Y is (L,U) where

$$L = \frac{\tilde{\beta}x_*}{2}\left(-x_*\left(\sum_i \frac{1}{x_i}\right) + \frac{\left(\sum_i \frac{1}{x_i}\right)\left[2 + x_*\left(\sum_i \frac{1}{x_i}\right)\right] - \sqrt{\tilde{\beta}DF\left(\sum_i \frac{1}{x_i}\right)\left[x_*\left(\sum_i \frac{1}{x_i}\right)(\tilde{\beta}DFx_* + 4) + 4\right]}}{\sum_i \frac{1}{x_i} - \tilde{\beta}DF}\right)$$

$$U = \frac{\tilde{\beta}x_*}{2}\left(-x_*\left(\sum_i \frac{1}{x_i}\right) + \frac{\left(\sum_i \frac{1}{x_i}\right)\left[2 + x_*\left(\sum_i \frac{1}{x_i}\right)\right] + \sqrt{\tilde{\beta}DF\left(\sum_i \frac{1}{x_i}\right)\left[x_*\left(\sum_i \frac{1}{x_i}\right)(\tilde{\beta}DFx_* + 4) + 4\right]}}{\sum_i \frac{1}{x_i} - \tilde{\beta}DF}\right),$$

$F = F_{1,n-1,1-\alpha}$. If the coefficient of y^2 is negative then the interval is (L,∞).

Case (c) β known: Solving the inequality

$$\frac{n(y - \beta x_*)^2}{Dx_*^2 y} \leq F_{1,n,1-\alpha} \quad \text{where } D = \sum_{i=1}^{n} \frac{(y_i - \beta x_i)^2}{x_i^2 y_i},$$

for y, a 100(1-α)% PI on Y is

$$x_*\left(\beta + \frac{DFx_* - \sqrt{DFx_*(DFx_* + 4\beta n)}}{2n},\ \beta + \frac{DFx_* + \sqrt{DFx_*(DFx_* + 4\beta n)}}{2n}\right),$$

$F = F_{1,n,1-\alpha}$.

5. Confidence Interval for x_0

Suppose in addition to the n independent Y's, t independent Y's are observed at unknown x, say x_0, i.e., $Y_i \sim I(\beta x_0, \lambda)$, $i = n+1, \ldots, n+t$; $\beta > 0$, $\lambda > 0$, x_0 unknown but positive. The MLE are

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} \frac{Y_i}{x_i^2}}{\sum_{i=1}^{n} \frac{1}{x_i}},$$

$$\tilde{x}_0 = \frac{\overline{Y_0}}{\tilde{\beta}}, \quad \text{and}$$

$$\frac{1}{\tilde{\lambda}} = \frac{1}{n+t}\left[\sum_{i=1}^{n} \left(\frac{1}{Y_i} - \frac{1}{\tilde{Y}_i}\right) + \sum_{i=n+1}^{n+t} \left(\frac{1}{Y_i} - \frac{1}{\overline{Y_0}}\right)\right].$$

Note that $\frac{(n+t)\lambda}{\tilde{\lambda}}$ is a chi-square with n+t-2 d.f. The exponent term of the joint distribution of the Y's, excluding $-\frac{\lambda}{2}$, is

$$\sum_{i=1}^{n} \frac{(Y_i - \beta x_i)^2}{\beta^2 x_i^2 Y_i} + \sum_{i=n+1}^{n+t} \frac{(Y_i - \beta x_0)^2}{\beta^2 x_0^2 Y_i} \qquad 3.2$$

$$= \sum_{i=1}^{n} \left(\frac{1}{Y_i} - \frac{1}{\tilde{Y}_i}\right) + \sum_{i=n+1}^{n+t} \left(\frac{1}{Y_i} - \frac{1}{\overline{Y_0}}\right) + \frac{(\Sigma \frac{1}{x_i})(\tilde{\beta} - \beta)^2}{\beta^2 \tilde{\beta}} + \frac{t(\overline{Y_0} - \beta x_0)^2}{\beta^2 x_0^2 \overline{Y_0}}. \qquad 3.3$$

λ times the expression in 3.2 is a chi-square with n+t d.f. and it is decomposed into a sum of independent chi-squares with n-1, t-1, one and one d.f., respectively. The sum of the last two

terms of 3.3 can be rewritten as

$$\frac{\lambda t(\Sigma\frac{1}{x})(\overline{Y}_0 - \tilde{\beta}x_0)^2}{\tilde{\beta}\overline{Y}_0(t\overline{Y}_0 + \tilde{\beta}x_0^2\Sigma\frac{1}{x})} + \frac{\lambda t^2(t\overline{Y}_0 + \tilde{\beta}x_0^2\Sigma\frac{1}{x} - \beta x_0(t + x_0\Sigma\frac{1}{x}))^2}{\beta^2 x_0^2(t\overline{Y}_0 + \tilde{\beta}x_0^2\Sigma\frac{1}{x})}$$

which are independent chi-squares each with one d.f.

Case (a) λ known: Let's determine the set of x_0's for which

$$\frac{\lambda t(\Sigma\frac{1}{x})(\overline{y}_0 - \tilde{\beta}x_0)^2}{\tilde{\beta}\overline{y}_0(t\overline{y}_0 + \tilde{\beta}x_0^2(\Sigma\frac{1}{x}))} \leq \chi^2_{1,1-\alpha}.$$

Thus a $100(1-\alpha)\%$ PI on x_0 is (L,U) where

$$L = \frac{\overline{y}_0}{\tilde{\beta}}\left(\frac{1}{1 - \frac{X_1^2\overline{y}_0}{\lambda t}} - \frac{\sqrt{t(\Sigma\frac{1}{x})X_1^2[\lambda\tilde{\beta}t + \lambda(\Sigma\frac{1}{x})\overline{y}_0 - \tilde{\beta}\overline{y}_0X_1^2]}}{(\Sigma\frac{1}{x})(\lambda t - \overline{y}_0X_1^2)}\right)$$

$$U = \frac{\overline{y}_0}{\tilde{\beta}}\left(\frac{1}{1 - \frac{X_1^2\overline{y}_0}{\lambda t}} + \frac{\sqrt{t(\Sigma\frac{1}{x})X_1^2[\lambda\tilde{\beta}t + \lambda(\Sigma\frac{1}{x})\overline{y}_0 - \tilde{\beta}\overline{y}_0X_1^2]}}{(\Sigma\frac{1}{x})(\lambda t - \overline{y}_0X_1^2)}\right),$$

$X_1^2 = \chi^2_{1,1-\alpha}$, provided that the parabola opens upward and that we have real roots. If the parabola opens downward with the roots of x_0 being real then the interval is (L,∞). If the lower bound is negative then the interval is $(0,U)$ or $(0,\infty)$ if the parabola opens upward or downward, respectively. However, no interval estimate of β exists if x_0 does not have real roots.

Case (b) λ unknown: Solving the inequality

$$\frac{t(\sum_i \frac{1}{x_i})(n+t-2)(\overline{y_0} - \tilde{\beta}x_0)^2}{\tilde{\beta}D\overline{y_0}(t\overline{y_0} + \tilde{\beta}x_0^2\sum_i \frac{1}{x_i})} \leq F_{1,n+t-2,1-\alpha},$$

where

$$D = \sum_{i=1}^{n} (\frac{1}{y_i} - \frac{1}{\tilde{y}_i}) + \sum_{i=n+1}^{n+t} (\frac{1}{y_i} - \frac{1}{\overline{y_0}}), \text{ for } x_0,$$

then a $100(1-\alpha)\%$ CI on x_0 is (L,U) where

$$L = \frac{\overline{y_0}}{\tilde{\beta}}\left(\frac{1}{1-\frac{DF\overline{y}_0}{t(n+t-2)}} - \frac{\sqrt{DFt(\sum_i \frac{1}{x_i})\overline{y_0}[\tilde{\beta}t(n+t-2)+(n+t-2)(\sum_i \frac{1}{x_i})\overline{y_0}-\tilde{\beta}DF\overline{y_0}]}}{(\sum_i \frac{1}{x_i})(t(n+t-2) - DF\overline{y_0})}\right)$$

$$U = \frac{\overline{y_0}}{\tilde{\beta}}\left(\frac{1}{1-\frac{DF\overline{y}_0}{t(n+t-2)}} + \frac{\sqrt{DFt(\sum_i \frac{1}{x_i})\overline{y_0}[\tilde{\beta}t(n+t-2)+(n+t-2)(\sum_i \frac{1}{x_i})\overline{y_0}-\tilde{\beta}DF\overline{y_0}]}}{(\sum_i \frac{1}{x_i})(t(n+t-2) - DF\overline{y_0})}\right),$$

$F = F_{1,n+t-2,1-\alpha}$, provided that the parabola opens upward and we have real roots. If the parabola opens downward and the roots of x_0 are real then the interval is (L,∞). If the lower bound is negative then the interval is $(0,U)$ or $(0,\infty)$ provided that the parabola opens upward or downward, respectively. However, an interval estimate of x_0 does not exist if x_0 does not have real roots.

Case (c) β known. We want to determine the set of x_0's for which

$$\frac{t(n+t-1)(\overline{y_0} - \beta x_0)^2}{\beta^2 D x_0^2 \overline{y_0}} \leq F_{1,n+t-1,1-\alpha}, \qquad 3.4$$

where D is the sum of the first three terms of 3.3. Since 3.4 is quadratic in x_0 then the parabola opens upward, a $100(1-\alpha)\%$ PI on x_0 is (L,U) where

$$L = \frac{\overline{y_0}}{\beta}\left(\frac{1}{1-\frac{DF\overline{y_0}}{t(n+t-2)}} - \frac{\sqrt{t(n+t-1)DF\overline{y_0}}}{t(n+t-1)-DF\overline{y_0}}\right)$$

$$U = \frac{\overline{y_0}}{\beta}\left(\frac{1}{1-\frac{DF\overline{y_0}}{t(n+t-2)}} + \frac{\sqrt{t(n+t-1)DF\overline{y_0}}}{t(n+t-1)-DF\overline{y_0}}\right),$$

$F = F_{1,n+t-1,1-\alpha}$. If the parabola opens downward then the interval is (L,∞).

6. We know that $E(\frac{1}{Y_i}) = \frac{1}{\beta x_i}+\frac{1}{\lambda}$. Let us denote the estimate of $E(\frac{1}{Y_i})$ by $(\frac{\tilde{1}}{y_i})$. Thus,

$$\sum_{i=1}^{n} \left(\frac{1}{y_i} - \left(\frac{\tilde{1}}{y_i}\right)\right) = \sum_{i=1}^{n} \left(\frac{1}{y_i} - \frac{1}{\tilde{\beta}x_i} - \frac{1}{\tilde{\lambda}}\right) = 0$$

where $\tilde{\beta}$ and $\tilde{\lambda}$ are MLE. This result is also true for the general nonlinear regression model except that we should use the roots of the likelihood equation.

B. Different λ

$Y_i \sim I(\beta x_i, \lambda_i)$, where $(\beta^2 x_i^2)/\lambda_i = k$ for all $i = 1,\ldots,n$; $\beta > 0$, $x_i > 0$, $\lambda_i > 0$ and the Y's are independent.

1. UMVU Estimator of $Var(Y_i)$

Davis (1977) showed that $(\tilde{\beta},\tilde{\lambda}_1,\ldots,\tilde{\lambda}_n)$ is a complete sufficient

statistic for $(\beta,\lambda_1,\ldots,\lambda_n)$ where $\tilde{\beta}$ and $\tilde{\lambda}_i$, $i=1,\ldots,n$ are MLE. Applying the result given by Kõsei Iwase and Noriaki Setõ (1983) the UMVUE of $\mathrm{Var}(Y_i) = (\beta^3 x_i^3)/\lambda_i = \beta k x_i$ is

$$\frac{n(\tilde{\beta}x_i)^3}{(n-1)\tilde{\lambda}_i} F(1,1.5;\frac{n+1}{2};\frac{-\tilde{\beta}x_i}{\tilde{\lambda}_i}),$$

where $F(1,1.5;\frac{n+1}{2};\frac{-\tilde{\beta}x_i}{\tilde{\lambda}_i}) = \frac{2\Gamma(\frac{n+1}{2})}{\sqrt{\pi}\Gamma(\frac{n-2}{2})} \int_0^1 \frac{\sqrt{t}(1-t)^{\frac{n-4}{2}}}{1+t\frac{\tilde{\beta}x_i}{\tilde{\lambda}_i}}\, dt$

2. Power

The test statistic used for testing $\beta=\beta_0$ against $\beta\neq\beta_0$, if k is known, is

$$T = \frac{n\bar{x}(\tilde{\beta}-\beta_0)^2}{\tilde{\beta}k}$$

which is a chi-square with one d.f. We reject H_0 if T is greater than $\chi^2_{1,1-\alpha}$.

To obtain the power function we will first find the distribution of T. We know that $\tilde{\beta}$ is IG with parameters β and $(n\beta^2\bar{x})/k$.

a. Let

$$u = \frac{\sqrt{n\bar{x}}(\tilde{\beta}-\beta_0)}{\sqrt{\tilde{\beta}k}} \qquad \text{(i)}$$

We will first find the pdf of u. Expressing $\tilde{\beta}$ in terms of u, we get

$$\tilde{\beta} = \frac{(\sqrt{k}u + \sqrt{ku^2 + 4\beta_0 n\bar{x}})^2}{4n\bar{x}}$$

with a Jacobian of

$$\frac{d\tilde{\beta}}{du} = \frac{\sqrt{k}(\sqrt{k}u + \sqrt{ku^2 + 4\beta_0 n\bar{x}})^2}{2n\bar{x}\sqrt{ku^2 + 4\beta_0 n\bar{x}}}$$

The term in the exponent of the p.d.f. of $\tilde{\beta}$, $(n\bar{x}(\tilde{\beta} - \beta)^2)/k\tilde{\beta}$, can be expressed as

$$(1 + \frac{\beta^2}{\beta_0^2})\frac{u^2}{2} + (1 - \frac{\beta^2}{\beta_0^2})\frac{u}{2}\sqrt{u^2 + \frac{4\beta_0 n\bar{x}}{k}} + \frac{(\beta - \beta_0)^2 n\bar{x}}{\beta_0 k}.$$

Hence the p.d.f. of u is

$$g_U(u) = \frac{\beta}{\beta_0\sqrt{2\pi}}\left(1 - \frac{u}{\sqrt{u^2 + \frac{4\beta_0 n\bar{x}}{k}}}\right)$$

$$\exp\{-(1 + \frac{\beta^2}{\beta_0^2})\frac{u^2}{4} + (\frac{\beta^2}{\beta_0^2} - 1)\frac{u}{4}\sqrt{u^2 + \frac{4\beta_0 n\bar{x}}{k}} - \frac{\beta_0 n\bar{x}}{2k}(1 - \frac{\beta}{\beta_0})^2\}, \qquad u \in R.$$

b. $T = U^2$

Suppose we denote the p.d.f. of T by $h_T(t)$. Hence

$$h_T(t) = \frac{1}{2\sqrt{t}}(g_U(\sqrt{t}) + g_U(-\sqrt{t}))$$

$$= \frac{\beta}{2\beta_0\sqrt{2\pi t}} \exp\{-(1 + \frac{\beta^2}{\beta_0^2})\frac{t}{4} - \frac{\beta_0 n\bar{x}}{2k}(1 - \frac{\beta}{\beta_0})^2\}$$

$$[(1-\frac{t}{\sqrt{t^2+\frac{4\beta_0 nt\bar{x}}{k}}})\exp\{(\frac{\beta^2}{\beta_0^2}-1)\frac{\sqrt{t^2+\frac{4\beta_0 nt\bar{x}}{k}}}{4}\}$$

$$+(1+\frac{t}{\sqrt{t^2+\frac{4\beta_0 nt\bar{x}}{k}}})\exp\{(1-\frac{\beta^2}{\beta_0^2})\frac{\sqrt{t^2+\frac{4\beta_0 nt\bar{x}}{k}}}{4}\}], \quad t>0. \qquad 3.5$$

Thus, the power function

$$p(\beta) = \text{Prob}[T > \chi^2_{1,1-\alpha}] = \int_{\chi^2_{1,1-\alpha}}^{\infty} h_\beta(t)\,dt.$$

If k is unknown the statistic used for testing $\beta = \beta_0$ against $\beta \neq \beta_0$ is

$$F = \frac{(n-1)n\bar{x}(\tilde{\beta}-\beta_0)^2}{\tilde{\beta}\beta_0^2\Sigma(\frac{x_i^2}{y_i}-\frac{\bar{x}}{\tilde{\beta}})}.$$

We reject H_0: $\beta = \beta_0$ if $F > F_{1,n-1,1-\alpha}$. To evaluate the power of the test we need to find the p.d.f. of F. Let

$$f = \frac{(n-1)t}{y}$$

$$v = y$$

where the p.d.f. of T is given in 3.5 and $Y \sim \chi^2_{n-1}$. Then

$$t = \frac{fv}{n-1}$$

$$y = v.$$

The joint p.d.f. of (T,Y) is

$$f_{T,Y}(t,y) = h_T(t)f_Y(y).$$

Hence the joint distribution of (F,V) is

$$g_{F,V}(f,v) = h_T(\frac{fv}{n-1})f_Y(v)\frac{v}{n-1}, \quad f > 0 \text{ and } v > 0.$$

The marginal of F,

$$g_F(f) = \frac{\beta(n-1)^{-\frac{1}{2}} f^{-\frac{1}{2}}}{2\beta_0\sqrt{2\pi}\Gamma(\frac{n-1}{2})2^{\frac{n-1}{2}}} \int_0^\infty v^{\frac{n}{2}-1}$$

$$\exp\{-\frac{(\beta_0^2+\beta^2)f+2\beta_0^2(n-1)}{4\beta_0^2(n-1)}v - \frac{n\bar{x}(\beta-\beta_0)^2}{2\beta_0 k}\}$$

$$[(1-\frac{v}{d(v,\frac{\beta_0}{k})}) \exp\{(\frac{\beta^2}{\beta_0^2}-1)\frac{fd(v,\frac{\beta_0}{k})}{4(n-1)}\} +$$

$$(1+\frac{v}{d(v,\frac{\beta_0}{k})}) \exp\{(1-\frac{\beta^2}{\beta_0^2})\frac{fd(v,\frac{\beta_0}{k})}{4(n-1)}\}] \quad \text{where}$$

$$d(v,\frac{\beta_0}{k}) = \sqrt{v^2 + \frac{4\beta_0 n(n-1)\bar{x}v}{kf}}.$$

Let $z = \dfrac{(\beta_0^2+\beta^2)f+2\beta_0^2(n-1)}{4\beta_0^2(n-1)}v.$

Thus $\dfrac{dv}{dz} = (\dfrac{dz}{dv})^{-1} = \dfrac{4\beta_0^2(n-1)}{(\beta_0^2+\beta^2)f+2\beta_0^2(n-1)}.$

Hence,

$$g_F(f) = \frac{\beta \exp\left(-\frac{n\bar{x}(\beta-\beta_0)^2}{2\beta_0 k}\right)}{2\beta_0\sqrt{\pi f(n-1)}\,\Gamma\left(\frac{n-1}{2}\right)\left(\left(1+\frac{\beta^2}{\beta_0^2}\right)\frac{f}{2(n-1)}+1\right)^{\frac{n}{2}}}$$

$$\int_0^\infty z^{\frac{n}{2}-1}\exp(-z)\left\{\left(1-\frac{z}{d(z,\beta)}\right)\exp\left[\left(\frac{\beta^2}{\beta_0^2}-1\right)\frac{fd(z,\beta)}{4(n-1)}\right] +\right.$$

$$\left.\left[\left(1+\frac{z}{d(z,\beta)}\right)\exp\left[\left(1-\frac{\beta^2}{\beta_0^2}\right)\frac{fd(z,\beta)}{4(n-1)}\right]\right\}dz, \text{ where}$$

$$d(z,\beta) = \sqrt{z^2+\frac{n\bar{x}((\beta_0^2+\beta^2)f+2\beta_0^2(n-1))z}{\beta_0 kf}}\ .$$

Although it is not possible to find a closed form for the p.d.f. of F, it is quite analogous to a non-central F. Note that when $\beta = \beta_0$ we get a central F with one and n-1 d.f.

The power function is

$$p(\beta) = \text{Prob}(F > F_{1,n-1,1-\alpha}) = \int_{F_{1,n-1,1-\alpha}}^\infty g_\beta(f)\,df\ .$$

3. Confidence Interval for $E(Y_i)$

Let us find an interval estimate of the mean of Y_i.

Case (a) k known: Using the set of β's for which

$$\frac{n\bar{x}(\tilde{\beta}-\beta)^2}{k\tilde{\beta}} \leq \chi^2_{1,1-\alpha},$$ a 100(1-α)% CI on βx_i is (L,U) where

$$L = \tilde{\beta}x_i - x_i\sqrt{\frac{\tilde{\beta}k\chi^2_{1,1-\alpha}}{n\bar{x}}} \qquad U = \tilde{\beta}x_i + x_i\sqrt{\frac{\tilde{\beta}k\chi^2_{1,1-\alpha}}{n\bar{x}}}$$

If L is negative then the interval is (0,U).

Case (b) k unknown: Davis (1977) obtained the set of β's for which

$$\frac{(n-1)\left(\sum_{i=1}^{n}\frac{1}{x_i}\right)(\tilde{\beta}-\beta)^2\sum_{i=1}^{n}\frac{y_i}{x_i^2}}{\tilde{\beta}\beta^2\left(\sum_{i=1}^{n}\frac{1}{y_i}\right)\left(\sum_{i=1}^{n}\frac{y_i}{x_i^2}\right)-\left(\sum_{i=1}^{n}\frac{1}{x_i}\right)^2} \leq F_{1,n-1,1-\alpha}.$$

Thus a 100(1-α)% confidence interval on βx_i is (L,U) where

$$L = \frac{\tilde{\beta}x_i}{1+\sqrt{\dfrac{F_{1,n-1,1-\alpha}\left(\tilde{\beta}\sum_{i=1}^{n}\frac{x_i^2}{y_i}-n\bar{x}\right)}{n\bar{x}}}}$$

$$U = \frac{\tilde{\beta}x_i}{1-\sqrt{\dfrac{F_{1,n-1,1-\alpha}\left(\tilde{\beta}\sum_{i=1}^{n}\frac{x_i^2}{y_i}-n\bar{x}\right)}{n\bar{x}}}}$$

provided that

$$(n-1)\sum_{i=1}^{n}\frac{1}{x_i} - F_{1,n-1,1-\alpha}(\tilde{\beta}\sum_{i=1}^{n}\frac{1}{y_i} - \sum_{i=1}^{n}\frac{1}{x_i}) \qquad 3.6$$

is positive. If 3.6 is negative then the interval is (L,∞).

4. Prediction Intervals

Suppose in addition to the n-independent Y's we have a future observation Y which is IG with parameters βx_* and λ_* where $(\beta^2 x_*^2)/\lambda_* = k$. The exponent term of the joint distribution of the Y's is

$$\frac{1}{k}\sum_{i=1}^{n}\frac{(Y_i - \beta x_i)^2}{Y_i} + \frac{(Y - \beta x_*)^2}{kY}$$

which is a chi-square with n+1 d.f. Rewriting the first term as the sum of two independent chi-squares we get

$$\frac{\beta^2}{k}\sum_{i=1}^{n}(\frac{x_i^2}{Y_i} - \frac{\bar{x}}{\tilde{\beta}}) + \frac{n\bar{x}(\tilde{\beta} - \beta)^2}{\tilde{\beta}k} + \frac{(Y - \beta x_*)^2}{kY} .$$

By combining the last two terms we get

$$\frac{\beta^2}{k}\sum_{i=1}^{n}(\frac{x_i^2}{Y_i} - \frac{\bar{x}}{\tilde{\beta}}) + \frac{n\beta^2\bar{x}(\tilde{\beta}x_* - Y)^2}{k\tilde{\beta}Y(\tilde{\beta}n\bar{x} + Y)} + \frac{(\tilde{\beta}n\bar{x} + Y - \beta(n\bar{x} + x_*))^2}{k(Y + \tilde{\beta}n\bar{x})}$$

which are independent chi-squares with d.f. n-1, one and one, respectively.

Case (a) β and k are unknown: Solving the inequality

$$\frac{(n-1)\bar{x}(y - \tilde{\beta}x_*)^2}{\tilde{\beta}vy(\tilde{\beta}n\bar{x} + y)} \leq F_{1,n-1,1-\alpha} \quad \text{where } v = \frac{1}{n}\sum_{i=1}^{n}(\frac{x_i^2}{y_i} - \frac{\bar{x}}{\tilde{\beta}})$$

for y, then a $100(1-\alpha)\%$ PI on Y is (L,U), where

$$L=\frac{\tilde{\beta}}{2}\left(\bar{x}\left(-n+\frac{(n-1)(2x_*+n\bar{x}))-\sqrt{\tilde{\beta}FV\bar{x}[4n(n-1)\bar{x}x_*+\tilde{\beta}Fn^2v\bar{x}+4(n-1)x_*^2]}}{(n-1)\bar{x}-\tilde{\beta}FV}\right)\right)$$

$$U=\frac{\tilde{\beta}}{2}\left(\bar{x}\left(-n+\frac{(n-1)(2x_*+n\bar{x}))+\sqrt{\tilde{\beta}FV\bar{x}[4n(n-1)\bar{x}x_*+\tilde{\beta}Fn^2v\bar{x}+4(n-1)x_*^2]}}{(n-1)\bar{x}-\tilde{\beta}FV}\right)\right),$$

$F = F_{1,n-1,1-\alpha}$, provided that the parabola opens upward. If the parabola opens downward then the interval is (L,∞).

Case (b) β known:

$$\frac{n(Y-\beta x_*)^2}{YD},\ D=\sum_{i=1}^{n}\frac{(Y_i-\beta x_i)^2}{Y_i} \qquad 3.7$$

is an F with one and n d.f. Thus solving the inequality

$$\frac{n(Y-\beta x_*)^2}{YD} \le F_{1,n,1-\alpha}$$ for Y, a $100(1-\alpha)\%$ PI on Y is

$$\left(\beta x_* + \frac{DF-\sqrt{DF(DF+4\beta n x_*)}}{2n},\ \beta x_* + \frac{DF+\sqrt{DF(DF+4\beta n x_*)}}{2n}\right),$$

$F = F_{1,n,1-\alpha}$.

5. Confidence Interval on x_0

Suppose in addition to the n independent Y's observed at known x's we observe t independent Y's at an unknown x, say x_0, i.e. $Y_i \sim I(\beta x_0,\lambda_i)$, $i=n+1,\ldots,n+t$ where $\beta^2x_0^2/\lambda_i = k$ for all i.

If we estimate β from the first n observations and βx_0 from the

last t observations, then $\tilde{\beta} = \bar{Y}/\bar{x}$ and $\tilde{\beta}x_0 = \bar{Y}_0$. Thus $\tilde{x}_0 = \bar{Y}_0/\tilde{\beta}$. The exponent term of the joint distribution of the Y's, excluding $-\frac{1}{2}$,

$$\frac{1}{k}\sum_{i=1}^{n}\frac{(Y_i - \beta x_i)^2}{Y_i} + \frac{1}{k}\sum_{i=n+1}^{n+t}\frac{(Y_i - \beta x_0)^2}{Y_i}$$

which is a chi-square with n+t d.f. can be decomposed into

$$\frac{\beta^2}{k}\sum_{i=1}^{n}\left(\frac{x_i^2}{Y_i} - \frac{\bar{x}}{\tilde{\beta}}\right) + \frac{\beta^2 x_0^2}{k}\sum_{i=n+1}^{n+t}\left(\frac{1}{Y_i} - \frac{1}{\bar{Y}_0}\right) + \frac{n\bar{x}(\tilde{\beta} - \beta)^2}{\tilde{\beta}k} + \frac{t(\bar{Y}_0 - \beta x_0)^2}{k\bar{Y}_0} . \quad 3.8$$

The last four terms are independent chi-squares with d.f. n-1, t-1, one and one, respectively. The sum of the last two terms of 3.8 can be rewritten as

$$\frac{\beta^2 nt\bar{x}(\bar{Y}_0 - \tilde{\beta}x_0)^2}{\tilde{\beta}k\bar{Y}_0(\tilde{\beta}n\bar{x} + t\bar{Y}_0)} + \frac{(n\bar{x} + tx_0)(\tilde{\beta}n\bar{x} + t\bar{Y}_0 - \beta(n\bar{x} + tx_0))^2}{kn\bar{x}(\tilde{\beta}n\bar{x} + t\bar{Y}_0)} .$$

The two terms are also independent chi-squares each with one d.f.

Case (a) β and k are unknown--Exact CI: Solving the inequality

$$\frac{(n-1)t\bar{x}(\bar{Y}_0 - \tilde{\beta}x_0)^2}{\tilde{\beta}V\bar{Y}_0(\tilde{\beta}n\bar{x} + t\bar{Y}_0)} \leq F_{1,n-1,1-\alpha} , \quad V = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i^2}{Y_i} - \frac{\bar{x}}{\tilde{\beta}}\right),$$

for x_0, then a $100(1-\alpha)\%$ CI on x_0 is

$$\left(\frac{\overline{y_0}}{\tilde{\beta}} - \sqrt{\frac{(\tilde{\beta}n\bar{x} + t\overline{y_0})Fv\overline{y_0}}{(n-1)\tilde{\beta}t\bar{x}}}\,,\ \frac{\overline{y_0}}{\tilde{\beta}} + \sqrt{\frac{(\tilde{\beta}n\bar{x} + t\overline{y_0})Fv\overline{y_0}}{(n-1)\tilde{\beta}t\bar{x}}}\right), \quad F = F_{1,n-1,1-\alpha}.$$

Approximate CI: Let us find the set of x_0's for which

$$\frac{t\bar{x}(n+t-2)(\overline{y_0} - \tilde{\beta}x_0)^2}{\tilde{\beta}D\overline{y_0}(\tilde{\beta}n\bar{x} + t\overline{y_0})} \leq F_{1,n+t-2,1-\alpha}$$

where D is the sum of the first two terms of 3.8 with x_0^2 replaced by its estimate $\bar{y}_0^2/\tilde{\beta}^2$. Thus an approximate 100(1-α)% CI on x_0 is

$$\left(\frac{\bar{y}_0}{\tilde{\beta}} - \sqrt{\frac{(\tilde{\beta}n\bar{x} + t\overline{y_0})DF\overline{y_0}}{(n+t-2)\tilde{\beta}t\bar{x}}}\,,\ \frac{\bar{y}_0}{\tilde{\beta}} + \sqrt{\frac{(\tilde{\beta}n\bar{x} + t\overline{y_0})DF\overline{y_0}}{(n+t-2)\tilde{\beta}t\bar{x}}}\right), \quad F = F_{1,n+t-2,1-\alpha}.$$

Case (b) β known--Exact CI: Based on the set of x_0's for which

$$\frac{nt(\overline{y_0} - \beta x_0)^2}{D\overline{y_0}} \leq F_{1,n,1-\alpha},$$

where D is given in 3.7, then a 100(1-α)% CI on x_0 is

$$\left(\frac{\overline{y_0}}{\beta} - \frac{1}{\beta}\sqrt{\frac{DF\overline{y_0}}{nt}}\,,\ \frac{\overline{y_0}}{\beta} + \frac{1}{\beta}\sqrt{\frac{DF\overline{y_0}}{nt}}\right), \quad F = F_{1,n,1-\alpha}.$$

Approximate CI: Replacing x_0^2 by its estimate $\bar{y}_0^2/\beta^2$ in the second term of 3.8 and obtaining the set of x_0's for which

$$\frac{(n+t-1)t(\overline{y_0} - \beta x_0)^2}{V\overline{y_0}} \leq F_{1,n+t-1,1-\alpha},$$

then an approximate $100(1-\alpha)\%$ CI on x_0 is

$$\left(\frac{\bar{y}_0}{\beta} - \frac{1}{\beta}\sqrt{\frac{FV\overline{y_0}}{t(n+t-1)}}\ ,\ \frac{\overline{y_0}}{\beta} + \frac{1}{\beta}\sqrt{\frac{FV\overline{y_0}}{t(n+t-1)}}\right), \quad F = F_{1,n+t-1,1-\alpha},$$

where $V = D + \bar{y}_0^2 \sum_{i=n+1}^{n+t} (1/y_i - 1/\overline{y_0})$.

CHAPTER IV

SIMPLE REGRESSION MODELS WITH NONZERO INTERCEPT

A. Common λ

Consider $Y_i \sim I(\alpha + \beta x_i, \lambda)$, $i = 1,\ldots,n$, $\alpha + \beta x_i > 0$, $\lambda > 0$ and the Y's are independent.

As mentioned in the introduction it is not possible to obtain a closed expression for the root of the likelihood equation. However, using the estimator due to Davis (1977)

$$\hat{\beta} = \frac{1}{n} \sum_{i=1}^{n} \frac{Y_i - \bar{Y}}{x_i - \bar{x}}, \qquad x_i \neq \bar{x} \text{ for all } i,$$

and estimating α by $\bar{Y} - \hat{\beta}\bar{x}$ we obtain the following results.

1. $\hat{\beta} \xrightarrow{P} \beta$ provided the x's are bounded, i.e. $a \leq x_i \leq b$ for all $i = 1,\ldots,n$ and

$$\sum_{i=1}^{n} \frac{1}{(x_i - \bar{x})^2} \leq \frac{cn}{b-a},$$

where c is some positive constant.

<u>Proof</u>:

$$P[|\hat{\beta} - \beta| \geq \varepsilon] \leq \frac{1}{\varepsilon^2} \operatorname{Var}(\hat{\beta})$$

$$= \frac{1}{\lambda\varepsilon^2 n^4} \sum_{i=1}^{n} (\alpha+\beta x_i)^3 \left(\sum_{k=1}^{n} \frac{1}{x_k - \bar{x}} - \frac{n}{x_i - \bar{x}} \right)^2$$

$$\leq \frac{(\alpha + \beta x\max)^3}{\lambda\varepsilon^2 n^4} \sum_{i=1}^{n} \left(\sum_{k=1}^{n} \frac{(x_i - x_k)}{(x_k - \bar{x})(x_i - \bar{x})} \right)^2, \quad x\max = x_{\text{maximum}}$$

$$\leq \frac{(\alpha+\beta x\max)^3}{\lambda\varepsilon^2 n^4} \sum_{i=1}^{n} [(\sum_{k=1}^{n} (x_i - x_k)^2)(\sum_{k=1}^{n} \frac{1}{(x_k-\bar{x})^2(x_i-\bar{x})^2})],$$

applying Cauchy-Schwarz inequality

$$\leq \frac{(\alpha+\beta x\max)^3(b-a)^2}{\lambda\varepsilon^2 n^3}[\sum_{i=1}^{n} \frac{1}{(x_i-\bar{x})^2}]^2$$

$$\leq c^2 \frac{(\alpha+\beta x\max)^3}{\lambda\varepsilon^2 n} \to 0 \quad \text{as} \quad n\to\infty. \quad \text{Hence } \hat{\beta} \xrightarrow{P} \beta .$$

2. $\hat{\alpha} \xrightarrow{P} \alpha$ provided $a \leq x_i \leq b$ for all $i = 1,\ldots,n$ and

$$\sum_{i=1}^{n} \frac{1}{(x_i-\bar{x})^2} \leq \frac{cn}{b-a}, \quad \text{for some } c > 0.$$

<u>Proof</u>: $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$. Since $\hat{\beta} \xrightarrow{P} \beta$ then $-\hat{\beta}\bar{x}_n \xrightarrow{P} -\beta\bar{x}$.

$$P[|\bar{Y}_n - (\alpha+\beta\bar{x})| \geq \varepsilon] \leq \frac{1}{\varepsilon^2} \operatorname{Var}(\bar{Y}_n)$$

$$= \frac{1}{\lambda\varepsilon^2 n^2} \sum_{i=1}^{n} (\alpha+\beta x_i)^3 \leq \frac{(\alpha+\beta x\max)^3}{\lambda\varepsilon^2 n} \to 0 \quad \text{as} \quad n\to\infty.$$

Thus, $\bar{Y}_n \xrightarrow{P} \alpha+\beta\bar{x}$. Hence $\hat{\alpha} \xrightarrow{P} \alpha$.

3. $\hat{\lambda} \xrightarrow{P} \lambda$

<u>Proof</u>: Since $\hat{\beta} \xrightarrow{P} \beta$ and $\hat{\alpha} \xrightarrow{P} \alpha$, then

$$\hat{\beta}x_i \xrightarrow{P} \beta x_i$$

$$\hat{\alpha}+\hat{\beta}x_i \xrightarrow{P} \alpha+\beta x_i$$

$$(Y_i - \hat{\alpha} - \hat{\beta}x_i)^2 \xrightarrow{P} (Y_i - \alpha - \beta x_i)^2$$

$$\frac{1}{(\hat{\alpha}+\hat{\beta}x_i)^2 Y_i} \xrightarrow{P} \frac{1}{(\alpha+\beta x_i)^2 Y_i},$$

provided that $\hat{\alpha}+\hat{\beta}x_i$ is not 0 with probability 1.

$$\frac{(Y_i - \hat{\alpha} - \hat{\beta}x_i)^2}{(\hat{\alpha} + \hat{\beta}x_i)^2 Y_i} \xrightarrow{P} \frac{(Y_i - \alpha - \beta x_i)^2}{(\alpha + \beta x_i)^2 Y_i} . \quad \text{Thus,}$$

$$\frac{1}{\hat{\lambda}} = \frac{1}{n} \sum_{i=1}^{n} \frac{(Y_i - \hat{\alpha} - \hat{\beta}x_i)^2}{(\hat{\alpha} + \hat{\beta}x_i)^2 Y_i} \xrightarrow{P} \frac{1}{n} \sum_{i=1}^{n} \frac{(Y_i - \alpha - \beta x_i)^2}{(\alpha + \beta x_i)^2 Y_i} .$$

We can rewrite the expression on the right hand as

$$\frac{1}{n} \sum_{i=1}^{n} \left(\frac{Y_i}{(\alpha + \beta x_i)^2} - \frac{2}{\alpha + \beta x_i} + \frac{1}{Y_i} \right).$$

Since $\frac{1}{n} \sum_i \frac{Y_i}{(\alpha+\beta x_i)^2} \xrightarrow{P} \frac{1}{n} \sum_i \frac{1}{\alpha+\beta x_i}$ and $\frac{1}{n} \sum_i \frac{1}{Y_i} \xrightarrow{P} \frac{1}{n} \sum_i \frac{1}{\alpha+\beta x_i} + \frac{1}{\lambda}$

then $\frac{1}{\hat{\lambda}} \xrightarrow{P} \frac{1}{\lambda}$. Hence $\hat{\lambda} \xrightarrow{P} \lambda$.

4. Using the central limit theorem,

$$\frac{\sqrt{n_i}(\bar{Y}_{n_i} - (\alpha + \beta x_i))}{\sqrt{\frac{(\alpha + \beta x_i)^3}{\lambda}}} \xrightarrow{L} N(0,1). \quad \sqrt{\frac{(\hat{\alpha} + \hat{\beta}x_i)^3}{\hat{\lambda}}} \xrightarrow{P} \sqrt{\frac{(\alpha + \beta x_i)^3}{\lambda}}$$

implies $\sqrt{\frac{(\alpha + \beta x_i)^3 \hat{\lambda}}{(\hat{\alpha} + \hat{\beta}x_i)^3 \lambda}} \xrightarrow{P} 1$. Hence $\frac{\sqrt{n_i}(\bar{Y}_{n_i} - (\alpha + \beta x_i))}{\sqrt{\frac{(\hat{\alpha} + \hat{\beta}x_i)^3}{\hat{\lambda}}}} \xrightarrow{L} N(0,1)$.

Let $0 < \gamma < 1$. Thus a $100(1-\gamma)\%$ CI on $\alpha + \beta x_i$ is

$$\bar{y}_{n_i} \pm z_{\frac{\gamma}{2}} \sqrt{\frac{(\hat{\alpha} + \hat{\beta}x_i)^3}{n_i \hat{\lambda}}} .$$

5. $\frac{\hat{\beta} - \beta}{\sqrt{Var(\hat{\beta})}} \xrightarrow{\mathcal{L}} N(0,1)$ where $Var(\hat{\beta}) = \frac{1}{n^4 \lambda} \sum_{k=1}^{n} (\alpha+\beta x_k)^3 [\sum_{i=1}^{n} \frac{1}{x_i - \bar{x}} - \frac{n}{x_k - \bar{x}}]^2$ provided $a \leq x_i \leq b$ for all $i = 1,\ldots,n$ and $\sum_{k=1}^{n} \frac{1}{(x_k - \bar{x})^2} \leq \frac{cn}{b-a}$, for some $c > 0$.

<u>Proof</u>: The proof follows from Liapunov theorem. This result still holds even for multiple observations at each x. One can use this result to test hypothesis on β and construct CI on β. Unlike for the simple regression model with zero intercept it is only possible to obtain the following results for a special case where half of the x's are at x_1 and the remaining half at x_2 (n even).

a. MLE

Let $r = \frac{n}{2}$. Thus,

$$\widetilde{\alpha + \beta x_1} = \bar{Y}_L, \qquad \widetilde{\alpha + \beta x_2} = \bar{Y}_U \quad \text{and}$$

$$\frac{1}{\tilde{\lambda}} = \frac{1}{n}[\sum_{i=1}^{r} (\frac{1}{Y_i} - \frac{1}{\bar{Y}_L}) + \sum_{i=r+1}^{n} (\frac{1}{Y_i} - \frac{1}{\bar{Y}_U})]$$

where

$$\alpha^* = \frac{x_1\overline{Y_U} - x_2\overline{Y_L}}{x_1 - x_2} \quad , \quad \beta^* = \frac{\overline{Y_U} - \overline{Y_L}}{x_2 - x_1}$$

$\overline{Y_L}$ and $\overline{Y_U}$ are the mean of the Y's at x_1 and x_2, respectively.

b. UMVU Estimator of $Var(Y_i)$

Applying the result given by Kõsei Iwase and Noriaki Setõ (1983) the UMVUE of $Var(Y_i)$ is

$$\frac{n\bar{Y}_L^3}{(n-1)\tilde{\lambda}} F(1,1.5;\ \frac{n+1}{2};-\frac{\overline{Y_L}}{\tilde{\lambda}}), \qquad i = 1,\ldots,\frac{n}{2},$$

while for $i = \frac{n}{2}+1,\ldots,n$ one should replace $\overline{Y_L}$ by $\overline{Y_U}$.

c. CI for $\alpha + \beta x_1$

The exponent term of the joint distribution of the Y's

$$\lambda[\sum_{i=1}^{r} \frac{(Y_i-\alpha-\beta x_1)^2}{(\alpha+\beta x_1)^2 Y_i} + \sum_{i=r+1}^{n} \frac{(Y_i-\alpha-\beta x_2)^2}{(\alpha+\beta x_2)^2 Y_i}]$$

is a chi-square with n d.f. It can be partitioned into

$$\lambda[\sum_{i=1}^{r} (\frac{1}{Y_i} - \frac{1}{\bar{Y}_L}) + \sum_{i=r+1}^{n} (\frac{1}{Y_i} - \frac{1}{\bar{Y}_U}) + \frac{n}{2}\frac{(\bar{Y}_L-\alpha-\beta x_1)^2}{(\alpha+\beta x_1)^2\bar{Y}_L} + \frac{n}{2}\frac{(\bar{Y}_U-\alpha-\beta x_2)^2}{(\alpha+\beta x_2)^2\bar{Y}_U}]$$

which are independent chi-squares with d.f. $\frac{n}{2}-1$, $\frac{n}{2}-1$, one and one, respectively. Let $\mu_1 = \alpha + \beta x_1$.

Let's now find an interval estimate of μ_1.

Case (i) λ known: Solving the inequality

$$\frac{n\lambda(\bar{y}_L - \mu_1)^2}{2\mu_1^2\bar{y}_L} \leq \chi^2_{1,1-\alpha} \text{ for } \mu_1,$$

then a 100(1-α)% CI on μ_1 is (L,U) where

$$L = \bar{y}_L\left(1+\frac{2\chi_1^2\bar{y}_L(1-n\lambda)}{n\lambda - 2\chi_1^2\bar{y}_L}\right),\ U = \bar{y}_L\left(1+\frac{2\chi_1^2\bar{y}_L(1+n\lambda)}{n\lambda - 2\chi_1^2\bar{y}_L}\right)$$

$\chi_1^2 = \chi^2_{1,1-\alpha}$, provided that the parabola opens upward while if the parabola opens downward the interval is (L,∞).

Case (ii) λ unknown: Based on the set of μ_1's for which

$$\frac{(n-2)(\bar{y}_L - \mu_1)^2}{2\mu_1^2\bar{y}_L D} \leq F_{1,n-2,1-\alpha}$$

where $D = \frac{1}{\tilde{\lambda}}$, then a 100(1-α)% CI on μ_1 is (L,U) where

$$L = \bar{y}_L\left(\frac{n-2}{n-2-2DF\bar{y}_L} - \sqrt{\frac{2(n-2)DF\bar{y}_L}{n-2-2DF\bar{y}_L}}\right),$$

$$U = \bar{y}_L\left(\frac{n-2}{n-2-2DF\bar{y}_L} + \sqrt{\frac{2(n-2)DF\bar{y}_L}{n-2-2DF\bar{y}_L}}\right), \qquad F = F_{1,n-2,1-\alpha},$$

provided that the parabola opens upward. If the parabola opens downward then the interval is (L,∞). One can similarly construct CI on $\alpha + \beta x_2$.

d. Prediction Interval

Suppose in addition to the n independent Y's we have a

future observation Y which is IG with parameters $\alpha + \beta x_1$ and λ. Thus the term in the exponent of the joint distribution of the n Y's plus

$$\lambda \frac{(Y - \alpha - \beta x_1)^2}{(\alpha + \beta x_1)^2 Y}$$

can be rewritten as

$$\lambda [\sum_{i=1}^{r} (\frac{1}{Y_i} - \frac{1}{\bar{Y}_L}) + \sum_{i=r+1}^{n} (\frac{1}{Y_i} - \frac{1}{\bar{Y}_U}) + \frac{n(Y - \bar{Y}_L)^2}{Y\bar{Y}_L(2Y + n\bar{Y}_L)} +$$

$$\frac{(2Y + n\bar{Y}_L - (2+n)(\alpha + \beta x_1))^2}{2(\alpha + \beta x_1)^2 (2Y + n\bar{Y}_L)}]$$

which are independent chi-squares with $\frac{n}{2} - 1$, $\frac{n}{2} - 1$, one and one d.f., respectively.

We want to obtain a prediction interval for Y.

Case (i) λ known: Let us find the set of y's for which

$$\frac{n\lambda(y - \bar{y}_L)^2}{y\bar{y}_L(2y + n\bar{y}_L)} \leq x^2_{1,1-\alpha}$$

then a $100(1-\alpha)\%$ PI on Y is (L,U) where

$$L = \frac{\bar{y}_L}{4}(-n + \frac{(4+n)n\lambda - 2\sqrt{n\bar{y}_L x_1^2 [n\bar{y}_L x_1^2 + 4\lambda(n+2)]}}{n\lambda - 2\bar{y}_L x_1^2})$$

$$U = \frac{\bar{y}_L}{4}\left(-n + \frac{(4+n)n\lambda + 2\sqrt{n\bar{y}_L x_1^2[n\bar{y}_L x_1^2 + 4\lambda(n+2)]}}{n\lambda - 2\bar{y}_L x_1^2}\right),$$

$$x_1^2 = \chi^2_{1,1-\alpha},$$

provided that the parabola opens upward. If the parabola opens downward then the interval is (L,∞).

Case (ii) λ unknown: Based on the set of Y's for which

$$\frac{(n-2)(y-\bar{y}_L)^2}{Dy\bar{y}_L(2y+n\bar{y}_L)} \leq F_{1,n-2,1-\alpha}$$

where $D = 1/\tilde{\lambda}$, then a $100(1-\alpha)\%$ PI on Y is (L,U), provided that the coefficient of y^2 is positive, where

$$L = \frac{\bar{y}_L}{4}\left(-n + \frac{(4+n)(n-2) - 2\sqrt{DF\bar{y}_L(DFn^2\bar{y}_L + 4(n-2)(n+2))}}{n-2-2DF\bar{y}_L}\right)$$

$$U = \frac{\bar{y}_L}{4}\left(-n + \frac{(4+n)(n-2) + 2\sqrt{DF\bar{y}_L(DFn^2\bar{y}_L + 4(n-2)(n+2))}}{n-2-2DF\bar{y}_L}\right)$$

and $F = F_{1,n-2,1-\alpha}$.

If the coefficient of y^2 is negative then the interval is (L,∞).

Case (iii) α and β known: Based on the set of y's for which

$$\frac{n(y-\mu)^2}{Dy\mu^2} \leq F_{1,n,1-\alpha}$$

where $\mu = \alpha + \beta x_1$ and D is the exponent term of the joint distribution of the n Y's excluding $-\frac{\lambda}{2}$, then a $100(1-\alpha)\%$ PI on $\underset{\sim}{Y}$ is

$$\mu\left(1 + \frac{\mu DF - \sqrt{\mu DF(\mu DF + 4n)}}{2n}, 1 + \frac{\mu DF + \sqrt{\mu DF(\mu DF + 4n)}}{2n}\right).$$

One can also construct PI when Y is IG with parameters $\alpha + \beta x_2$ and λ.

e. Confidence Interval on x_0

In addition to the n Y's, suppose we have t Y's which are iid IG with parameters $\alpha + \beta x_0$ and λ, where x_0 is unknown but positive. Then

$$x_0^* = \frac{\bar{y}_0 - \alpha^*}{\beta^*}$$

where α^*, β^* are roots of the likelihood equation and $\bar{y}_0$ is the mean of the t y's at x_0. Thus,

$$\sum_{i=n+1}^{n+t} \frac{(y_i - (\alpha + \beta x_0))^2}{(\alpha + \beta x_0)^2 y_i} = \sum_{i=n+1}^{n+t} \left(\frac{1}{y_i} - \frac{1}{\bar{y}_0}\right) + \frac{t(\bar{y}_0 - (\alpha + \beta x_0))^2}{(\alpha + \beta x_0)^2 \bar{y}_0}.$$

Suppose α and β are known. Since

$$\frac{nt(\bar{y}_0 - (\alpha + \beta x_0))^2}{D\bar{y}_0(\alpha + \beta x_0)^2}, \quad D = \sum_{i=1}^{n} \frac{(y_i - (\alpha + \beta x_i))^2}{(\alpha + \beta x_i)^2 y_i}$$

is an F with d.f. one and n, then by obtaining the set of x_0's for which

$$\frac{nt(\bar{y}_0 - (\alpha + \beta x_0))^2}{D\bar{y}_0(\alpha + \beta x_0)^2} \leq F_{1,n,1-\alpha}$$

a $100(1-\alpha)\%$ CI on x_0 is (L,U) where

$$L = -\frac{\alpha}{\beta} + \frac{\bar{y}_0(nt - \sqrt{DFnt\bar{y}_0})}{\beta(nt - DF\bar{y}_0)},$$

$$U = -\frac{\alpha}{\beta} + \frac{\bar{y}_0(nt + \sqrt{DFnt\bar{y}_0})}{\beta(nt - DF\bar{y}_0)} \quad \text{and} \quad F = F_{1,n,1-\alpha},$$

provided that $nt-DF\bar{y}_0 > 0$. If $nt-DF\bar{y}_0 < 0$ then the interval is (L,∞). If L is negative, then the interval is $(0,\infty)$.

B. Different λ

Consider $Y_i \sim I(\alpha + \beta x_i, \lambda_i)$ with

$$\frac{(\alpha + \beta x_i)^2}{\lambda_i} = k$$

for all $i = 1,\ldots,n$, $\alpha + \beta x_i > 0$, $\lambda_i > 0$ and the Y's are independent.

1. We will try to obtain estimates of the parameters for several cases.

Case (a) all parameters unknown: The log-likelihood is

$$\mathcal{L}(\alpha,\beta,k;\underset{\sim}{y}) = \sum_{i=1}^{n} \ln(\alpha+\beta x_i) - \frac{n}{2}\ln k - \frac{n}{2}\ln(2\pi) - \frac{3}{2}\sum_{i=1}^{n}\ln y_i$$

$$- \frac{1}{2k}\sum_{i=1}^{n}\frac{(y_i-\alpha-\beta x_i)^2}{y_i}. \qquad 4.1$$

After obtaining the partials of the log-likelihood with respect to α, β and k and setting them equal to 0 we get

$$\sum_{i=1}^{n}\frac{1}{\alpha^*+\beta^* x_i} + \frac{1}{k^*}[n - \alpha^*\sum_{i=1}^{n}\frac{1}{y_i} - \beta^*\sum_{i=1}^{n}\frac{x_i}{y_i}] = 0 \qquad 4.2$$

$$\sum_{i=1}^{n}\frac{x_i}{\alpha^*+\beta^* x_i} + \frac{1}{k^*}[\sum_{i=1}^{n} x_i - \alpha^*\sum_{i=1}^{n}\frac{x_i}{y_i} - \beta^*\sum_{i=1}^{n}\frac{x_i^2}{y_i}] = 0 \qquad 4.3$$

$$k^* = \frac{1}{n}\sum_{i=1}^{n}\frac{(y_i-\alpha^*-\beta^* x_i)^2}{y_i} \qquad 4.4$$

Replacing

$$\sum_{i=1}^{n}\frac{x_i}{\alpha^*+\beta^* x_i} \text{ by } \frac{n}{\beta^*} - \frac{\alpha^*}{\beta^*}\sum_{i=1}^{n}\frac{1}{\alpha^*+\beta^* x_i}$$

in 4.3 and adding the resulting expression to $\frac{\alpha^*}{\beta^*}$ times the equa-

tion given in 4.2 and simplifying further we get

$$\alpha^*[n - \alpha^* \sum_{i=1}^{n} \frac{1}{y_i} - \beta^* \sum_{i=1}^{n} \frac{x_i}{y_i}] + nk^* +$$

$$\beta^*[\sum_{i=1}^{n} x_i - \alpha^* \sum_{i=1}^{n} \frac{x_i}{y_i} - \beta^* \sum_{i=1}^{n} \frac{x_i^2}{y_i}] = 0.$$

If we replace the expression for nk^* in the above equation then it reduces to

$$\alpha^* = \bar{y} - \beta^*\bar{x}.$$

However, it was not possible to obtain a closed expression for β^*. One can obtain the MLE using iterative techniques.

Case (b) Two parameters known:

(i) α and β known, k unknown

$$\frac{\partial \mathcal{L}(\alpha_0, \beta_0, k)}{\partial k} = -\frac{n}{2k} + \frac{1}{2k^2} \sum_{i=1}^{n} \frac{(y_i - \alpha_0 - \beta_0 x_i)^2}{y_i}.$$

Setting this equal to zero,

$$\tilde{k} = \frac{1}{n} \sum_{i=1}^{n} \frac{(y_i - \alpha_0 - \beta_0 x_i)^2}{y_i}.$$

Since the second derivative of $\mathcal{L}(\alpha_0,\beta_0,k)$ at $k = \tilde{k}$ is negative and $k > 0$ then $\tilde{k}$ is the MLE.

However, there is no explicit solution for

(ii) α and k known, β unknown,

(iii) β and k known, α unknown.

Case (c) One parameter known:

(i) α known, β and k unknown

$$k^* = \frac{1}{n} \sum_{i=1}^{n} \frac{(y_i - \alpha_0 - \beta^* x_i)^2}{y_i}$$

There is no explicit solution for β^*.

(ii) β known, α and k unknown

Simply interchange α and β in (i).

(iii) k known, α and β unknown

Although there are no explicit expressions for α and β it is still possible to show that the matrix of second partials is negative definite.

From 4.1 we obtain

$$\frac{\partial^2 \mathcal{L}}{\partial^2 \alpha} = - \sum_{i=1}^{n} \left(\frac{1}{(\alpha + \beta x_i)^2} + \frac{1}{ky_i}\right)$$

$$\frac{\partial^2 \mathcal{L}}{\partial \beta \partial \alpha} = - \sum_{i=1}^{n} \left(\frac{x_i}{(\alpha + \beta x_i)^2} + \frac{x_i}{ky_i}\right)$$

$$\frac{\partial^2 \mathcal{L}}{\partial \beta^2} = - \sum_{i=1}^{n} \left(\frac{x_i^2}{(\alpha + \beta x_i)^2} + \frac{x_i^2}{k y_i} \right)$$

The matrix of second partials is

$$M = \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial \alpha^2} & \frac{\partial^2 \mathcal{L}}{\partial \beta \partial \alpha} \\ \frac{\partial^2 \mathcal{L}}{\partial \beta \partial \alpha} & \frac{\partial^2 \mathcal{L}}{\partial \beta^2} \end{pmatrix}$$

Note that the (1,1) entry of M is negative and its determinant

$$\frac{(x_1 - x_2)^2}{(\alpha+\beta x_1)^2 (\alpha+\beta x_2)^2} + \frac{(x_1 - x_3)^2}{(\alpha+\beta x_1)^2 (\alpha+\beta x_3)^3} + \ldots + \frac{(x_1 - x_n)^2}{(\alpha+\beta x_1)^2 (\alpha+\beta x_n)^2}$$

$$+ \ldots + \frac{(x_{n-1} - x_n)^2}{(\alpha + \beta x_{n-1})^2 (\alpha + \beta x_n)^2} + \frac{(x_1 - x_2)^2}{k y_1 (\alpha + \beta x_2)^2} + \ldots +$$

$$\frac{(x_1 - x_n)^2}{k y_1 (\alpha + \beta x_n)^2} + \ldots + \frac{(x_{n-1} - x_n)^2}{k y_{n-1} (\alpha + \beta x_n)^2} + \frac{(x_1 - x_2)^2}{k y_2 (\alpha + \beta x_1)^2} + \ldots +$$

$$\frac{(x_1 - x_n)^2}{k y_n (\alpha + \beta x_1)^2} + \ldots + \frac{(x_{n-1} - x_n)^2}{k y_n (\alpha + \beta x_{n-1})^2} + \frac{(x_1 - x_2)^2}{k^2 y_1 y_2} + \ldots +$$

$$\frac{(x_1 - x_n)^2}{k^2 y_1 y_n} + \ldots + \frac{(x_{n-1} - x_n)^2}{k^2 y_{n-1} y_n}$$

is positive. Thus M is negative definite. Hence $\mathcal{L}(\alpha,\beta,k_0)$ attains its maximum value at (α^*,β^*,k_0) where α^* and β^* are obtained using iterative techniques. α^* and β^* will be MLE and we denote them by $\tilde{\alpha}$ and $\tilde{\beta}$ if (α^*, β^*, k_0) lie in the parameter space $\{(\alpha,\beta,k_0) \mid \alpha+\beta x_i > 0 \text{ for all } i,\ k_0 > 0,\ x_i > 0\}$. Otherwise the MLE will be on the boundary of the parameter space. It was not possible to determine the locus of points formed by $\mathcal{L}(\alpha,\beta,k_0)$. Nevertheless,

$$f(\alpha,\beta) = -\frac{1}{2k_0}\sum_{i=1}^{n}\frac{(y_i-\alpha-\beta x_i)^2}{y_i}$$

is the locus of points that form an ellipse.

2. Asymptotic Results

In addition to the expression given for the estimator of λ it is possible in this case to show that $\alpha^* = \bar{y} - \beta^*\bar{x}$. Although it is not possible to find an explicit expression for β^* from the likelihood equation, using

$$\frac{1}{n}\sum_{i=1}^{n}\frac{Y_i-\bar{Y}}{x_i-\bar{x}},\quad x_i \neq \bar{x} \text{ for all } i,$$

as an estimator of β which is denoted by $\hat{\beta}$ one can obtain the following asymptotic results.

If $a \leq x_i \leq b$ for all $i = 1,\ldots,n$ and $\sum_{i=1}^{n}\frac{1}{(x_i-\bar{x})^2} \leq \frac{cn}{b-a}$, $c > 0$ then

a. $\hat{\beta} \xrightarrow{P} \beta$

<u>Proof</u>:

$$P[|\hat{\beta} - \beta| \geq \varepsilon] \leq \frac{1}{\varepsilon^2} \mathrm{Var}\,(\hat{\beta})$$

$$= \frac{k}{\varepsilon^2 n^4} \sum_{i=1}^{n} (\alpha + \beta x_i)\left(\sum_{k=1}^{n} \frac{1}{x_k - \bar{x}} - \frac{n}{x_i - \bar{x}}\right)^2$$

$$\leq \frac{c^2 k(\alpha + \beta x\max)}{\varepsilon^2 n} \to 0 \quad \text{as } n \to \infty.$$

Hence $\hat{\beta} \xrightarrow{P} \beta$.

b. $\hat{\alpha} \xrightarrow{P} \alpha$

<u>Proof</u>: Since $\hat{\beta} \xrightarrow{P} \beta$ then $-\hat{\beta}\bar{x}_n \xrightarrow{P} -\beta\bar{x}$

$$P[|\bar{Y}_n - (\alpha + \beta\bar{x})| \geq \varepsilon] \leq \frac{1}{\varepsilon^2} \mathrm{Var}\,(\bar{Y}_n)$$

$$= \frac{k(\alpha + \beta x\max)}{\varepsilon^2 n} \to 0 \quad \text{as } n \to \infty.$$

Thus $\bar{Y}_n \xrightarrow{P} \alpha + \beta\bar{x}$ and $\hat{\alpha} \xrightarrow{P} \alpha$.

c. $\hat{k} \xrightarrow{P} k$,

$$\hat{k} = \frac{1}{n} \sum_{i=1}^{n} \frac{(Y_i - \hat{\alpha} - \hat{\beta}x_i)^2}{Y_i}$$

<u>Proof</u>: Since $\hat{\beta} \xrightarrow{P} \beta$ and $\hat{\alpha} \xrightarrow{P} \alpha$ then

$$\hat{\alpha} + \hat{\beta}x_i \xrightarrow{P} \alpha + \beta x_i$$

$$\frac{(Y_i - \hat{\alpha} - \hat{\beta}x_i)^2}{Y_i} \xrightarrow{P} \frac{(Y_i - \alpha - \beta x_i)^2}{Y_i}$$

Thus

$$\hat{k} = \frac{1}{n}\sum_{i=1}^{n}\frac{(Y_i - \hat{\alpha} - \hat{\beta}x_i)^2}{Y_i} \xrightarrow{P} \frac{1}{n}\sum_{i=1}^{n}\frac{(Y_i - \alpha - \beta x_i)^2}{Y_i}$$

We can rewrite

$$\frac{1}{n}\sum_{i=1}^{n}\frac{(Y_i - \alpha - \beta x_i)^2}{Y_i} \quad \text{as} \quad \bar{Y} - 2(\alpha + \beta\bar{x}) + \frac{1}{n}\sum_{i=1}^{n}\frac{(\alpha + \beta x_i)^2}{Y_i}.$$

Since $\bar{Y}_n \xrightarrow{P} \alpha + \beta\bar{x}$ and

$$\frac{1}{n}\sum_{i=1}^{n}\frac{(\alpha + \beta x_i)^2}{Y_i} \xrightarrow{P} \alpha + \beta\bar{x} + k$$

then $\hat{k} \xrightarrow{P} k$.

d. Using the central limit theorem,

$$\frac{\sqrt{n_i}(\bar{Y}_{n_i} - (\alpha + \beta x_i))}{\sqrt{k(\alpha + \beta x_i)}} \xrightarrow{\mathcal{L}} N(0,1).$$

Since $\sqrt{\hat{k}(\hat{\alpha} + \hat{\beta}x_i)} \xrightarrow{P} \sqrt{k(\alpha + \beta x_i)}$ then

$$\sqrt{\frac{k(\alpha + \beta x_i)}{\hat{k}(\hat{\alpha} + \hat{\beta}x_i)}} \xrightarrow{P} 1.$$

Thus

$$\frac{\sqrt{n_i}(\bar{Y}_{n_i} - (\alpha + \beta x_i))}{\sqrt{\hat{k}(\hat{\alpha} + \hat{\beta}x_i)}} \xrightarrow{\mathcal{L}} N(0,1).$$

Let $0 < \gamma < 1$. Hence a $100(1-\gamma)\%$ CI on $\alpha + \beta x_i$ is

$$\bar{Y}_{n_i} \pm z_{\frac{\gamma}{2}} \sqrt{\frac{k(\hat{\alpha} + \hat{\beta}x_i)}{n_i}}$$

e. $\dfrac{\hat{\beta} - \beta}{\sqrt{Var(\hat{\beta})}} \xrightarrow{\mathcal{L}} N(0,1)$ where

$$Var(\hat{\beta}) = \frac{k}{n^4} \sum_{j=1}^{n} \{(\alpha + \beta x_j)(\sum_{i=1}^{n} \frac{1}{x_i - \bar{x}} - \frac{n}{x_j - \bar{x}})^2\}.$$

<u>Proof</u>: This result follows from Liapunov theorem. That $\hat{\beta}$ is asymptotically normal also holds for multiple observations. One can do hypothesis testing on β and construct CI on β.

CHAPTER V

GENERAL NONLINEAR REGRESSION MODEL

Consider $Y_{ij} \sim I(\frac{1}{\underset{\sim}{x}_i'\underset{\sim}{\beta}}, \lambda)$ where $\underset{\sim}{x}_i'$ is a $1\times p$ vector, $\underset{\sim}{\beta}$ is $p\times 1$, $\underset{\sim}{x}_i'\underset{\sim}{\beta} > 0$ and $\lambda > 0$; $i = 1,\ldots,k$, $j = 1,\ldots,n$.

Whitmore (1980) showed that the roots of the likelihood equation are

$$\underset{\sim}{\beta}^* = (X'DX)^{-1}X'\underset{\sim}{j}$$

$$\frac{1}{\lambda^*} = \frac{1}{kn}(\sum_i \sum_j \frac{1}{y_{ij}} - n\underset{\sim}{j}'X\underset{\sim}{\beta}^*)$$

where $X = (\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_k)'$

$$D = \mathrm{diag}(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_k),$$

$$\bar{y}_i = \frac{1}{n}\sum_{j=1}^{n} y_{ij}$$

and $\underset{\sim}{j}$ is a column vector of ones. He also showed that the exponent term of the joint distribution of the Y's can be expressed as

$$n\lambda(\underset{\sim}{\beta}^* - \underset{\sim}{\beta})'X'DX(\underset{\sim}{\beta}^* - \underset{\sim}{\beta}) + \frac{kn\lambda}{\lambda^*}.$$

Although their sum is a chi-square with kn d.f., the two components are not, in general, independent chi-square variables. The question whether

$\underset{\sim}{x}_i'\underset{\sim}{\beta}^*$ lies in the parameter space is discussed for $p = 2$ later in the chapter.

A. Asymptotic Theory for the General Nonlinear Regression Model

1. Modified Least Squares Estimator (MLSE)

Instead of finding the LSE for the General Nonlinear Regression Model (GNRM) let us find the least squares estimator (LSE) for the following model.

Let $\frac{1}{\bar{y}_i} = \frac{1}{n\lambda} + \underset{\sim}{x}_i'\underset{\sim}{\beta} + \varepsilon_i$ where $E(\varepsilon_i) = 0$, $\bar{y}_i = \frac{1}{n}\sum_{j=1}^{n} y_{ij}$, $\underset{\sim}{x}_i'$ is a $1\times p$ vector, $i = 1,\ldots,k$ $(k > p+1)$, $\underset{\sim}{\beta}$ is $p\times 1$,

$$X = \begin{bmatrix} \frac{1}{n} & \underset{\sim}{x}_1' \\ \vdots & \vdots \\ \frac{1}{n} & \underset{\sim}{x}_k' \end{bmatrix}$$

is a $k\times(p+1)$ matrix of known constants with rank $p+1$. Now

$$\frac{1}{\bar{y}_i} = \frac{1}{n\lambda} + \sum_{j=1}^{p} \bar{x}_j\beta_j + \sum_{j=1}^{p} (x_{ij} - \bar{x}_j)\beta_j + \varepsilon_i$$

$$= \alpha + \sum_{j=1}^{p} (x_{ij} - \bar{x}_j)\beta_j + \varepsilon_i \quad \text{where } \alpha = \frac{1}{n\lambda} + \sum_{j=1}^{p} \bar{x}_j\beta_j. \quad \text{Thus}$$

$$\hat{\alpha} = \frac{1}{k}\sum_{i=1}^{k} \frac{1}{\bar{y}_i}$$

$$\frac{1}{\hat{\lambda}} = n\left[\frac{1}{k}\sum_{i=1}^{k} \frac{1}{\bar{y}_i} - \sum_{j=1}^{p} \bar{x}_j\hat{\beta}_j\right]$$

and $\hat{\underset{\sim}{\beta}} = (X_c'X_c)^{-1}X_c'\underset{\sim}{w}$ where $X_c = (\underset{\sim}{x}_1 - \underset{\sim}{j}\bar{x}_1, \ldots, \underset{\sim}{x}_p - \underset{\sim}{j}\bar{x}_p)$ with

$$\bar{x}_j = \frac{1}{k}\sum_{i=1}^{k} x_{ij},\ j = 1,\ldots,p \text{ and}$$

$$\underset{\sim}{w} = \left(\frac{1}{\bar{y}_1},\ldots,\frac{1}{\bar{y}_k}\right)'.$$

$\hat{\underset{\sim}{\beta}}$ and $\frac{1}{\hat{\lambda}}$ are unbiased for $\underset{\sim}{\beta}$ and $\frac{1}{\lambda}$, respectively. Using the central limit theorem,

$$\sqrt{n}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \xrightarrow[n\to\infty]{L} N_p(\underset{\sim}{0}, G'\Sigma G) \quad \text{where}$$

$$\Sigma = \frac{1}{\lambda}\,\mathrm{Diag}\left(\underset{\sim}{x}_1'\underset{\sim}{\beta} + \frac{2}{n\lambda}, \underset{\sim}{x}_2'\underset{\sim}{\beta} + \frac{2}{n\lambda}, \ldots, \underset{\sim}{x}_k'\underset{\sim}{\beta} + \frac{2}{n\lambda}\right) \text{ and } G = X_c(X_c'X_c)^{-1}.$$

Note that

$$E\left(\frac{1}{\bar{y}_i}\right) = \underset{\sim}{x}_i'\underset{\sim}{\beta} + \frac{1}{n\lambda}, \quad E\left(\frac{1}{\bar{y}_i^2}\right) = (\underset{\sim}{x}_i'\underset{\sim}{\beta})^5\, E(\bar{y}_i^{-3}) \quad \text{and}$$

$$\mathrm{Var}\left(\frac{1}{\bar{y}_i}\right) = \frac{1}{n\lambda}\left(\underset{\sim}{x}_i'\underset{\sim}{\beta} + \frac{2}{n\lambda}\right).$$

Suppose the first entry of each of the $\underset{\sim}{x}_i$'s is one, then our model after reparameterizing becomes

$$\frac{1}{\bar{y}_i} = \alpha + \sum_{j=2}^{p}(x_{ij} - \bar{x}_j)\beta_j + \varepsilon_i, \quad \text{where}$$

$$\alpha = \beta_1 + \frac{1}{n\lambda} + \sum_{j=2}^{p}\bar{x}_j\beta_j$$

and all the assumptions given earlier will hold except that in this case the rank of X is p. Suppose we denote

$$(\underset{\sim}{x}_2 - j\underset{\sim}{\bar{x}}_2 , \underset{\sim}{x}_3 - j\underset{\sim}{\bar{x}}_3 , \ldots , \underset{\sim}{x}_p - j\underset{\sim}{\bar{x}}_p) , \quad \bar{x}_j = \frac{1}{k} \sum_{i=1}^{k} x_{ij},$$

by X_c, then

$$\hat{\alpha} = \frac{1}{k} \sum_{i=1}^{k} \frac{1}{\bar{y}_i} \quad \text{and}$$

$$(\hat{\beta}_2, \ldots, \hat{\beta}_p)' = (X_c' X_c)^{-1} X_c' \underset{\sim}{w}.$$

It is not possible to find an explicit expression for the estimates of β_1 and $\frac{1}{\lambda}$. Suppose we estimate $\frac{1}{\lambda}$ by

$$\frac{1}{k(n-1)} \sum_{i=1}^{k} \sum_{j=1}^{n} \left(\frac{1}{y_{ij}} - \frac{1}{\bar{y}_i}\right), \; n > 1 \text{ and } k > p, \text{ then}$$

$$\hat{\beta}_1 = \frac{1}{k} \sum_{i=1}^{k} \frac{1}{\bar{y}_i} - \frac{1}{n\hat{\lambda}} - \sum_{j=2}^{p} \bar{x}_j \hat{\beta}_j$$

and is unbiased for β_1. Since an estimator of $\frac{1}{\lambda}$ is used it seems reasonable to call $\hat{\underset{\sim}{\beta}}$ a modified least squares estimator. Using the central limit theorem,

$$\sqrt{n}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \xrightarrow[n\to\infty]{L} N_p\left(\underset{\sim}{0} , \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix}\right) \text{ where}$$

$$\sigma_{11} = \frac{1}{k\lambda}(\beta_1 + \sum_{j=2}^{p} \bar{x}_j\beta_j + \frac{2}{n\lambda}) + \frac{2}{k(n-1)n} +$$

$$(\bar{\underset{\sim}{x}}'(X_c'X_c)^{-1} X_c' - \frac{2}{k}\underset{\sim}{j}')\Sigma X_c(X_c'X_c)^{-1}\bar{\underset{\sim}{x}}$$

and $\sigma_{22} = (X_c'X_c)^{-1}X_c'\Sigma X_c(X_c'X_c)^{-1}$ with $\Sigma = \frac{1}{\lambda}\text{Diag}\,(\underset{\sim}{x}_1'\beta + \frac{2}{n\lambda},\ldots,\underset{\sim}{x}_k'\underset{\sim}{\beta} + \frac{2}{n\lambda})$. Note that

$$\text{Var}(\frac{1}{\bar{y}_i}) = \frac{1}{n\lambda}(\underset{\sim}{x}_i'\underset{\sim}{\beta} + \frac{2}{n\lambda}).$$

We also know that $(\hat{\beta}_2,\ldots,\hat{\beta}_p)'$ and $\frac{1}{\hat{\lambda}}$ are independent for the former is a function of $\bar{y}_1,\ldots,\bar{y}_k$ which are independent of the latter while $\hat{\beta}_1$ is asymptotically independent of $\frac{1}{\hat{\lambda}}$.

2. Root of the Log-Likelihood Equation

One can also derive an asymptotic distribution for the root of the likelihood equation. However, it is not even possible to find the exact expression for the expected value of the root. Thus, we can neither determine its bias nor measure the quality of our estimate. Nevertheless, one can obtain the following asymptotic distribution. The log-likelihood is

$$\mathcal{L}(\underset{\sim}{\beta},\lambda;\underset{\sim}{y}) = \frac{kn}{2}\ln\lambda - \frac{\lambda}{2}[n\underset{\sim}{\beta}'X'DX\underset{\sim}{\beta} - 2n\underset{\sim}{j}'X\underset{\sim}{\beta} + \sum_{i,j}\frac{1}{y_{ij}}]$$

where

$$D = \text{Diag}(\bar{y}_1,\ldots,\bar{y}_k)$$

and

$$\bar{y}_i = \frac{1}{n}\sum_{j=1}^{n} y_{ij}, \qquad i = 1,\ldots,k.$$

Let us obtain the first and second partials of $\mathcal{L}(\underset{\sim}{\beta},\lambda;\underset{\sim}{y})$ with respect to $\underset{\sim}{\beta}$ and λ. Thus

$$\frac{\partial\mathcal{L}}{\partial\underset{\sim}{\beta}} = -\frac{\lambda}{2}[2nX'DX\underset{\sim}{\beta} - 2nX'\underset{\sim}{j}]$$

$$\frac{\partial\mathcal{L}}{\partial\lambda} = \frac{kn}{2\lambda} - \frac{1}{2}[n\underset{\sim}{\beta}'X'DX\underset{\sim}{\beta} - 2n\underset{\sim}{j}'X\underset{\sim}{\beta} + \sum_{i,j}\frac{1}{y_{ij}}]$$

$$\frac{\partial^2\mathcal{L}}{(\partial\underset{\sim}{\beta})\partial\underset{\sim}{\beta}'} = -n\lambda X'DX$$

$$\frac{\partial^2\mathcal{L}}{\partial\lambda^2} = -\frac{kn}{2\lambda^2}$$

$$\left.\frac{\partial^2\mathcal{L}}{(\partial\underset{\sim}{\beta})\partial\lambda}\right]_{\underset{\sim}{\beta}=\underset{\sim}{\beta}^*} = 0.$$

Thus, using the result given in Cox and Hinkley (1974, p. 294)

$$\sqrt{n}(\underset{\sim}{\beta}^* - \underset{\sim}{\beta}) \xrightarrow{L} N_p(\underset{\sim}{0}, \frac{1}{\lambda}(X'DX)^{-1}),$$

$$\sqrt{n}(\frac{1}{\lambda^*} - \frac{1}{\lambda}) \xrightarrow{L} N(0,\frac{2}{k}\lambda^2)$$

and β^* and λ^* are asymptotically independent. If we estimate $\frac{1}{\lambda}$ by

$$\frac{1}{\hat{\lambda}} = \frac{1}{k(n-1)} \sum_{i=1}^{k} \sum_{j=1}^{n} \left(\frac{1}{y_{ij}} - \frac{1}{\bar{y}_i}\right),$$

then $\frac{k(n-1)\lambda}{\hat{\lambda}}$ is a $\chi^2_{k(n-1)}$ and

$$\frac{\sqrt{n}(\underset{\sim}{\ell}'\underset{\sim}{\beta}^* - \underset{\sim}{\ell}'\underset{\sim}{\beta})}{\sqrt{\frac{1}{\hat{\lambda}}\underset{\sim}{\ell}'(X'DX)^{-1}\underset{\sim}{\ell}}} \xrightarrow{L} \text{t distribution with } k(n-1) \text{ d.f.}$$

The results given above and in the previous section can be used to test hypotheses and construct confidence regions on $\underset{\sim}{\beta}$, a subvector of $\underset{\sim}{\beta}$ or a linear function of $\underset{\sim}{\beta}$.

Let us consider two particular cases of the GNRM.

B. Simple Nonlinear Regression Model

Consider

$$Y_i \sim I\left(\frac{1}{\alpha + \beta x_i}, \lambda\right), \qquad i = 1,\ldots,n.$$

1. Folks and Chhikara (unpublished monograph) assumed that the model is linear for the reciprocal of y_i, i.e.

$$\frac{1}{y_i} = E\left(\frac{1}{y_i}\right) + \frac{1}{y_i} - E\left(\frac{1}{y_i}\right) = \alpha + \beta x_i + e_{\lambda i}$$

where

$$E(e_{\lambda i}) = \frac{1}{\lambda}$$

$$\mathrm{Var}(e_{\lambda i}) = \frac{\alpha + \beta x_i}{\lambda} + \frac{2}{\lambda^2}.$$

In this case,

$$\alpha^* = \frac{1}{\bar{y}} - \beta^* \frac{\sum_i x_i y_i}{\sum_i y_i} = \frac{1}{\bar{y}}(1 - \frac{\beta^*}{n} \sum_i x_i y_i)$$

$$\beta^* = \frac{n \sum_{i=1} (x_i - \bar{x})(y_i - \bar{y})}{(\sum_{i=1}^{n} x_i y_i)^2 - (\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} x_i^2 y_i)} .$$

The estimate of β can be rewritten as

$$\frac{-n \sum_{i=1}^{n} (x_i - \bar{x}) y_i}{\sum_{i<j} (x_i - x_j) y_i y_j} .$$

$$\frac{1}{\lambda^*} = \frac{1}{n} \sum_{i=1}^{n} \frac{((\alpha^* + \beta^* x_i) y_i - 1)^2}{y_i}$$

Thus, the fitted model is

$$\frac{1}{y^*} = \alpha^* + \beta^* x$$

$$\sum_{j=1}^{n} (\frac{1}{y_j} - \frac{1}{y_j^*}) = \sum_j (\frac{1}{y_j} - \alpha^* - \beta^* x_j)$$

$$= \sum_j \frac{1}{y_j} - \frac{1}{\bar{y}}(n - \beta^* \sum_i x_i y_i) - \beta^* \sum_j x_j$$

$$= \sum_j (\frac{1}{y_j} - \frac{1}{\bar{y}}) - \frac{n(\sum_i (x_i - \bar{x}) y_i)^2}{\bar{y} \sum_{i<j} (x_i - x_j)^2 y_i y_j}$$

Hence

$$\sum_{j=1}^{n}\left(\frac{1}{y_j}-\frac{1}{\bar{y}}\right) = \sum_{j=1}^{n}\left(\frac{1}{y_j}-\frac{1}{y_j^*}\right)+\frac{n\left(\sum_i (x_i-\bar{x})y_i\right)^2}{\bar{y}\sum_{i<j}(x_i-x_j)^2 y_i y_j} \qquad 5.2$$

These results can be presented in the following analysis of reciprocals table which is quite analogous to the result of the normal regression model.

AOR

Source	Sum of Reciprocals
Total	$\sum_i \frac{1}{y_i}$
$R(\alpha)$	$\frac{n}{\bar{y}}$
Total Corrected	$\sum_i \frac{1}{y_i}-\frac{n}{\bar{y}}$
$R(\beta\|\alpha)$	$\frac{n\left(\sum_i (x_i-\bar{x})y_i\right)^2}{\bar{y}\sum_{i<j}(x_i-x_j)^2 y_i y_j}$
Residual	$\sum_i\left(\frac{1}{y_i}-\frac{1}{y_i^*}\right)$

Although the term on the left-hand side of 5.2 is a Chi-square with n-1 d.f. it is not yet possible to determine the distribution of the two terms on the right-hand side of 5.2. However, some results are obtained for the following special case.

2. Special Case. Suppose half of the x's are x_1 and the remaining half x_2. Without loss of generality let $x_2 > x_1$. Let $r = \frac{n}{2}$ (n even).

a. $$\sum_i (x_i-\bar{x})y_i = \frac{x_2-x_1}{2}\left(\sum_{i=r+1}^{n} y_i - \sum_{i=1}^{r} y_i\right)$$

$$\sum_{i<j} (x_i - x_j)^2 y_i y_j = (x_1 - x_2)^2 (\sum_{i=1}^{r} y_i)(\sum_{i=r+1}^{n} y_i)$$

$$\sum_{i=1}^{n} (\frac{1}{y_i} - \frac{1}{\tilde{y}_i}) = \sum_{i=1}^{r} (\frac{1}{y_i} - \frac{1}{\bar{y}_L}) + \sum_{i=r+1}^{n} (\frac{1}{y_i} - \frac{1}{\bar{y}_U}) \qquad 5.3$$

$$R(\beta|\alpha) = \frac{n(\sum_{i=1}^{r} y_i - \sum_{i=r+1}^{n} y_i)^2}{4\bar{y}(\sum_{i=1}^{r} y_i)(\sum_{i=r+1}^{n} y_i)} \qquad 5.4$$

λ times the expression given in the right-hand side of 5.3 is a chi-square with n-2 d.f. If $\beta = 0$ then

$$\sum_{i=1}^{r} y_i \sim I(\frac{r}{\alpha}, r^2\lambda)$$

$$\sum_{i=r+1}^{n} y_i \sim I(\frac{r}{\alpha}, r^2\lambda)$$

and they are independent. Thus $\lambda R(\beta|\alpha)$ is a chi-square with one d.f. Hence we can use these results to test $\beta = 0$.

(i) For λ known we will reject H_0: $\beta = 0$ if $\lambda R(\beta|\alpha) > \chi^2_{1,1-\alpha}$.

(ii) For λ unknown we will reject H_0: $\beta = 0$ if $(n-2)R(\beta|\alpha)$ divided by the quantity in the right-hand side of 5.3 is greater than $F_{1,n-2,1-\alpha}$.

Note that $E(\beta^*) = \beta$, $E(\alpha^*) = \alpha + \frac{2}{n\lambda}$ and $E(\lambda^*) = \frac{n\lambda}{n-4}$. Thus β^* is unbiased for β while α^* and λ^* are biased for α and λ.

b. UMVU Estimator of $Var(Y_i)$. The UMVUE of $Var(Y_i) = \frac{1}{(\alpha + \beta x_1)^3 \lambda}$, $i = 1, \ldots, \frac{n}{2}$ is

$$\frac{n\bar{y}_L^3}{(n-1)\tilde{\lambda}} F(1,1.5;\frac{n+1}{2};-\frac{\bar{y}_L}{\tilde{\lambda}}), \quad i=1,\ldots,\frac{n}{2}.$$

For $i=\frac{n}{2}+1,\ldots,n$ one should replace $\bar{y}_L$ by $\bar{y}_U$.

c. Confidence Interval for $\alpha+\beta x_1$. The exponent term of the joint distribution of the Y's is

$$\sum_{i=1}^{r} \frac{((\alpha+\beta x_1)y_i-1)^2}{y_i} + \sum_{i=r+1}^{n} \frac{((\alpha+\beta x_2)y_i-1)^2}{y_i} \qquad 5.5$$

$$= \sum_{i=1}^{r} (\frac{1}{y_i}-\frac{1}{\bar{y}_L}) + \sum_{i=r+1}^{n} (\frac{1}{y_i}-\frac{1}{\bar{y}_U}) + \frac{n}{2}\frac{(\bar{y}_L(\alpha+\beta x_1)-1)^2}{\bar{y}_L}$$

$$+ \frac{n}{2}\frac{(\bar{y}_U(\alpha+\beta x_2)-1)^2}{\bar{y}_U} \qquad 5.6$$

λ times the four terms in 5.6 are independent chi-squares with d.f. $\frac{n}{2}-1$, $\frac{n}{2}-1$, one and one, respectively.

Let's now find an interval estimate of μ where $\mu=\alpha+\beta x_1$.

Case (i) λ known: Let's now determine the set of μ's for which

$$\frac{n\lambda}{2}\frac{(\bar{y}_L\mu-1)^2}{\bar{y}_L} \le \chi^2_{1,1-\alpha}.$$

Thus a $100(1-\alpha)\%$ CI on μ is (L,U), where

$$L = \frac{1}{\bar{y}_L} - \sqrt{\frac{2\chi^2_{1,1-\alpha}}{n\lambda\bar{y}_L}} \quad \text{and} \quad U = \frac{1}{\bar{y}_L} + \sqrt{\frac{2\chi^2_{1,1-\alpha}}{n\lambda\bar{y}_L}}.$$

If L is negative then the interval on μ is $(0,U)$.

Case (ii) λ unknown: Based on the set of μ's for which

$$\frac{(n-2)(\bar{y}_L\mu - 1)^2}{2D\bar{y}_L} \leq F_{1,n-2,1-\alpha}$$

where D is the sum of the first two terms of 5.6 divided by n, a $100(1-\alpha)\%$ CI on μ is (L,U), where

$$L = \frac{1}{\bar{y}_L} - \sqrt{\frac{2DF}{(n-2)\bar{y}_L}} \quad \text{and} \quad U = \frac{1}{\bar{y}_L} + \sqrt{\frac{2DF}{(n-2)\bar{y}_L}}, \quad F = F_{1,n-2,1-\alpha}.$$

If L is negative then the interval on μ is (0,U).

One can similarly construct CI on $\alpha + \beta x_2$.

d. Prediction Intervals. Suppose in addition to the n Y's, we have a future observation Y which is IG with parameters the reciprocal of $\alpha + \beta x_1$ and λ. Thus 5.6 plus $((\alpha + \beta x_1)y - 1)^2/y$ can be rewritten as

$$\sum_{i=1}^{r} \left(\frac{1}{y_i} - \frac{1}{\bar{y}_L}\right) + \sum_{i=r+1}^{n} \left(\frac{1}{y_i} - \frac{1}{\bar{y}_u}\right) + \frac{n}{2}\,\frac{((\alpha + \beta x_2)\bar{y}_u - 1)^2}{\bar{y}_u} + \frac{n(y - \bar{y}_L)^2}{y\bar{y}_L(2y + n\bar{y}_L)}$$

$$+ \frac{((2y + n\bar{y}_L)(\alpha + \beta x_1) - 2 - n)^2}{2(2y + n\bar{y}_L)}.$$

λ times all of the five terms are independent chi-squares with d.f. $\frac{n}{2} - 1$, $\frac{n}{2} - 1$, one, one and one, respectively.

Case (i) λ known: Based on the set of y's for which

$$\frac{n\lambda(y - \bar{y}_L)^2}{y\bar{y}_L(2y + n\bar{y}_L)} \leq \chi^2_{1,1-\alpha}$$

a 100(1-α)% PI on y is (L,U), where

$$L = \frac{\bar{y}_L}{4}\left(-n + \frac{n\lambda(4+n) - 2\bar{y}_L\sqrt{n\bar{y}_L X_1^2[n\bar{y}_L X_1^2 + 4\lambda(n+2)]}}{n\lambda - 2\bar{y}_L X_1^2}\right), \quad X_1^2 = \chi^2_{1,1-\alpha},$$

$$U = \frac{\bar{y}_L}{4}\left(-n + \frac{n\lambda(4+n) + 2\bar{y}_L\sqrt{n\bar{y}_L X_1^2[n\bar{y}_L X_1^2 + 4\lambda(n+2)]}}{n\lambda - 2\bar{y}_L X_1^2}\right), \quad X_1^2 = \chi^2_{1,1-\alpha},$$

provided that the parabola opens upward. If the parabola opens downward then the interval is (L,∞).

Case (ii) λ unknown: Solving the inequality

$$\frac{(n-2)(y-\bar{y}_L)^2}{Dy\bar{y}_L(2y+n\bar{y}_L)} \leq F_{1,n-2,1-\alpha}$$

where D is the sum of the first two terms of 5.6 divided by n, for y, a 100(1-α)% PI on y is (L,U), where

$$L = \frac{\bar{y}_L}{4}\left(-n + \frac{(n-2)(4+n) - 2\sqrt{DF\bar{y}_L(DFn^2\bar{y}_L + 4(n-2)(n+2))}}{n-2-2DF\bar{y}_L}\right)$$

$$U = \frac{\bar{y}_L}{4}\left(-n + \frac{(n-2)(4+n) + 2\sqrt{DF\bar{y}_L(DFn^2\bar{y}_L + 4(n-2)(n+2))}}{n-2-2DF\bar{y}_L}\right),$$

$$F = F_{1,n-2,1-\alpha},$$

provided that the parabola opens upward. If the parabola opens downward

then the interval is (L,∞).

Case (iii) α and β known: Solving the inequality

$$\frac{n((\alpha+\beta x_1)y-1)^2}{Dy} \leq F_{1,n,1-\alpha},$$

for y, where D is the expression given in 5.5 then a $100(1-\alpha)\%$ PI on y is

$$\frac{1}{\mu}\left(1+\frac{DF-\sqrt{DF(DF+4n\mu)}}{2n\mu},\ 1+\frac{DF+\sqrt{DF(DF+4n\mu)}}{2n\mu}\right)$$

where $\mu=\alpha+\beta x_1$ and $F = F_{1,n-2,1-\alpha}$.

3. Estimate of x_0 and CI on x_0. In addition to the n independent Y's suppose we have t Y's which are iid with parameters the reciprocal of $\alpha+\beta x_0$ and λ with x_0 being unknown but positive. Thus

$$x_0^* = \frac{1-\alpha^*\bar{y}_0}{\beta^*\bar{y}_0}$$

where α^* and β^* are solutions of the likelihood equation and $\bar{y}_0$ is the mean of those y's at x_0.

In this case we can only construct CI on x_0 when α and β are both known. Note that

$$\frac{nt(\bar{y}_0(\alpha+\beta x_0)-1)^2}{D\bar{y}_0} \sim F_{1,n} \quad \text{where} \quad D = \sum_{i=1}^{n}\frac{(y_i(\alpha+\beta x_i)-1)^2}{y_i}.$$

Solving the inequality

$$\frac{nt(\bar{y}_0(\alpha+\beta x_0)-1)^2}{D\bar{y}_0} \leq F_{1,n,1-\alpha} \text{ for } x_0,$$

then a $100(1-\alpha)\%$ CI on x_0 is (L,U) where

$$L = \frac{1}{\beta\bar{y}_0}(1-\alpha\bar{y}_0 - \sqrt{\frac{DF\bar{y}_0}{nt}}) \quad \text{and } U = \frac{1}{\beta\bar{y}_0}(1-\alpha\bar{y}_0 + \sqrt{\frac{DF\bar{y}_0}{nt}}), \quad F = F_{1,n,1-\alpha}.$$

If L is negative and U is positive then the interval is $(0,U)$. However, if U is negative then there is no interval estimate of x_0.

4. Consider the model:

$$E(Y) = \frac{1}{\alpha+\beta x}$$

We can have four different graphs when

a) $\alpha > 0,\ \beta > 0$

b) $\alpha > 0,\ \beta < 0$

c) $\alpha < 0,\ \beta > 0$

d) $\alpha < 0,\ \beta < 0$

Suppose we assume (a) and $x > 0$. It is possible to come up with an example where the estimates of one of the parameters is negative

Let

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 20 \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Thus $\alpha^* = \frac{2132}{3029}$, $\beta^* = -\frac{57}{3029}$.

Nevertheless, $\alpha^* + \beta^* x_i$ is positive for all x_i. We might still assume any of the remaining cases and still $\alpha^* + \beta^* x_i$ is positive for all x_i used in generating the data. Is $\alpha^* + \beta^* x_i$ always positive?

Consider the case where

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{pmatrix}, \qquad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$\alpha^* + x_i\beta^* = \frac{1}{\bar{y}} - \frac{\beta^*}{n\bar{y}} \sum_{i=1}^{n} x_i y_i + x_i\beta^* = \frac{1}{\bar{y}} + (x_i - \frac{\sum_i x_i y_i}{n\bar{y}})\beta^* \qquad 5.7$$

replacing β^* by $\dfrac{n\sum_i (x_i - \bar{x})y_i}{(\sum_i x_i y_i)^2 - (\sum_i y_i)(\sum_i x_i^2 y_i)}$ and letting $x_i = x_1$ then for

$x_1 = 2$, $x_2 = 3$, $x_3 = 5$, $y_1 = 1$ and $y_2 = 10$, then 5.7 reduces to

$$15y_3^2 + 155y_3 - 110.$$

Equating this resulting expression to zero (note that y_3 is positive), the values of y_3 for which 5.7 will be negative are between 0 and $\frac{2}{3}$. If we let $y_3 = 0.5$ then

$$\alpha^* = -.7391304$$

$$\beta^* = \frac{1}{3}.$$

In this example $\alpha^* + 2\beta^*$ is negative which unfortunately does not lie in the parameter space. What should we do to overcome this problem?

When the estimates of $\alpha + \beta x_i$ lie outside the parameter space it is either to the left of $\alpha + (\min x_i)\beta = 0$ or below $\alpha + (\max x_i)\beta = 0$ (assuming the horizontal is the α-axis while the vertical is the β-axis).

Let us consider the former. We want to find a value for α and β where

$$\mathcal{L}(\alpha,\beta,\lambda;y) = \frac{n}{2}\ln\lambda - \frac{\lambda}{2}\sum_{i=1}^{n}\frac{(y_i(\alpha+\beta x_i)-1)^2}{y_i} + c, \qquad 5.8$$

and $\quad \alpha + \min x_i\,\beta = 0 \qquad 5.9$

intersect. Replacing the expression for α obtained from 5.9 into 5.8 and then obtaining the derivative w.r.t. β, finally solving for β we get

$$\tilde{\beta} = \frac{\Sigma x_i - n(\min x_i)}{\sum_i y_i(x_i - \min x_i)^2}$$

and

$$\tilde{\alpha} = -(\min x_i)\tilde{\beta}.$$

We can similarly obtain $\tilde{\alpha}$ and $\tilde{\beta}$ when the point, whose coordinates are the the root of the likelihood, is below $\alpha + (\max x_i)\beta = 0$.

5. Four Estimation Techniques -- Consider $Y_i \sim I(\frac{1}{\alpha+\beta x_i}, \lambda)$ $i = 1, \ldots, n$. To estimate the mean of the simple nonlinear regression model one can use

a. Root Method. The estimates of α and β are

$$\beta^* = \frac{n\sum_i (x_i - \bar{x})y_i}{(\sum_i x_i y_i)^2 - (\Sigma y_i)(\sum_i x_i^2 y_i)}$$

$$\alpha^* = \frac{1}{\bar{y}} - \beta^* \frac{\sum_i x_i y_i}{\sum_i y_i} .$$

b. MLE. The estimates of the parameters are

$$\tilde{\beta} = \frac{\Sigma x_i - n(\min x_i)}{\Sigma y_i (x_i - \min x_i)^2}$$

$$\tilde{\alpha} = -(\min x_i)\tilde{\beta} .$$

Assuming the horizontal is the α-axis and the vertical is the β-axis, one should replace $\min x_i$ by $\max x_i$ if $(\tilde{\alpha},\tilde{\beta})$ lies below $\alpha + (\max x_i)\beta = 0$.

c. MLSE. In this case

$$\hat{\beta} = \frac{k\sum_i \frac{x_i}{\bar{y}_i} - (\sum_i x_i)(\sum_i \frac{1}{\bar{y}_i})}{k\sum_i x_i^2 - (\sum_i x_i)^2} \quad \text{and}$$

$$\hat{\alpha} = \frac{1}{k} \sum_{i=1}^{k} \frac{1}{\bar{y}_i} - \hat{\beta}\bar{x} - \frac{1}{n\hat{\lambda}} \quad \text{where}$$

$$\frac{1}{\hat{\lambda}} = \frac{1}{k(r-1)} \sum_{i=1}^{k} \sum_{j=1}^{n} \left(\frac{1}{y_{ij}} - \frac{1}{\bar{y}_i}\right)$$

d. Least Squares Estimator (LSE). To obtain the least squares estimator we should minimize $\sum_i (y_i - \frac{1}{\alpha + \beta x_i})^2$ with respect to α and β.

To determine which of these techniques is more efficient a simulation of 500 samples with three different numbers of observations, three λ's and three α, β values was run. The following criteria are used in comparing the different methods.

(i) error sum of squares (ESS),

(ii) mean square error (MSE),

(iii) determinant of the mean square error matrix (Det(MSE)).

To obtain some of these results the IMSL nonlinear least squares subroutine and the SAS package are used.

The following tables contain the four estimates of the mean of the simple nonlinear regression model, the error sum of squares (ESS), the mean square error (MSE) and the determinant of the mean square error matrix.

For a fixed sample size with increase in λ, the general trend is that

a) the quality of the estimates improve,

b) there is a decrease in the error sum of squares,

c) there is a decrease in the mean square error and the determinant of the mean square error matrix.

It turned out that the subroutine used for computing the LSE is dependent upon the initial values. As the initial values get closer to the true values, the quality of the estimates improved. One can also observe a smaller error sum of squares, smaller than even all of the ESS and yet the estimates are considerably off from the true value. This might be due to the weakness of the subroutine. If result of such nature is not due to the subroutine then it may not be advisable to use the LSE for one can't tell in reality how close the estimates are to the

TABLE I

A COMPARISON OF THE ESTIMATES OF THE MEAN OF THE SIMPLE NONLINEAR REGRESSION MODEL USING SIMULATED DATA FOR $\underset{\sim}{\beta} = \binom{.02}{.06}$, TEN OBSERVATIONS AND λ TAKING VALUES 1, 10 AND 50, RESPECTIVELY

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-5}$)
Root -.028 .027	31,969,854	887 139	1.2
MLE -.021 .026	30,707,514	695 141	96
LSE -153.47 38.4	6092.5	5.9693×10^{9} 3.730×10^{8}	-5.008×10^{8}
MLSE -.273 .051	1,565,772	17493 200	24.8

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-6}$)
Root -.023 .056	1383.2	812 25	1.34
MLE -.023 .056	733.5	809* 25	1.33
LSE -.009 .056	38.2	1259 105	4.22
MLSE -.188 .062	14,798,429	7002 45	17.99

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-8}$)
Root .003 .059	9.5	221 6	5
MLE .003 .059	9.5	221 6	5
LSE -.001 .062	5.4	482 31	23
MLSE -.109 .061	98941.2	2276 9	143

*Increase in MSE.

TABLE II

A COMPARISON OF THE ESTIMATES OF THE MEAN OF THE SIMPLE NONLINEAR REGRESSION MODEL USING SIMULATED DATA FOR $\underset{\sim}{\beta} = \binom{.02}{.06}$, THIRTY OBSERVATIONS AND λ TAKING VALUES 1, 10 AND 50, RESPECTIVELY

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-6}$)
Root -.037 .026	2,918,004	408 124	1.8
MLE -.036 .025	2,403,681	379 126	1.6
LSE -.017 .02	16,790.8	258 170	3.2
MLSE -.124 .036	207,040,953	2845 87	20.2

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-7}$)
Root -.031 .055	394.6	435* 10	4
MLE -.031 .055	394.6	435* 10	4
LSE -.02 .054	311.9	582* 37	15
MLSE -.049 .058	36,959,571	807 11	8

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-8}$)
Root .003 .059	41.3	93 2	1
MLE .003 .059	41.3	93 2	1
LSE .003 .06	37.1	198 10	6
MLSE -.0003 .059	48.2	122 3	2

*Increase in MSE

TABLE III

A COMPARISON OF THE ESTIMATES OF THE MEAN OF THE SIMPLE NONLINEAR REGRESSION MODEL USING SIMULATED DATA FOR $\underset{\sim}{\beta} = \binom{.02}{.06}$, FIFTY OBSERVATIONS AND λ TAKING VALUES 1, 10 AND 50, RESPECTIVELY

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-7}$)
Root -.04 .026	16,285,754	396 121	9
MLE -.039 .025	1,369,564	382 122	82
LSE -.02 .02	36,897	214 168	16
MLSE -.101 .033	33,376,690	1803 87	85

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-7}$)
Root -.038 .056	736.5	444* 6	3
MLE -.038 .056	736.5	444* 6	3
LSE -.026 .054	648.6	457* 24	10
MLSE -.051 .058	18078.1	709 7	5

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-9}$)
Root .0003 .059	77.3	80 1	8
MLE .0003 .059	77.3	80 1	8
LSE -.001 .06	73	151 7	40
MLSE -.001 .059	84.3	98 2	11

*Increase in MSE

TABLE IV

A COMPARISON OF THE ESTIMATES OF THE MEAN OF THE SIMPLE NONLINEAR REGRESSION MODEL USING SIMULATED DATA FOR $\beta = \binom{.26}{-.02}$, TEN OBSERVATIONS AND λ TAKING VALUES 1, 10 AND 50, RESPECTIVELY

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-8}$)
Root .049 -.004	221,393,177	4543 26	5.69
MLE .047 -.004	118,570,847	4625 27	4.2
LSE 2478.8 -309.8	971,827.9	1.41×10^{12} 2.20×10^{10}	-2.52×10^{15}
MLSE .104 -.016	1,973,746,016	3600 31	506

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-8}$)
Root 0.174 -.014	3,363,855	1052 6	5.7
MLE 0.173 -.014	3,165,965	1058 6	5.3
LSE -2743812100* -236928480	25,558	10^{24}** 7.465×10^{21}**	-3.3×10^{39}
MLSE 0.162 -.019	81,526,070	1312 9	79.3

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-8}$)
Root .234 -.019	141,954	194 2	.6
MLE .234 -.019	141,954	194 2	.6
LSE .224 -.017	263.2	516 5	1[a]
MLSE .2 -.019	12,952,341	454 2	8.5

*Quality of $\hat{\beta}$ did not improve. **Increase in MSE. [a]Increase in Det(MSE).

TABLE V

A COMPARISON OF THE ESTIMATES OF THE MEAN OF THE SIMPLE NONLINEAR REGRESSION MODEL USING SIMULATED DATA FOR $\underset{\sim}{\beta} = \begin{pmatrix} .26 \\ -.02 \end{pmatrix}$, THIRTY OBSERVATIONS AND λ TAKING VALUES 1, 10 AND 50, RESPECTIVELY

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-9}$)
Root .042 -.004	16,448,973,395	4778 27	5
MLE .041 -.004	16,380,905,230	4803 27	3
LSE .029 -.002	2,173,276	5369 31	0.0
MLSE .05 -.007	173,051,815	4626 19	350

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-8}$)
Root 0.167 -.014	7,326,359	952 4	1.42
MLE .167 -.014	6,300,118	954 4	1.36
LSE .146 -.012	26,907.6	1552 9	3
MLSE .181 -.017	4,870,042	785 3	6.1

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-9}$)
Root 0.234 -.019	2432.6	107 1	1.5
MLE 0.234 -.019	2432.6	107 1	1.5
LSE .227 -.018	2024.6	332 3	10
MLSE .237 -.019	588,031	101 1	2.6

TABLE VI

A COMPARISON OF THE ESTIMATES OF THE MEAN OF THE SIMPLE NONLINEAR REGRESSION MODEL USING SIMULATED DATA FOR $\underset{\sim}{\beta} = \begin{pmatrix} .26 \\ -.02 \end{pmatrix}$, FIFTY OBSERVATIONS AND λ TAKING VALUES 1, 10 AND 50, RESPECTIVELY

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-9}$)
Root .04 -.004	893,852,767	4834 27	2.4
MLE .04 -.004	553,416,286	4847 27	1.9
LSE .026 -.002	4,630,292	5493 32	0.0
MLSE .042 -.005	2.8082×10^{11}	4806 22	87

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-9}$)
Root 0.164 -.014	61,498,369	967 4	8.73
MLE .164 -.014	58,491,743	967 4	8.67
LSE .143 -.012	57,461.1	1536 8	20
MLSE .174 -.016	34,520,448	839 3	35.4

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-9}$)
Root 0.234 -.019	4,482.7	95 0.4	1
MLE 0.234 -.019	4,482.7	95 0.4	1
LSE 0.228 -.018	4,196.2	257 2	0.0
MLSE .235 -.019	24,623,836	93 1	1.6

TABLE VII

A COMPARISON OF THE ESTIMATES OF THE MEAN OF THE SIMPLE NONLINEAR REGRESSION MODEL USING SIMULATED DATA FOR $\underset{\sim}{\beta} = \binom{-.02}{.06}$, TEN OBSERVATIONS AND λ TAKING VALUES 1, 10 AND 50, RESPECTIVELY

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-6}$)
Root -.036 .025	828,195	430 145	6
MLE -.028 .024	376,818	315 150	5
LSE -8324145150 -1733017600	40,784.8	2.275×10^{25} 9.8591×10^{23}	4.08×10^{10}
MLSE -.287 .049	705,796	14,981 181	197

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-7}$)
Root -.052 .054	33,832.7	528 22	8.4
MLE -.051 .054	13,143.3	520 22	8.2
LSE -150.1 50.1	765	3.0492×10^{9} 3.3878×10^{8}	2.4×10^{9}
MLSE -.238 .062	8,183,798*	7183 41	184

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-7}$)
Root -.033 .059	23.2	165 5	0.4
MLE -.033 .059	23.2	165 5	0.4
LSE -.036 .062	8.9	398 33	2
MLSE -.165 .060	885,197	2695 8	17

*Increase in ESS.

TABLE VIII

A COMPARISON OF THE ESTIMATES OF THE MEAN OF THE SIMPLE NONLINEAR REGRESSION MODEL USING SIMULATED DATA FOR $\underset{\sim}{\beta} = \binom{-.02}{.06}$, THIRTY OBSERVATIONS AND λ TAKING VALUES 1, 10 AND 50, RESPECTIVELY

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-7}$)
Root -.041 .023	5,344,713,881	85 139	7.4
MLE -.039 .023	5,020,637,223	66 142	5.6
LSE -43300.5 904.9	65279.7	9.026×10^{14} 3.942×10^{11}	3.908×10^{18}
MLSE -.131 .034	2,282,874,530	1863 91	137.6

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-7}$)
Root -.06 .054	1524.6	273 10	2.6
MLE -.06 .054	1524.6	273 10	2.6
LSE -.044 .050	852.3	360 39	10.0
MLSE -.083 .057	1,855,410	677 11	6.4

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE ($\times 10^{-8}$)
Root -.035 .059	101	72 2	1
MLE -.035 .059	101	72 2	1
LSE -.035 .059	88	168 11	5.0
MLSE -.039 .059	139	107 2.3	1.6

TABLE IX

A COMPARISON OF THE ESTIMATES OF THE MEAN OF THE SIMPLE NONLINEAR REGRESSION MODEL USING SIMULATED DATA FOR $\underset{\sim}{\beta} = \begin{pmatrix} -.02 \\ .06 \end{pmatrix}$, FIFTY OBSERVATIONS AND λ TAKING VALUES 1, 10 AND 50, RESPECTIVELY

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-7}$)
Root -.042 .023	92,458,683	69 138	3.6
MLE -.041 .023	31,092,887	58 140	2.8
LSE -.021 .016	119,092.4	29 198	5.7
MLSE -.107 .031	16,340,254	1034 95	58.4

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-7}$)
Root -.066 .055	2,433.6	273 6	1.6
MLE -.066 .055	2,433.6	273 6	1.6
LSE -.050 .050	1,857.7	253 27	6.5
MLSE -.083 .057	452,810	567 7	3.8

Estimate	ESS	MSE ($\times 10^{-5}$)	Det(MSE) ($\times 10^{-9}$)
Root -.038 .059	193.8	63 1	6.3
MLE -.038 .059	193.8	63 1	6.3
LSE -.037 .059	179.6	124 8	40
MLSE -.040 .059	230.8	85 2	9.2

true value.

The MLE, unlike the LSE, has a closed expression which gives estimates that lie in the parameter space.

Using the root method it is possible to obtain estimates which lie outside the parameter space. Although it is not yet known how likely it is for this to happen, the simulation studies reveal that it is quite rare. One nice thing about the root method is that it is asymptotically normal.

Even though the MLSE is unbiased, the results indicate that it is the least efficient of all.

For a known λ with increase in the number of observations one can see a reduction in the determinant of the MSE matrix. It hasn't been possible to detect the general trend in the other cases.

C. Classification Model

Let $Y_{ijk} \sim I(\mu_{ij},\lambda)$, $\quad i=1,\ldots,I;\ j=1,\ldots,J;\ k=1,\ldots,N$

and the Y_{ijk}'s are independent.

Fries and Bhattacharyya (1983) assumed the following model

$$\frac{1}{\mu_{ij}} = \mu + \alpha_i + \beta_j$$

where $\sum_{i=1}^{I} \alpha_i = \sum_{j=1}^{J} \beta_j = 0$ for testing additivity. Consider the following example for $I=2$ and $J=3$. If we let $\mu=10$, $\alpha_1=5$, $\beta_1=2$ and $\beta_2=1$ then based on the above constraint $\alpha_2=-5$ and $\beta_3=-3$. Thus the μ_{ij} values are

	J = 1	J = 2	J = 3
I = 1	$\frac{1}{17}$	$\frac{1}{16}$	$\frac{1}{12}$
I = 2	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{2}$

Since $(\mu_{12} - \mu_{11}) \neq (\mu_{22} - \mu_{21})$ then the above model is not an additive or interaction model. It seems reasonable to call it classification model.

Although they showed that for $I = J = 2$ the estimates of $\frac{1}{\mu_{ij}}$ lie in the parameter space, the following example demonstrates that this result is not true in general. Let

$$X = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \quad \text{and} \quad Y = (1,1,10,10,0.2,0.2)'.$$

Then $(\mu^*, \alpha_1^*, \beta_1^*, \beta_2^*) = (0.767344, -.31815, -.84583, 1.11142)$ and $\mu^* + \alpha_1^* + \beta_1^* = -0.396636$. It still remains to determine the conditions under which the estimates of μ_{ij} lie in the parameter space.

CHAPTER VI

MINIMAL SUFFICIENT STATISTIC

Our aim is to find

$$\{\underset{\sim}{y}: \frac{f(\underset{\sim}{y};\alpha,\beta,\underset{\sim}{\lambda})}{f(\underset{\sim}{y}^\circ;\alpha,\beta,\underset{\sim}{\lambda})} \text{ is independent of } (\alpha,\beta,\underset{\sim}{\lambda})\}.$$

A. Common λ

Let $Y_i \sim I(\alpha+\beta x_i,\lambda)$ $i=1,\ldots,n$ and the Y's are independent. Thus the joint density function of the Y's is

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y};\alpha,\beta,\lambda) = (\frac{\lambda}{2\pi})^{\frac{n}{2}} \prod_{i=1}^{n} y_i^{-\frac{3}{2}} \exp[-\frac{\lambda}{2} \sum_{i=1}^{n} \frac{(y_i-\alpha-\beta x_i)^2}{(\alpha+\beta x_i)^2 y_i}].$$

Thus

$$\frac{f(\underset{\sim}{y};\alpha,\beta,\lambda)}{f(\underset{\sim}{y}^\circ;\alpha,\beta,\lambda)} = \prod_{i=1}^{n} (\frac{y_i^\circ}{y_i})^{\frac{3}{2}} \exp[-\frac{\lambda}{2} \sum_{i=1}^{n} \{\frac{(y_i-\alpha-\beta x_i)^2}{(\alpha+\beta x_i)^2 y_i} - \frac{(y_i^\circ-\alpha-\beta x_i)^2}{(\alpha+\beta x_i)^2 y_i^\circ}\}]$$

$$= \prod_{i=1}^{n} (\frac{y_i^\circ}{y_i})^{\frac{3}{2}} \exp[-\frac{\lambda}{2} \sum_{i=1}^{n} \{\frac{y_i-y_i^\circ}{(\alpha+\beta x_i)^2} + \frac{1}{y_i} - \frac{1}{y_i^\circ}\}].$$

Suppose the ratio does not depend on α, β or λ. This is equivalent to the term in the exponent being independent of α, β or λ. If all the x's are distinct then $y_i = y_i^\circ$ for all $i=1,\ldots,n$. If some of the x's are equal, say $x_1 = x_2 = x_3$ then (y_1,y_2,y_3) is a permutation of $(y_1^\circ,y_2^\circ,y_3^\circ)$ and

$y_i = y_i^\circ$, $i = 4,\ldots,n$.

Consider $\{\underset{\sim}{y}: y_i = y_i^\circ, i = 1,\ldots,n\}$. Then the ratio does not depend on α, β or λ. Hence $(y_1, y_2, \ldots, y_n)$ is minimal sufficient for (α,β,λ).

B. Different λ

Given $Y_i \sim I(\alpha + \beta x_i, \lambda_i)$ $i = 1,\ldots,n$ and the Y's are independent.

Case (a): Assume $\dfrac{\alpha + \beta x_i}{\lambda_i} = k$ for all i.

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y};\alpha,\beta,\underset{\sim}{\lambda}) = (2\pi k)^{-\frac{n}{2}} \prod_{i=1}^{n} \sqrt{\frac{\alpha+\beta x_i}{y_i^3}} \exp[-\frac{1}{2k}\{\sum_{i=1}^{n} \frac{y_i}{\alpha+\beta x_i} - 2n + \alpha \sum_{i=1}^{n} \frac{1}{y_i} + \beta \sum_{i=1}^{n} \frac{x_i}{y_i}\}].$$

Thus

$$\frac{f(\underset{\sim}{y};\alpha,\beta,\underset{\sim}{\lambda})}{f(\underset{\sim}{y}^\circ;\alpha,\beta,\underset{\sim}{\lambda})} = \prod_{i=1}^{n} (\frac{y_i^\circ}{y_i})^{\frac{3}{2}} \exp[-\frac{1}{2k}\{\sum_{i=1}^{n} \frac{y_i - y_i^\circ}{\alpha+\beta x_i} + \alpha(\sum_i \frac{1}{y_i} - \sum_i \frac{1}{y_i^\circ}) + \beta(\sum_i \frac{x_i}{y_i} - \sum_i \frac{x_i}{y_i^\circ})\}].$$

Suppose the ratio does not depend on α, β or k. This is equivalent to the term in the exponent being independent of α, β or k. Assume all the x's are distinct. Then $y_i = y_i^\circ$ for all $i = 1,\ldots,n$,

$$\sum_i \frac{1}{y_i} - \sum_i \frac{1}{y_i^\circ} = 0 \quad \text{and} \quad \sum_i \frac{x_i}{y_i} - \sum_i \frac{x_i}{y_i^\circ} = 0.$$

$y_i = y_i^\circ$ for all $i = 1,\ldots,n$ implies $\sum_i \frac{1}{y_i} = \sum_i \frac{1}{y_i^\circ}$. Since the x's are known constants then

$\sum_i \frac{x_i}{y_i} = \sum_i \frac{x_i}{y_i^\circ}$. Thus it suffices to conclude that $y_i = y_i^\circ$ for all $i = 1,\ldots,n$. If some of the x's are equal, say $x_1 = x_2 = x_3$, then (y_1, y_2, y_3) is a permutation of $(y_1^\circ, y_2^\circ, y_3^\circ)$ and $y_i = y_i^\circ$, $i = 4,\ldots,n$.

Consider $\{\underset{\sim}{y}: y_i = y_i^\circ, i = 1,\ldots,n\}$. Then the ratio is independent of

α, β or k. Hence $(y_1, y_2, \ldots, y_n)$ is minimal sufficient for (α, β, k).

Case (b): Assume

$$\frac{(\alpha + \beta x_i)^2}{\lambda_i} = k \quad \text{for all } i$$

$$\frac{f(\underset{\sim}{y};\alpha,\beta,\underset{\sim}{\lambda})}{f(\underset{\sim}{y}^\circ;\alpha,\beta,\underset{\sim}{\lambda})} = \prod_{i=1}^{n} \left(\frac{y_i^\circ}{y_i}\right)^{\frac{3}{2}}$$

$$\exp\left[-\frac{1}{2k}\left\{\left(\sum_i y_i - \sum_i y_i^\circ\right) + \alpha^2\left(\sum_i \frac{1}{y_i} - \sum_i \frac{1}{y_i^\circ}\right) + \beta^2\left(\sum_i \frac{x_i^2}{y_i} - \sum_i \frac{x_i^2}{y_i^\circ}\right)\right.\right.$$

$$\left.\left. + 2\alpha\beta\left(\sum_i \frac{x_i}{y_i} - \sum_i \frac{x_i}{y_i^\circ}\right)\right\}\right].$$

Suppose this resulting ratio is independent of α, β or k. This is equivalent to saying that the term in the exponent does not depend on α, β or k. Thus

$$\sum y_i = \sum y_i^\circ, \quad \sum_i \frac{1}{y_i} = \sum_i \frac{1}{y_i^\circ}, \quad \sum_i \frac{x_i}{y_i} = \sum_i \frac{x_i}{y_i^\circ} \quad \text{and} \quad \sum_i \frac{x_i^2}{y_i} = \sum_i \frac{x_i^2}{y_i^\circ}.$$

Note that the converse is also true. Hence

$$\left(\sum_i y_i, \sum_i \frac{1}{y_i}, \sum_i \frac{x_i}{y_i}, \sum_i \frac{x_i^2}{y_i}\right)$$

is minimal sufficient for (α, β, k).

Case (c): Assume

$$\frac{(\alpha + \beta x_i)^3}{\lambda_i} = k \quad \text{for all } i$$

$$\frac{f(\underset{\sim}{y};\alpha,\beta,\underset{\sim}{\lambda})}{f(\underset{\sim}{y}^{\circ};\alpha,\beta,\underset{\sim}{\lambda})} = \prod_{i=1}^{n} \left(\frac{y_i^{\circ}}{y_i}\right)^{\frac{3}{2}} \frac{\exp[-\frac{1}{2k}\{\sum_i(\alpha+\beta x_i)y_i + \sum_i \frac{(\alpha+\beta x_i)^3}{y_i}\}]}{\exp[-\frac{1}{2k}\{\sum_i(\alpha+\beta x_i)y_i^{\circ} + \sum_i \frac{(\alpha+\beta x_i)^3}{y_i^{\circ}}\}]}$$

$$= \prod_{i=1}^{n} \left(\frac{y_i^{\circ}}{y_i}\right)^{\frac{3}{2}} \exp[-\frac{1}{2k}\{\alpha(\sum_i y_i - \sum_i y_i^{\circ}) + \beta(\sum_i x_i y_i - \sum_i x_i y_i^{\circ})$$

$$+ \alpha^3(\sum_i \frac{1}{y_i} - \sum_i \frac{1}{y_i^{\circ}}) + 3\alpha^2\beta(\sum_i \frac{x_i}{y_i} - \sum_i \frac{x_i}{y_i^{\circ}})$$

$$+ 3\alpha\beta^2(\sum_i \frac{x_i^2}{y_i} - \sum_i \frac{x_i^2}{y_i^{\circ}}) + \beta^3(\sum_i \frac{x_i^3}{y_i} - \sum_i \frac{x_i^3}{y_i^{\circ}})\}.$$

Assume that the ratio is independent of α, β or k. Then the term in the exponent does not depend on α, β or k. Thus

$$\sum_i y_i = \sum_i y_i^{\circ}, \quad \sum_i x_i y_i = \sum_i x_i y_i^{\circ}. \quad \sum_i \frac{1}{y_i} = \sum_i \frac{1}{y_i^{\circ}}, \quad \sum_i \frac{x_i}{y_i} = \sum_i \frac{x_i}{y_i^{\circ}},$$

$$\sum_i \frac{x_i^2}{y_i} = \sum_i \frac{x_i^2}{y_i^{\circ}} \quad \text{and} \quad \sum_i \frac{x_i^3}{y_i} = \sum_i \frac{x_i^3}{y_i^{\circ}}.$$

The converse is also true. Hence

$$(\sum_i y_i, \ \sum_i x_i y_i, \ \sum_i \frac{1}{y_i}, \ \sum_i \frac{x_i}{y_i}, \ \sum_i \frac{x_i^2}{y_i}, \ \sum_i \frac{x_i^3}{y_i})$$

is minimal sufficient for (α,β,k).

Case (d): Assume

$$\frac{(\alpha+\beta x_i)^2 x_i}{\lambda_i} = k \quad \text{for all } i.$$

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y};\alpha,\beta,\underset{\sim}{\lambda}) = (2\pi k)^{-\frac{n}{2}} \prod_{i=1}^{n} [(\alpha+\beta x_i)\sqrt{\frac{x_i}{y_i^3}}]$$

$$\exp[-\frac{1}{2k}\{\sum_i x_i y_i - 2\sum_i (\alpha+\beta x_i)x_i + \sum_i \frac{(\alpha+\beta x_i)^2 x_i}{y_i}\}]$$

$$\frac{f(\underset{\sim}{y};\alpha,\beta,\underset{\sim}{\lambda})}{f(\underset{\sim}{y}^\circ;\alpha,\beta,\underset{\sim}{\lambda})} = \prod_{i=1}^{n} (\frac{y_i^\circ}{y_i})^{\frac{3}{2}}$$

$$\exp[-\frac{1}{2k}\{\sum_i x_i y_i - \sum_i x_i y_i^\circ) + \alpha^2(\sum_i \frac{x_i}{y_i} - \sum_i \frac{x_i}{y_i^\circ})$$

$$+ 2\alpha\beta(\sum_i \frac{x_i^2}{y_i} - \sum_i \frac{x_i^2}{y_i^\circ}) + \beta^2(\sum_i \frac{x_i^3}{y_i} - \sum_i \frac{x_i^3}{y_i^\circ})\}].$$

If the ratio is independent of α, β or k then the term in the exponent is also independent of α, β or k. Thus

$$\sum_i x_i y_i = \sum_i x_i y_i^\circ, \quad \sum_i \frac{x_i}{y_i} = \sum_i \frac{x_i}{y_i^\circ}, \quad \sum_i \frac{x_i^2}{y_i} = \sum_i \frac{x_i^2}{y_i^\circ} \quad \text{and} \quad \sum_i \frac{x_i^3}{y_i} = \sum \frac{x_i^3}{y_i^\circ}.$$

The converse also holds. Hence

$$(\sum_i x_i y_i, \quad \sum_i \frac{x_i}{y_i}, \quad \sum_i \frac{x_i^2}{y_i}, \quad \sum_i \frac{x_i^3}{y_i})$$

is minimal sufficient for (α,β,k).

CHAPTER VII

TRIALS OF THESE MODELS USING REAL DATA AND SIMULATED DATA

In this chapter a summary of the rsults obtained using simulated and real data is given. In the tables and figures that follow, common λ refers to $Y_i \sim I(\beta x_i, \lambda)$ while different λ's refers to $Y_i \sim I(\beta x_i, \lambda_i)$ where $(\beta^2 x_i^2)/\lambda_i = k$ for all i. For simple (non-zero intercept) regression models replace βx_i by $\alpha + \beta x_i$.

A. Simulated Data

1. From the results given in Table X one can conclude that if the number of observations is fixed but λ is increasing, then

   a. the point estimate of a future observation y gets closer to the true value and

   b. the PI of a future observation y gets narrower.

2. If we look at (a) of Table XI, for the fixed λ case, the point estimate of x_0 (its true value is 8) is not contained in the three interval estimates. The interval estimates are unbounded too. But in (c) the point estimates for x_0 seem to be reasonable and they are contained in all the three CI's. Furthermore, each interval estimate of x_0 gets narrower as the number of observations increase. For the case where

TABLE X

POINT AND INTERVAL ESTIMATE OF A PREDICTED VALUE
y AT $x_* = 8.5$ USING SIMULATED DATA FOR $\beta = 2$ AND
λ TAKING VALUES ONE, THIRTY AND FIFTY

a) ten observations

	$\lambda = 1$		$\lambda = 30$		$\lambda = 50$	
	Predicted Value	95% PI	Predicted Value	95% PI	Predicted Value	95% PI
fixed λ	100.3	$(3.24,\infty)$	17.53	(5.4,62.41)	17.15	(6.79,45.69)
different λ's	167.27	$(6.79,\infty)$	23	(8.17,75.88)	21.12	(9.36,52.11)

b) fifty observations

	$\lambda = 1$		$\lambda = 30$		$\lambda = 50$	
	Predicted Value	95% PI	Predicted Value	95% PI	Predicted Value	95% PI
fixed λ	211.5	(3.94,28533.2)	24.99	(6.07,105.86)	21.83	(6.86,70.69)
different λ's	352.75	(10,90504.7)	32.16	(10.64,100.78)	26.47	(10.38,69.16)

c) two hundred observations

	$\lambda = 1$		$\lambda = 30$		$\lambda = 50$	
	Predicted Value	95% PI	Predicted Value	95% PI	Predicted Value	95% PI
fixed λ	181.27	(4.18,8934.27)	23	(6.22,85.53)	20.12	(6.94,58.53)
different λ's	264.72	(9.06,9410.31)	25.77	(8.57,78.17)	21.91	(9.03,53.45)

TABLE XI

POINT AND INTERVAL ESTIMATE OF x_0 USING SIMULATED DATA FOR $\beta = 2$ AND λ TAKING VALUES ONE, THIRTY AND FIFTY

a) $\lambda = 1$	ten observations		fifty observations		two hundred observations	
	Point Estimate	95% CI on x_0	Point Estimate	95% CI on x_0	Point Estimate	95% CI on x_0
fixed λ	0.46	$(1.45,\infty)$	16.33	$(50.57,\infty)$	15.52	$(186.66,\infty)$
different λ's	0.26	(0,5.42)	9.47	(0,33.63)	10.42	(0,21.33)

b) $\lambda = 30$	ten observations		fifty observations		two hundred observations	
	Point Estimate	95% CI on x_0	Point Estimate	95% CI on x_0	Point Estimate	95% CI on x_0
fixed λ	6.31	$(62.57,\infty)$	11.06	(0,164.24)	10.18	(0,32.26)
different λ's	4.62	(0,11.76)	8.89	(4.15,13.63)	9.08	(6.72,11.45)

c) $\lambda = 50$	ten observations		fifty observations		two hundred observations	
	Point Estimate	95% CI on x_0	Point Estimate	95% CI on x_0	Point Estimate	95% CI on x_0
fixed λ	6.76	(0,175.47)	11.03	(0,73.71)	10.06	(0,25.2)
different λ's	5.32	(0,11.23)	9.03	(5.03,13.04)	9.27	(7.37,11.18)

the ratio between the variance and the mean is constant all the interval estimates contain the estimates of x_0 and their widths are smaller. It seems that one can get a sharper result for a large λ with many observations.

Suppose the number of observations is fixed. As λ gets larger the interval estimate of x_0 is consistently narrower for different λ's than the fixed λ case. However, both cases give a reasonable point and interval estimate of x_0 for large λ and many observations.

B. Real Data

1. Folks and Chhikara (1978) fitted the IG distribution to the data of precipitation from Jug Bridge, Maryland. For those data, the MLE of μ^3/λ is 1.247 while its UMVUE is 1.261.

2. They also fitted the IG distribution to the run-off amounts for the same place. Based on those data, the MLE of the variance is 0.356 while its UMVUE is 0.349.

3. Snedecor and Cochran (1967) describe number of acres in corn (response variable) and size of farm in acres (independent variable) on 25 farms. The results they obtained using weighted least squares on

$$y_i = \beta x_i + \varepsilon_i, \quad \varepsilon_i \sim n(0,\sigma^2 x_i^2)$$

are presented in Table XII with those of the IG.

In Table XII, the interval estimate of β and the mean of y_i are fairly close. Although the bounds of the prediction intervals of the normal and the IG with common λ are different, their widths are almost equal. However, the PI for the IG with different λ's differ considerably.

TABLE XII

POINT AND INTERVAL ESTIMATE OF THE SLOPE, THE MEAN OF THE RESPONSE VARIABLE AND A PREDICTION INTERVAL USING REAL DATA

		Estimate of β		95% CI on the Mean of y_i at $x_* = 100$	95% PI on a Future Observation at $x_* = 100$
		Point Estimate	95% CI		
Normal (weighted Least Squares)		0.243	(0.206,0.28)	(20.6,28)	(5.4,43.2)
IG	Common λ	0.253	(0.215,0.308)	(21.5,30.8)	12.98,51.07)
	Different λ's	0.235	(0.197,0.29)	(19.7,29)	(5.89,96.97)

In Figure 4 the graph of the fitted line for common λ slightly overestimates the mean of y_i for large x's than the other fitted lines.

In Figure 5 one can observe that the prediction bands about the weighted least squares prediction line are narrower in width than the rest. For more information one should refer to Figures 1, 2 and 3.

Point and interval estimates of x_0 are also obtained. This is done by treating $x = 320$ as unknown and excluding the five y's at $x = 320$ in calculating β. For the common λ case $\tilde{x}_0 = 256.53$ and a 95% CI on x_0 is (0,1671.98) while for different λ's, $\tilde{x}_0 = 277.19$ and a 95% CI on x_0 is (144.9,383.76).

4. Montgomery and Peck (1982) give data on energy usage (independent variable) and demand (dependent variable) for 53 residential customers. They found that the least squares line was not a good fit. Thus they transformed the response variable by taking its square root. However, interpretation of the analysis of this transformed variable is not that easy.

In Figure 6, the line

$$\hat{y}_i = 0.1645 + 0.00282\, x_i,$$

where the slope is

$$\frac{1}{n} \sum_{i=1}^{n} \frac{y_i - \bar{y}}{x_i - \bar{x}}, \qquad x_i \neq \bar{x} \text{ for all } i,$$

and the intercept is $\bar{y}$-(slope) $\bar{x}$, is fitted to the original data. In addition a 95% CI on the slope for common λ and different λ's are (-0.0965,0.1022) and (-0.0262,0.03185), respectively.

In figure 7 one can observe that the prediction bands for different λ's are narrower than the common λ.

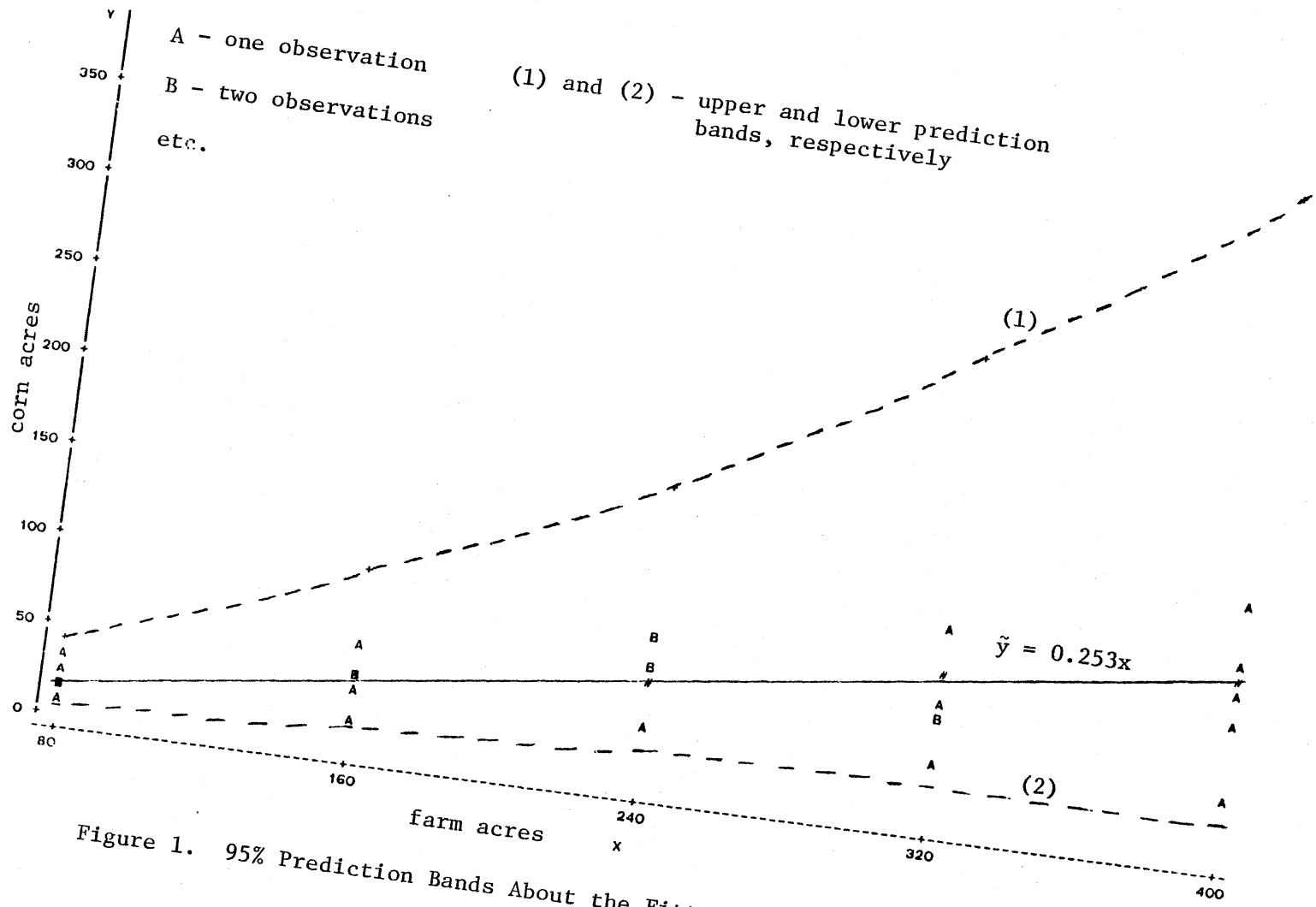

Figure 1. 95% Prediction Bands About the Fitted Line $\tilde{y} = 0.253\,x$ for Common λ

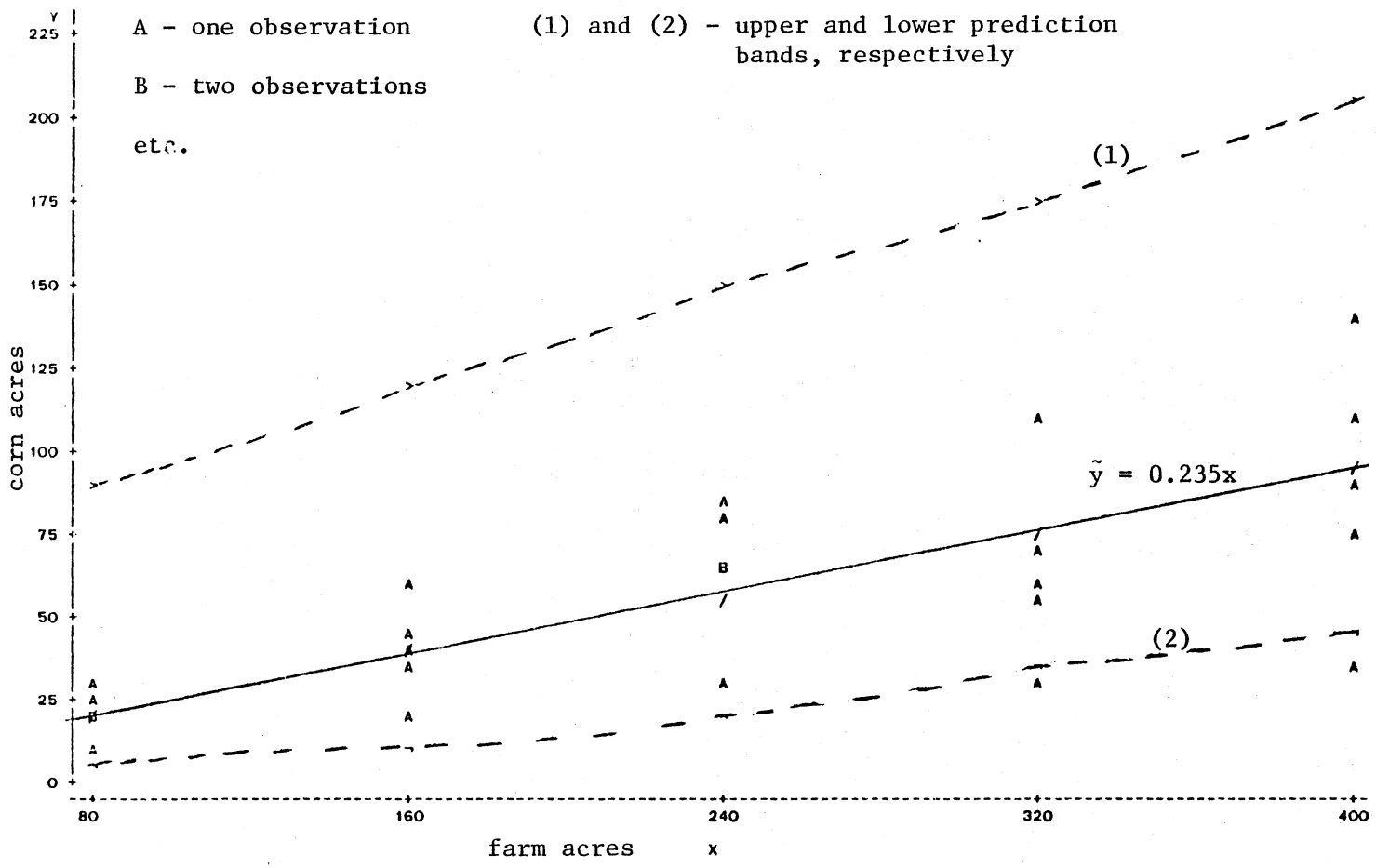

Figure 2. 95% Prediction Bands About the Fitted Line $\tilde{y} = 0.235\,x$ for Different λ's

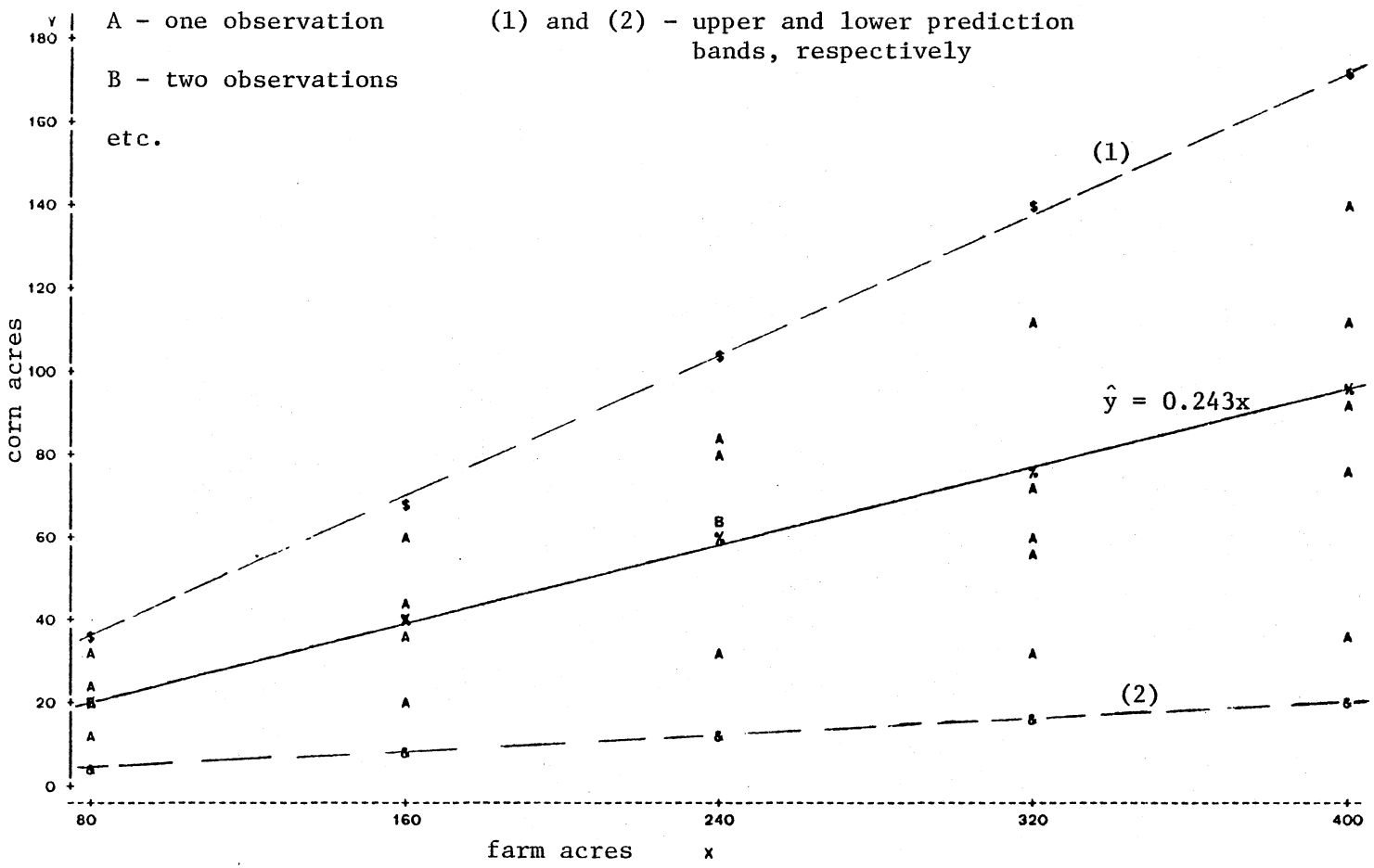

Figure 3. 95% Prediction Bands About the Weighted Least Squares Line $\hat{y} = 0.243\,x$ Assuming Normality

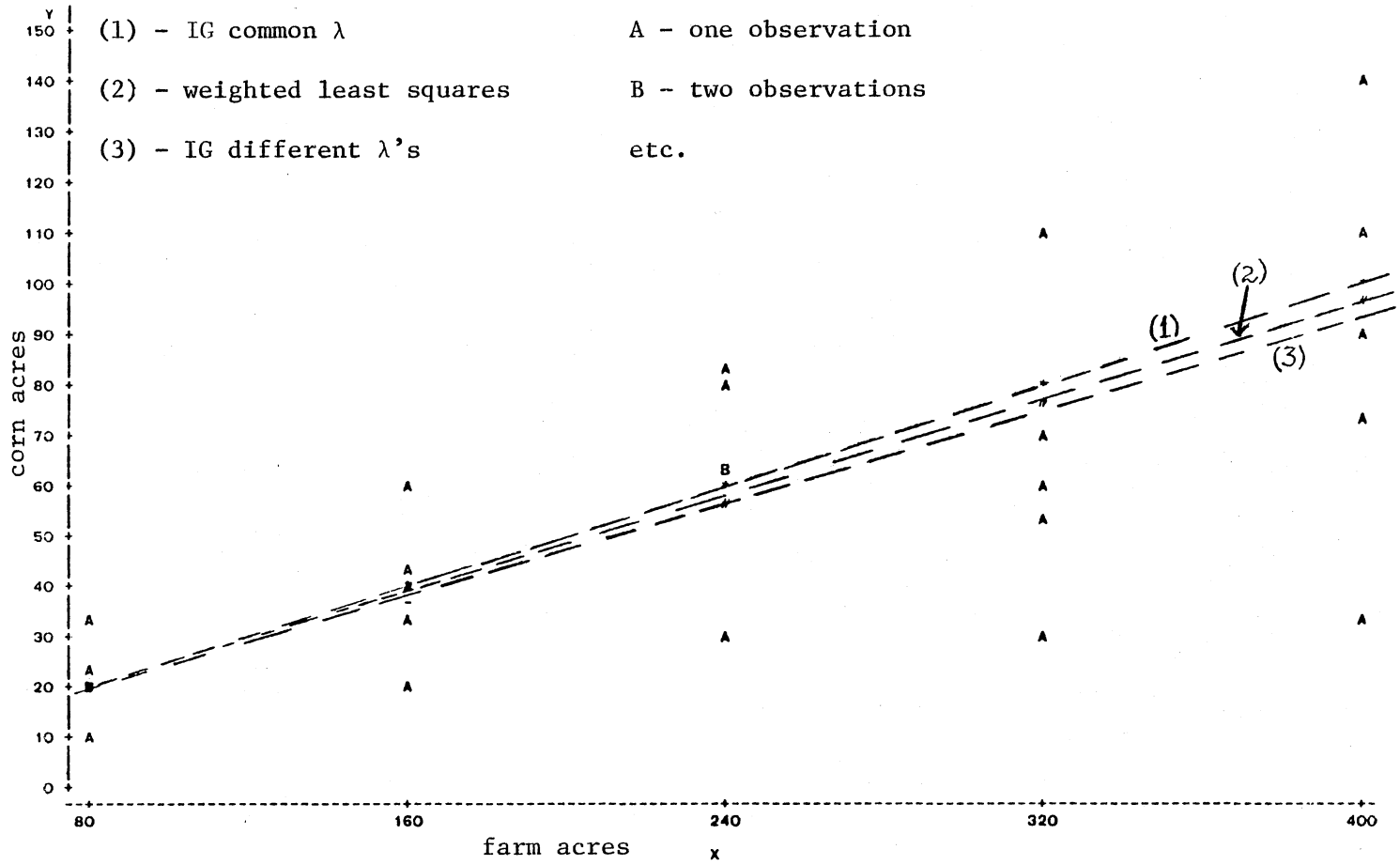

Figure 4. Graphs of Three Fitted Lines

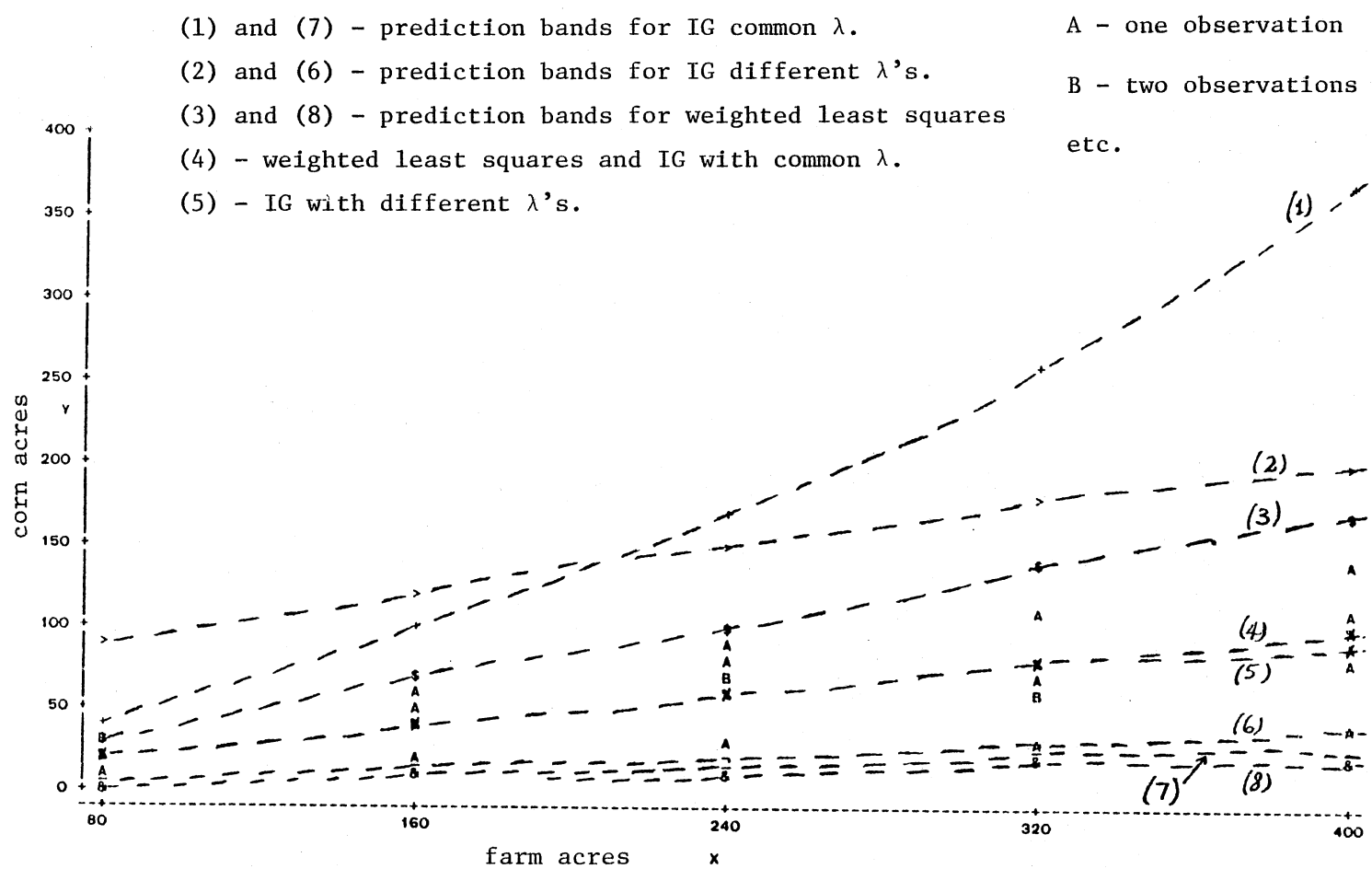

Figure 5. Graphs of Three Fitted Lines and the Corresponding 95% Prediction Bands

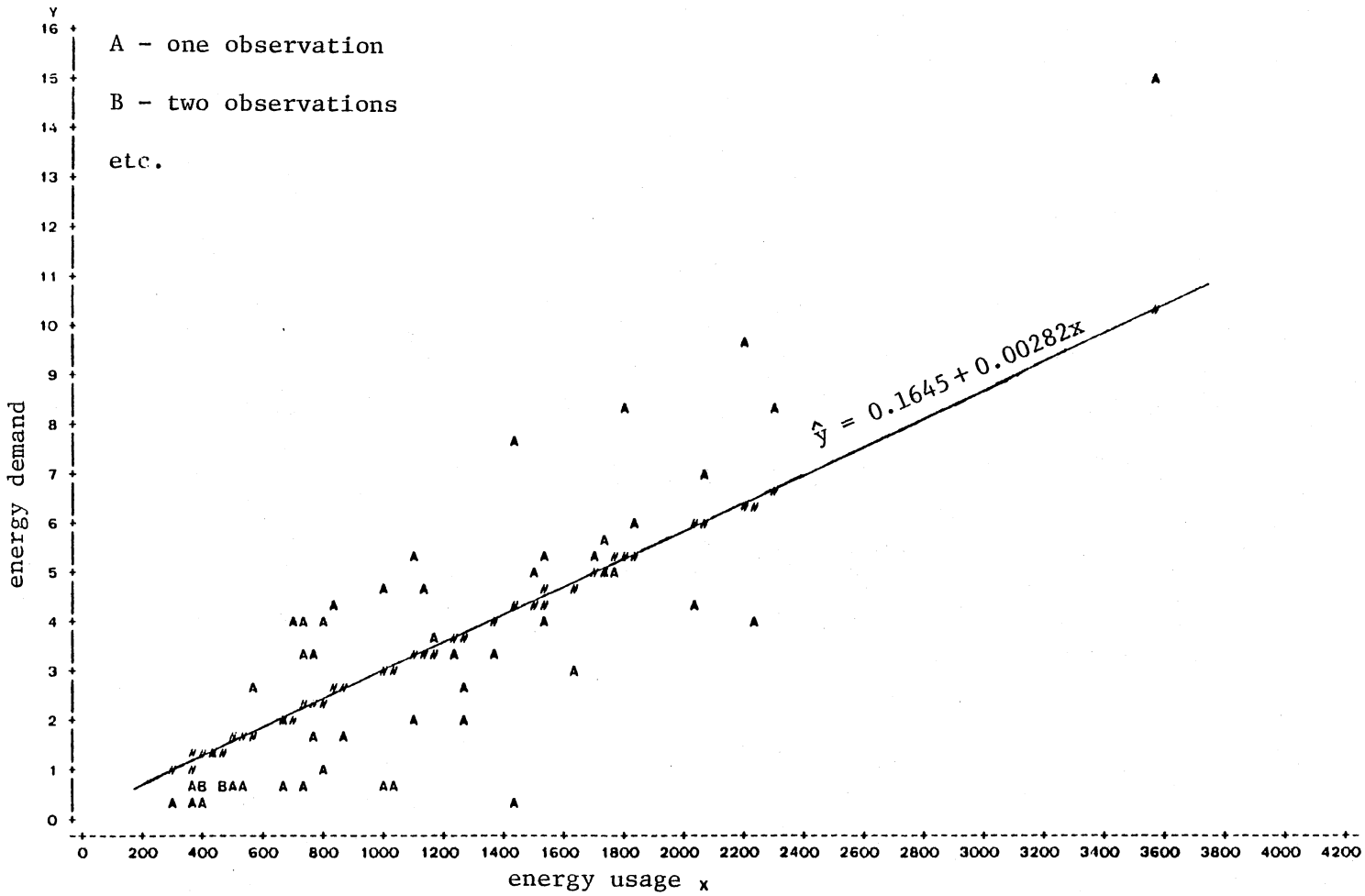

Figure 6. Graph of the Fitted Line $\hat{y} = 0.1645 + 0.00282\,x$ Which is Obtained Using the Normal Approximation

'#' - $\hat{y} = 0.1645 + 0.00282x$

'+' and '¬' - prediction bands for IG common λ.

'$' and '&' - prediction bands for IG different λ's.

A - one observation

B - two observations

etc.

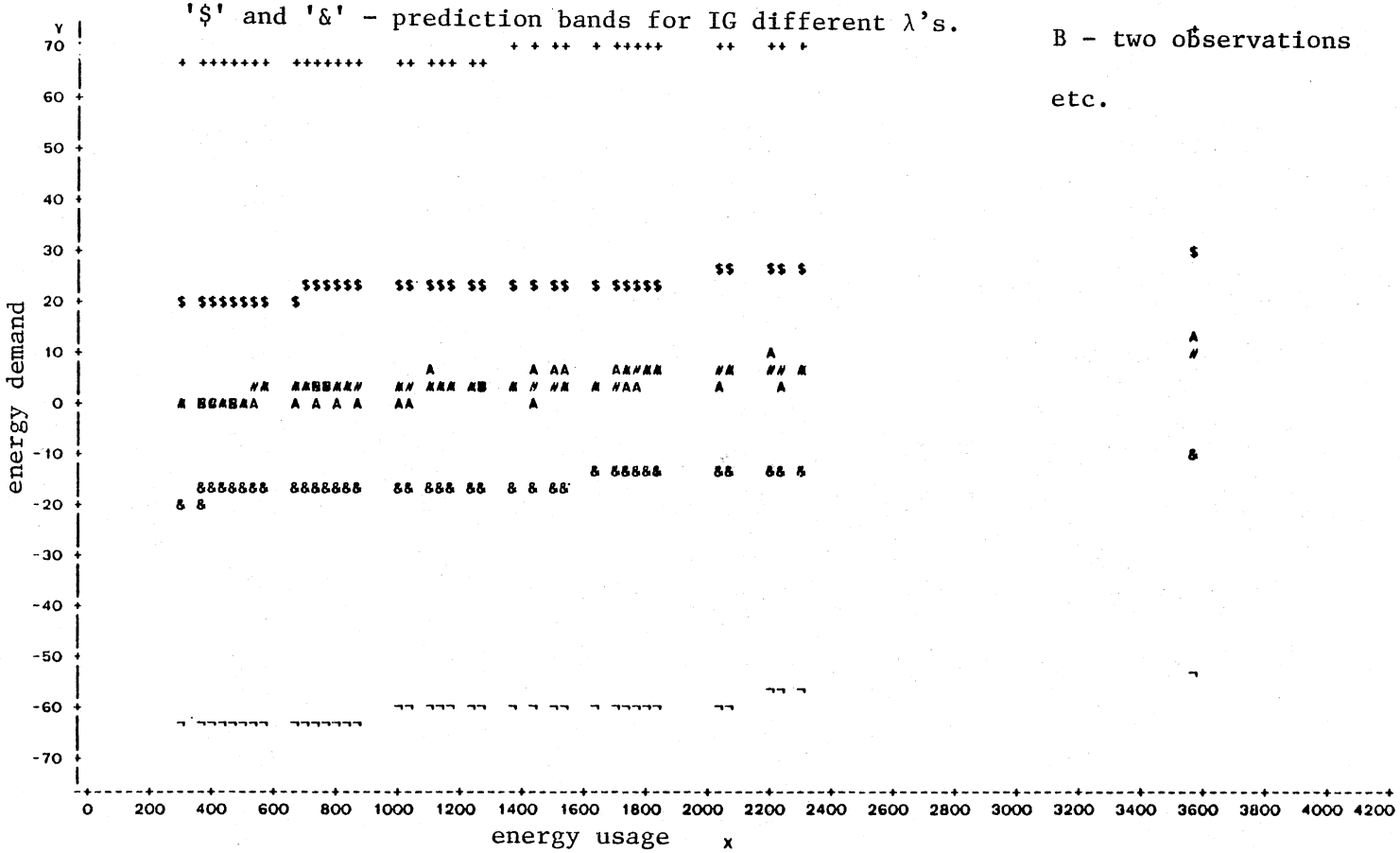

Figure 7. 95% Prediction Bands About the Line $\hat{y} = 0.1645 + 0.00282\,x$

5. The weighted least square estimates obtained using the data from Draper and Smith (1981) along with the asymptotic results of the simple non-zero intercept models are given.

TABLE XIII

FITTED LINES AND INTERVAL ESTIMATES OF THE SLOPE USING REAL DATA

		Equation of Fitted Line	95% CI on β
Normal (Weighted Least Squares)		$\hat{y}_i = 1.264 + 0.925\, x_i$	(0.5604,1.289)
IG	Common λ	$\hat{y}_i = -2.448 + 1.39\, x_i$	(-2.6292,5.4091)
	Different λ's	Same as common λ	(0.8868,1.8931)

From Figure 8 one can observe that the IG fitted line underestimates the response variable for small values of the independent variable while it slightly overestimates for large values of x. The reverse is true for the weighted least squares line.

In Figure 11 one can notice that the prediction bands for the IG with different λ's are narrower than the rest. For more information one should refer to Figures 9 and 10.

Although it is hard to tell in this case which of the lines gives a better fit, the preceding examples demonstrate clearly that the IG distribution can indeed serve as an alternative to the normal in some situations.

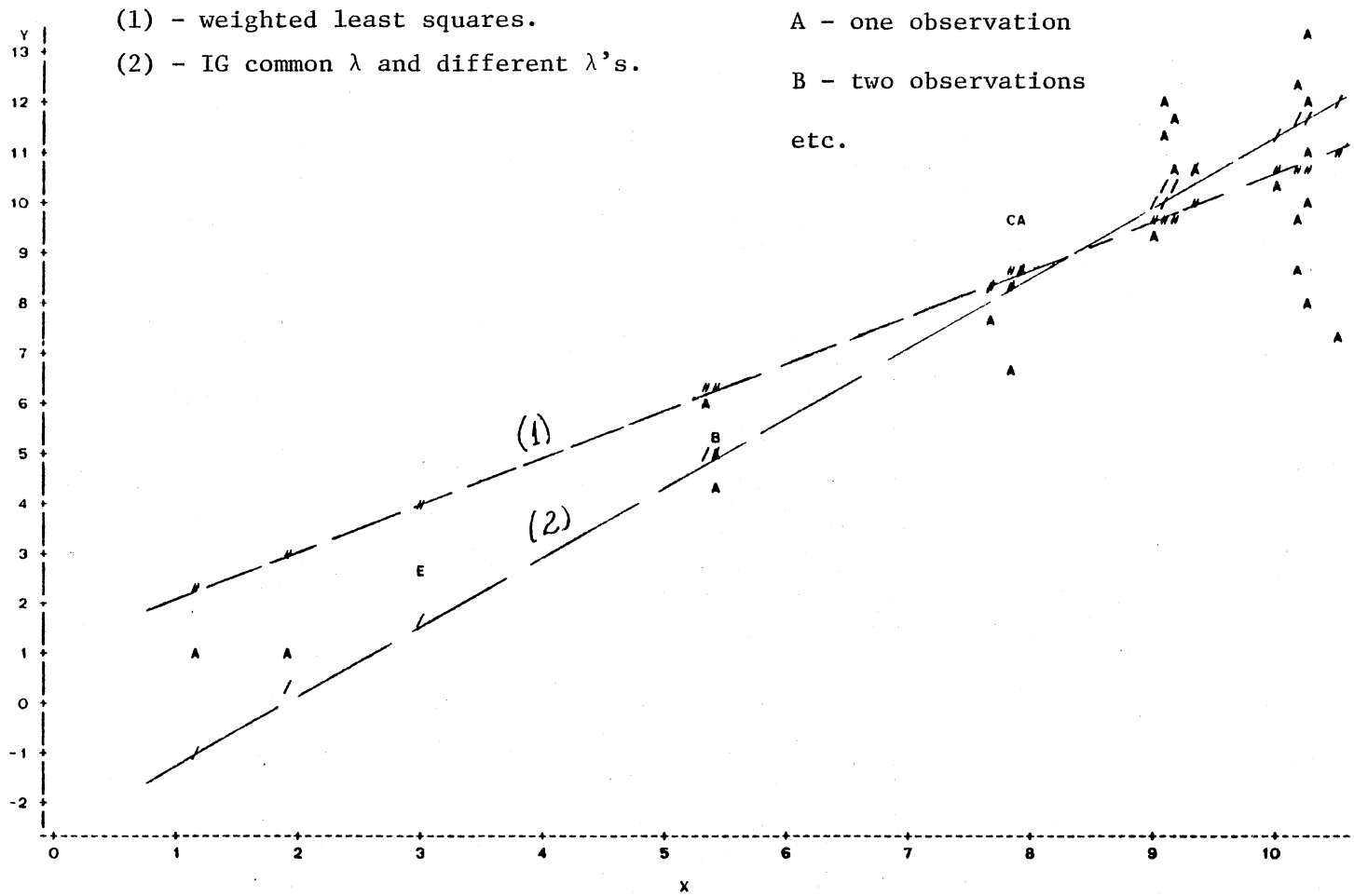

Figure 8. Graphs of a Weighted Least Squares Line and Another One Obtained Using Asymptotics

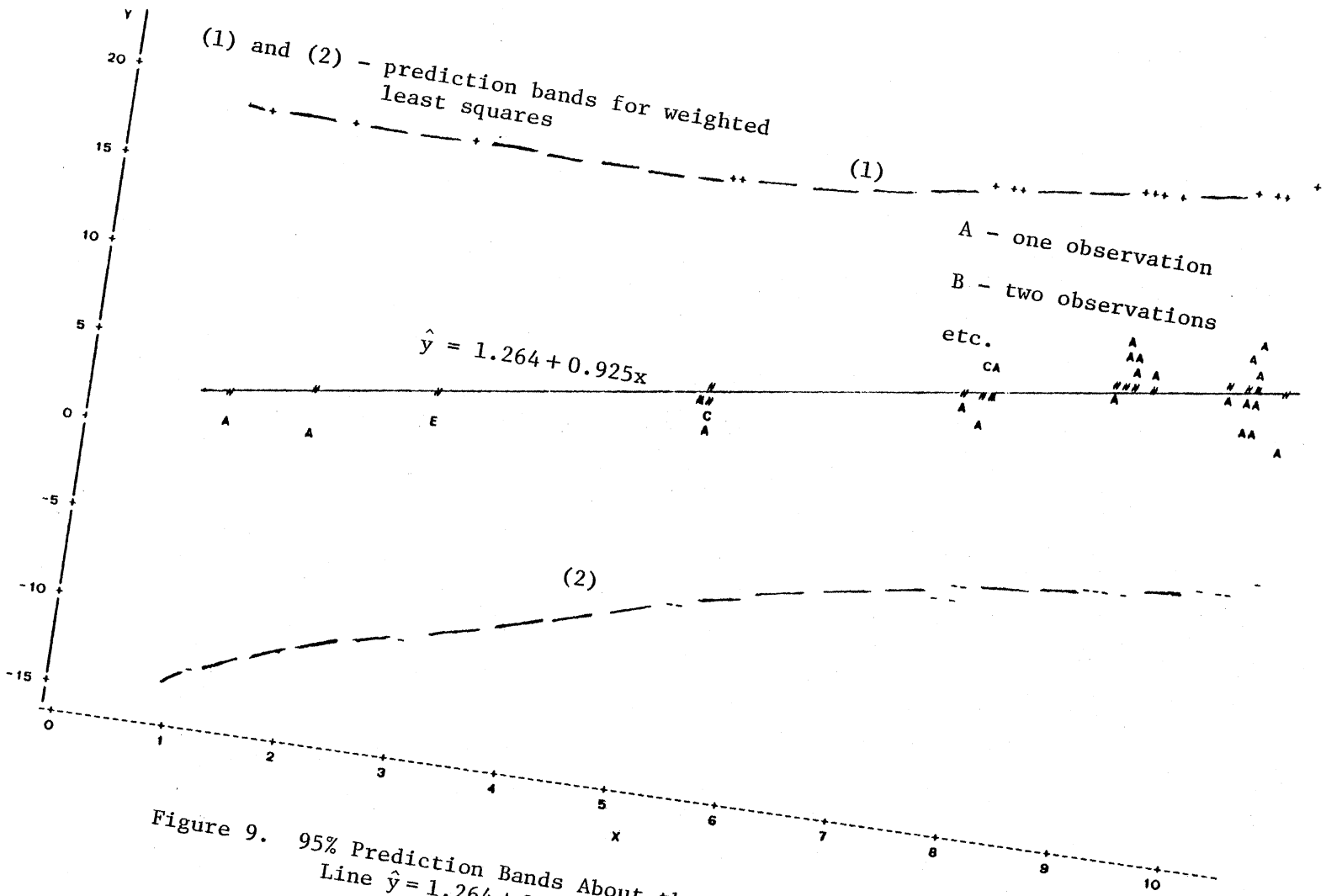

Figure 9. 95% Prediction Bands About the Weighted Least Squares Line $\hat{y} = 1.264 + 0.925\,x$

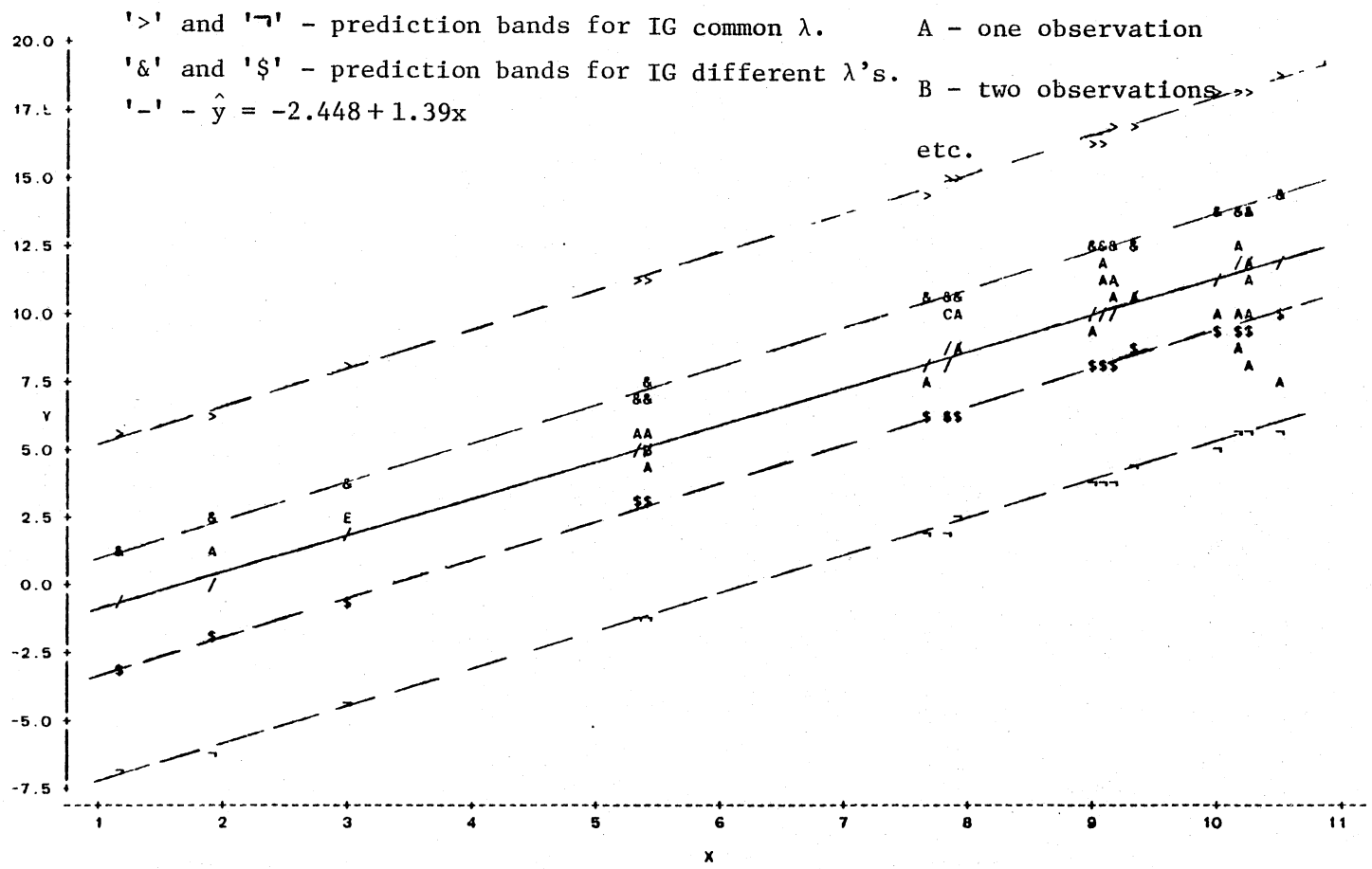

Figure 10. 95% Prediction Bands About the Line $\hat{y} = -2.448 + 1.39\,x$

(1) - weighted least squares.
(2) - IG common λ and different λ's.
'+' and '_' - prediction bands for weighted least squares.
'>' and '¬' - prediction bands for IG common λ.
'&' and '$' - prediction bands for IG different λ's.

A - one observation

B - two observations

etc.

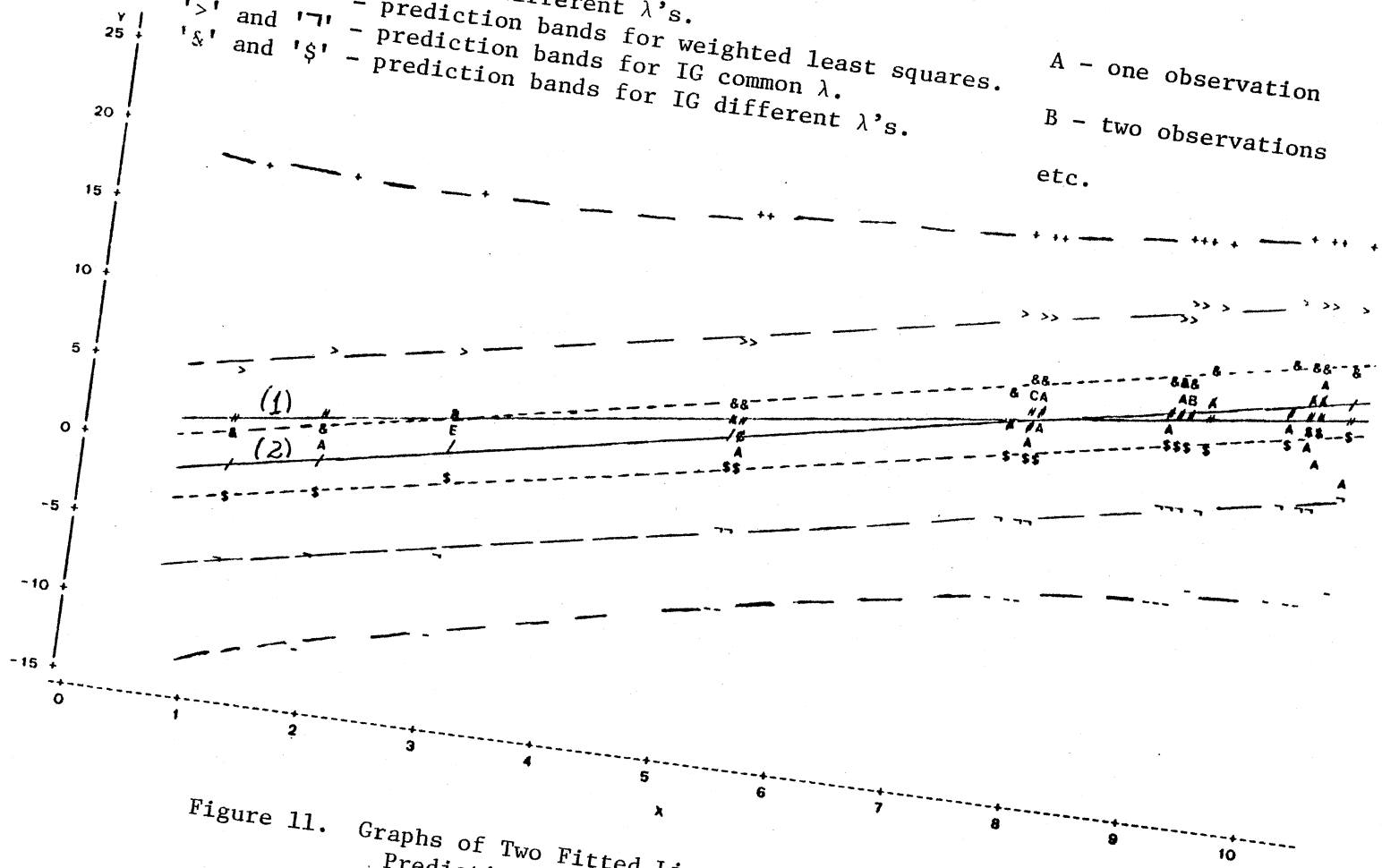

Figure 11. Graphs of Two Fitted Lines and the Corresponding 95% Prediction Bands

CHAPTER VIII

OUTLINE

1. Results for the regression models

   a. $Y_i \sim I(\beta x_i, \lambda)$

   b. $Y_i \sim I(\beta x_i, \lambda_i)$ where $\frac{\beta^2 x_i^2}{\lambda_i} = k$ for all i are

      i. UMVUE of $Var(Y_i)$

      ii. Power

      iii. Confidence Interval on $E(Y_i)$

      iv. Prediction Interval on Y and

      v. Confidence Interval on x.

2. Results for special cases of the regression models

   a. $Y_i \sim I(\alpha + \beta x_i, \lambda)$

   b. $Y_i \sim I(\frac{1}{\alpha + \beta x_i}, \lambda)$ are

      i. UMVUE of $Var(Y_i)$

      ii. Confidence Interval on $\alpha + \beta x_i$

      iii. Prediction Interval on Y and

      iv. Confidence Interval on x.

3. Asymptotic results for

   a. $Y_i \sim I(\alpha + \beta x_i, \lambda)$

   b. $Y_i \sim I(\alpha + \beta x_i, \lambda_i)$ where $(\alpha + \beta x_i)^2/\lambda_i = k$ for all i

   c. $Y_i \sim I(\frac{1}{\underset{\sim}{x}_i' \underset{\sim}{\beta}}, \lambda)$.

4. Estimates of α, β, and k where $Y_i \sim I(\alpha + \beta x_i, \lambda_i)$ subject to $(\alpha + \beta x_i)^2/\lambda_i = k$ for all i.

5. For the simple nonlinear regression model, i.e.

$$Y_i \sim I(\frac{1}{\alpha + \beta x_i}, \lambda)$$

   a. $\alpha^* + \beta^* x_i$ does not necessarily lie in the parameter space for some x_i used in generating the data

   b. The MLSE and MLE are given

   c. A comparison of the different methods of estimating α and β using the root of the likelihood, the MLE, the LSE and the MLSE are given

   d. One can test $\beta = 0$ for a special case.

6. Let

$$Y \sim (1 + h(y))\frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2})$$

where h(y) is any odd function and $|h(y)| \leq 1$ then $Y^2 \sim \chi_1^2$.

7. Minimal sufficient statistic for

$$Y_i \sim I(\alpha + \beta x_i, \lambda_i)$$

   a. Common λ

   b. Different λ

      i. $\frac{(\alpha + \beta x_i)^2}{\lambda_i} = k$ for all i

      ii. $\frac{\alpha + \beta x_i}{\lambda_i} = k$ for all i

      iii. $\frac{(\alpha + \beta x_i)^3}{\lambda_i} = k$ for all i

iv. $\dfrac{(\alpha + \beta x_i)^2 x_i}{\lambda_i} = \xi$ for all i.

8. Trials of these models on real data and simulated data.

BIBLIOGRAPHY

1. Abramowitz, M. and Stegun, I. A. (1964) Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. U. S. Department of Commerce, National Bureau of Standards, Applied Mathematics Series, 55.

2. Banerjee, A. K. and Bhattacharyya, G. K. (1976) "A purchase incidence model with inverse Gaussian interpurchase times". J. Amer. Stat. Assoc., 71, 823-829.

3. Bradsley, W. E. (1980) "Note on the use of the inverse Gaussian distribution for wind energy applications". J. Applied Meteorology, 19, 1126-1130.

4. Chhikara, R. S. (1972) Statistical inference related to the inverse Gaussian distribution. Unpublished Ph.D. Dissertation, Oklahoma State University.

5. Chhikara, R. S. (1975) "Optimum tests for the comparison of two inverse Gaussian distribution means". Austral. J. Stat., 17, 77-83.

6. Chhikara, R. S. and Folks, J. L. (1974) "Estimation of the inverse Gaussian distribution function". J. Amer. Stat. Assoc., 69, 250-254.

7. Chhikara, R. S. and Folks, J. L. (1975) "Statistical distribution related to the inverse Gaussian". Commun. in Statistics, 4, 1081-1091.

8. Chhikara, R. S. and Folks, J. L. (1976) "Optimum tests procedures for the mean of first passage time in Brownian motion with positive drift (inverse Gaussian)". Technometrics, 18, 189-193.

9. Chhikara, R. S. and Folks, J. L. (1977) "The inverse Gaussian distribution as a life time model". Technometrics, 19, 461-468.

10. Chhikara, R. S. and Guttman, Irwin. (1982) "Prediction limits for the inverse Gaussian distribution". Technometrics, 24, 319-324.

11. Cox, D. R. and Hinkley, D. V. (1974) Theoretical Statistics. London: Chapman and Hall.

12. Cox, D. R. and Miller, H. D. (1965) The Theory of Stochastic Processes. London: Methuen.

13. Davis, A. S. (1977) Linear statistical inference as related to the inverse Gaussian distribution. Unpublished Ph.D. Dissertation, Oklahoma State University.

14. Draper, N. R. and Smith, H. (1981) Applied Regression Analysis. New York: John Wiley.

15. Folks, J. L. and Chhikara, R. S. (1978) "The inverse Gaussian distribution and its statistical application - A review". J. Royal Stat. Society, Series B, 40, 263-289.

16. Folks, J. L. and Chhikara, R. S. (1981) The Inverse Gaussian Distribution. Unpublished monograph.

17. Folks, J. L. and Davis, A. S. (1980) "Regression Models for the Inverse Gaussian Distribution". Statistical Distributions in Scientific Work, 4, 91-97, Dorcrecht: D. Reidel.

18. Folks, J. L. and Doan, D. J. (1976) "The use of regression equations for demonstrating reliability". The Relia-com Review, 1, 1-3.

19. Fries, A. and Bhattacharyya, G. K. (1983) "Analysis of two factor life tests under an inverse Gaussian model". J. Amer. Stat. Assoc., 78, 820-826.

20. Graybill, F. A. (1976) Theory and Application of the Linear Model. North Scitnate: Duxbury Press.

21. Hasofer, A. M. (1964) "A dam with inverse Gaussian input". Proc. Camb. Phil. Soc., 60, 931-933.

22. Iwase, K. and Noriaki, S. (1983) "Uniformly minimum variance unbiased estimation for the inverse Gaussian distribution". J. Amer. Stat. Assoc., 78, 660-663.

23. Johnson, N. L. and Kotz, S. (1970) Distributions in Statistics: Univariate Distributions 1. Boston: Houghton-Mifflin.

24. Korwar, R. M. (1980) "On the uniformly minimum variance unbiased estimations of the variance and its reciprocal of an inverse Gaussian distribution". J. Amer. Stat. Assoc., 75, 734-735.

25. Lancaster, A. (1972) "A stochastic model for the duration of a strike". J. Royal Stat. Society, A, 135, 257-271.

26. Lehmann, E. L. and Scheffe, H. (1950) "Completeness, similar regions and unbiased estimation". Sankhyā, 10, 305-340.

27. Michael, J. R., Schucany, W. R. and Haas, R. W. (1976) "Generating random variables using transformations with multiple roots". American Statist., 30, 88-90.

28. Montgomery, D. C. and Peck, E. A. (1982) Introduction to Linear Regression Analysis. New York: John Wiley.

29. Rao, C. R. (1973) Linear Statistical Inference and Its Applications. New York: John Wiley.

30. Roy, L. K. and Wasan, M. T. (1968) "The first passage time distribution of Brownian Motion with positive drift". Math Bioscience, 3, 191-204.

31. Schrödinger, E. (1915) "Zur Theorie der Fallund Steigversuche an Teilchen mit Brownschen Bewegung". Physikalische Zeitschrift, 16, 289-295.

32. Seshadri, V. (1981). "A note on the inverse Gaussian distribution". Statistical Distributions in Scientific Work, 4, 99-103, Dordrecht: D. Reidel.

33. Shuster, J. J. (1968) "On the inverse Gaussian distribution function". J. Amer. Stat. Assoc., 63, 1514-1516.

34. Shuster, J. J. and Muira, C. (1972) "Two way analysis of reciprocals". Biometrics, 59, 478-481.

35. Smoluchowsky, M. V. (1915) "Notiz über die Berechrung der Brownschen Molekular-bewegung bei der Ehrenhaft-Millikanschen Versuchsanordnung". Phy. Z., 16, 318-321.

36. Snedecor, G. W. and Cochran, W. G. (1967) Statistical Methods. Ames: Iowa State University Press.

37. Tweedie, M. C. K. (1957) "Statistical properties of inverse Gaussian distributions I". Annals of Math. Stat., 28, 362-377.

38. Tweedie, M. C. K. (1957) "Statistical properties of inverse Gaussian distributions II". Annals of Math. Stat., 28, 696-705.

39. Wald, A. (1947) Sequential Analysis. New York: John Wiley.

40. Wasan, M. T. (1969) "First passage time distribution of Brownian Motion with positive drift (inverse Gaussian distribution)". Queen's Paper in Pure and Applied Mathematics, 19, Queen's University, Canada.

41. Whitmore, G. A. (1980) "An exact multiple regression model for inverse Gaussian data". McGill University.

42. Zigangirov, K. S. (1962) "Expression for the Wald distribution in terms of normal distribution". Radiotekhnika Electronika, 7, 164-166.

APPENDIX

Consider Y_i which is IG with parameters the reciprocal of $\underset{\sim}{x}_i'\underset{\sim}{\beta}$ and λ. Let

$$Z_i = \underset{\sim}{x}_i'\underset{\sim}{\beta}\sqrt{Y_i} - \frac{1}{\sqrt{Y_i}}$$

then Z_i will have a distribution $(1+h(z_i))\frac{1}{\sqrt{2\pi}}\exp(-\frac{z_i^2}{2})$ and $Z_i^2 \sim \chi_1^2$ where $h(z_i)$ is an odd function and $|h(z_i)| \leq 1$.

2

VITA

Mammo Woldie

Candidate for the Degree of

Doctor of Philosophy

Thesis: INVERSE GAUSSIAN REGRESSION MODELS

Major Field: Statistics

Biographical:

Personal Data: Born in Arur Amba, Shewa, Ethiopia, April 5, 1947, the son of Ato Woldie Mermara and Woizero Atsede Begashaw.

Education: Completed secondary education in Haile Mariam Mammo Secondary School and Special Teachers' Education Program in June, 1965; received Bachelor of Science degree in Mathematics from Education Faculty of Haile Selassie I University, Ethiopia, in July, 1970; attended summer course in the Department of Mathematics, Perugia, Italy, July 16 - August 15, 1978; received Master of Science degree in Statistics from Western Michigan University, Kalamazoo, Michigan, in August, 1980; completed the requirements for the Doctor of Philosophy degree at Oklahoma State University, Stillwater, Oklahoma, in July, 1984.

Professional Experience: Ethiopian University Service, September, 1968 - July, 1969; Mathematics teacher in Dej. Wolde Gebriel Aba Seitan Secondary School, Haile Mariam Mammo Secondary School and Special Teachers' Education Program, July, 1970 - December, 1972; Mathematics Lecturer, Science Faculty, Addis Ababa University, January, 1973 - July, 1978; Graduate Teaching Associate, Department of Statistics, Oklahoma State University, August, 1980 - January, 1984.

Professional Organization: Mu Sigma Rho.