SEQUENTIAL ROOT ESTIMATION FOR THE

BETA-BINOMIAL DISTRIBUTION

By

LAURA PRICE COOMBS

Bachelor of Science
Florida Southern College
Lakeland, Florida
1987

Master of Statistics
University of Florida
Gainesville, Florida
1989

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 1999

SEQUENTIAL ROOT ESTIMATION FOR THE

BETA-BINOMIAL DISTRIBUTION

Thesis Approved:

Barry Kurt Moser
Thesis Adviser

Melinda H. McCann

Mark E Payton

[illegible]

Wayne B. Powell
Dean of the Graduate College

ACKNOWLEDGEMENTS

First and foremost, I would like to express my heartfelt thanks to my advisor, Dr. Moser, for providing support, guidance, encouragement, and friendship during these last four years. It would be impossible to overstate his contributions to both my professional and personal development. Dr. Moser epitomizes, in theory and practice, the ideals of scholarship, teaching, mentoring, and integrity.

Second, I would like to thank the members of my committee for reviewing my work and providing many helpful, constructive comments. In addition to their roles as academic jurors, I additionally wish to thank Dr. McCann for helping me with Fortran programming, Dr. Payton for convincing me I could succeed when I didn't believe I would, and Dr. Fuqua for personal counsel and fellowship.

Third, I would like to thank my close friend and fellow student, Jim Blum, for being my personal tutor, sounding board, and companion. I could have survived neither advanced calculus nor real analysis without his instruction and wit.

Finally, I would like to thank my family. That my parents would move to Stillwater to care for my children so that I could concentrate on my studies involves such a degree of personal sacrifice that any form of thank you seems grossly inadequate. That I am not surprised, but still overwhelmed, by their generosity speaks to their history as parents. It is my singular good fortune

to have been so blessed. My children, Tommy, Billy, and Rachel have provided me with the opportunity to relive my childhood, to recognize the values I cherish, and to identify what really is important. My husband, Tom, has sacrificed his freedom and sanity and yet still has managed to provide me with the laughter, affection, and love that I needed to make this possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Let $Y(x)$ be a random variable representing the results of an experiment whose outcomes are dichotomous, response or no response. Let $\mu(x)$ be the expectation of $Y(x)$, the probability of response at a given design point, x. Consider the problem of estimating the roots of $\mu(x)$. Define the root, L_p, to be the value of x where $\mu(x) = p$. For example, in drug testing, $L_{.5}$ is the dose level at which the drug produces toxicity in 50% of the subjects.

Various procedures for finding roots of $\mu(x)$ by sequentially selecting values of the design point x have been developed. One such sequential approximation method called the Up and Down Method was developed by Dixon and Mood (1948). After n updates of the process, the response at the current design point, X_n, is used to determine the next design point, X_{n+1}, using a predetermined constant step size. Specifically, on the n^{th} update if Y_n=0 represents "no response" and Y_n=1 represents "response" then the next design point is determined by:

$$X_{n+1} = X_n + \Delta \quad \text{if } Y_n = 0 \quad \text{and}$$
$$X_{n+1} = X_n - \Delta \quad \text{if } Y_n = 1$$

where Δ is the step size. This and other Up and Down methods described by Storer (1989) have the undesirable property of being "memoryless" in that information that has accrued prior to the current design point is not used in determining the next design point. Another undesirable property of these types of procedures is that X_n does not converge to L_p.

In 1951, Robbins and Monro introduced a procedure for estimating L_p known as stochastic approximation. They suggested using the following rule to determine successive design points for estimating L_p:

$$X_{n+1} = X_n - a_n(Y_n - p)$$

where a_n is a fixed sequence of positive constants and Y_n is the response associated with X_n. Like the Up and Down Method, the Robbins-Monro procedure generates X_{n+1} as a linear function of X_n. However, instead of using a single, constant step size, the step size is allowed to vary from update to update. For estimating a single root, L_p, Robbins and Monro demonstrated that X_n converges to L_p in L^2 under the following conditions:

a) $0 < \sum_{n=1}^{\infty} a_n^2 < \infty$ and $\sum_{n=1}^{\infty} a_n = \infty$.

b) There exists a positive constant c such that $P(|Y(x)| \leq c) = \int_{-c}^{c} \partial F(y|x) = 1$ for all x where $F(y|x)$ is the cumulative distribution function of $Y(x)$.

c) $\mu(x)$ is nondecreasing, $\mu(L_p) = p$ and $\mu'(L_p) > 0$.

Blum (1954) showed that X_n converges to L_p almost surely assuming that conditions a) and c) above hold plus the following conditions on the Lebesgue-measurable function $\mu(x)$:

a) $|\mu(x)| \leq c + d|x| \qquad c, d \geq 0$

b) $\int_{-\infty}^{\infty} [y - \mu(x)]^2 \partial F(y|x) \leq \sigma^2 < \infty$

c) $\mu(x)$ is bounded away from p outside every neighborhood of $X=L_p$.

Using $a_n = a/n$ where a is a positive constant, Chung (1954) showed that

$\sqrt{n}(X_n - L_p)$ is asymptotically normal with mean zero and variance

$$\text{var}[\sqrt{n}(X_n - L_p)] = \frac{a^2\sigma^2}{2a\mu'(L_p) - 1}$$

where σ^2 is assumed to be a constant which does not depend on x

and $\mu'(L_p) = \left.\frac{\partial\mu(x)}{\partial(x)}\right|_{x=L_p}$. The following conditions were required for the proof:

a) $\mu'(L_p) > 0$

b) $\mu(x)$ is bounded away from p outside every neighborhood of $x=L_p$

c) $\mu(x)$ is bounded on any finite interval of x.

d) $\text{E}[(Y(x) - \mu(x))^p] < \infty$ for every integer $p>1$.

Thus when $a_n = [n\mu'(L_p)]^{-1}$, the Robbins-Monro process is optimal (i.e., the asymptotic variance is minimized).

The optimal Robbins-Monro procedure requires knowledge of the derivative of the true response curve, $\mu'(x)$. Since $\mu(x)$ and $\mu'(x)$ are not usually known, Anbar (1978) suggested replacing $\mu'(x)$ with a strongly consistent estimator, namely the slope of the least squares line:

$$b_n^* = \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \Big/ \sum_{i=1}^n (X_i - \bar{X})^2$$

The next design points are determined by the rule

$$X_{n+1} = X_n - (nb_n)^{-1}(Y_n - p) \qquad (1)$$

where

$$b_n = \begin{cases} \delta_1 & \text{if } b_n^* \le \delta_1 \\ b_n^* & \text{if } \delta_1 < b_n^* < \delta_2 \\ \delta_2 & \text{if } \delta_2 \le b_n^* \end{cases} \qquad 0 < \delta_1 < \delta_2 < \infty$$

Anbar proved that X_n converges to L_p almost surely and b_n converges to $\mu'(L_p)$ almost surely. He also proved that $\sqrt{n}(X_n - L_p)$ has the same asymptotic distribution as the optimal Robbins-Monro process.

Wu (1985) suggested estimating the root L_p from an estimate of the entire response curve, $\mu(x)$. He used a parametric form, $H(x|\theta)$ where $\theta=(\theta_1, \ldots, \theta_k)$, to represent the expectation of Y. The updating rule is described below:

1) Find an efficient estimate $\hat{\theta}_n$ for θ based on the n pairs of observations $[(x_1, y_1),$ $(x_2, y_2) \ldots (x_n, y_n)]$.

2) Define the estimated response curve $\hat{H}_n(x) = H(x \mid \hat{\theta}_n)$ and choose x_{n+1} so that $\hat{H}_n(x_{n+1}) = p$.

Using maximum likelihood estimators (MLEs) as the efficient estimators and using a one-parameter logit model for $H(x|\theta)$, Wu demonstrated that his procedure is equivalent to a Robbins-Monro procedure and hence X_n converges to L_p almost surely regardless of whether $H(x|\theta)=\mu(x)$. He was unable, however, to show consistency when using a two-parameter logit model for $H(x|\theta)$.

Shen and O'Quigley (1996) used a procedure similar to Wu's called the continual reassessment method (CRM) to estimate target dose levels in dose finding studies. This method uses a one-parameter model, $\psi(x, \theta)$ to represent $\mu(x)$ where $\mu(x)$ is considered unknown. The difference between CRM and Wu's procedure is that CRM limits its design points to a small panel of discrete dose levels. Shen and O'Quigley established consistency and asymptotic normality of the MLEs under model misspecification and proved that the dose level will converge to the closest discrete dose level to L_p under the following conditions:

1) a) For each θ, $\psi(., \theta)$ is strictly increasing.

   b) The function $\psi(x, .)$ is continuous and is strictly monotone in θ in the same direction for all x.

2) For each $0<t<1$ and each x, the function

$$s(t,x,\theta) = t\frac{\psi'}{1-\psi}(x,\theta) + (1-t)\frac{-\psi'}{1-\psi}(x,\theta)$$

   is continuous and strictly increasing in θ.

3) The parameter θ belongs to the finite interval $[A, B]$.

4) The target dose level is L_p, that is $\mu(L_p)=p$.

5) The probabilities of toxicity at $x_1,\ldots,x_m$ satisfy $0<\mu(x_1)<\ldots<\mu(x_m)<1$.

6) For $i=1,\ldots,m$, $\theta_i \in S$ where the θ_i are such that $\psi(x, \theta_i) = \mu(x_i)$ and the set S is defined as:

$$S = \{\theta : |\psi(L_p,\theta) - p| < |\psi(L_p,\theta) - p| \text{ for all } x_i \neq L_p\}$$

Moser and Faries (1996) developed the Sequential Approximation Method (SAM) which provides a method for estimating any number of roots L_p using a parametric model to represent $\mu(x)$ where $\mu(x)$ is assumed unknown. On each update of the two parameter version of SAM, observations are taken at two different x values, x_{n1} and x_{n2}. The procedure searches for two roots, L_{p_1} and L_{p_2}. Moser and Faries extended the results of Shen and O'Quigley for the discrete set of x values to the two-parameter logit model, $G(x, \theta_1, \theta_2)$ proving convergence provided $G(x, \theta_1, \theta_2)$ is sufficiently close to the true model $\mu(x)$. In particular, let $\tilde{\theta}_1^n$ and $\tilde{\theta}_2^n$ be solutions to

$\tilde{I}_n^1(\theta_1,\theta_2)=0$ and $\tilde{I}_n^2(\theta_1,\theta_2)=0$ where $\tilde{I}_n^1$ and $\tilde{I}_n^2$ are the expected value of the derivative of the likelihood function with $\mu(x)$ replaced by $G(x,\theta_1,\theta_2)$. Let (θ_1^0,θ_2^0) be the value of θ_1 and θ_2 such that $G(L_{p_j},\theta_1^0,\theta_2^0)=\mu(L_{p_j})$ for j=1,2. Define the set

$$S(\theta_1^0,\theta_2^0)=$$

$$\{(\theta_1,\theta_2):\forall i=1,\ldots m[x_{ij}:\forall j=1,2\, x_{ij}\neq x_j^0 \Rightarrow |G(L_{p_j},\theta_1,\theta_2)-p_j|<|G(x_{ij},\theta_1,\theta_2)-p_j|]\}$$

Then the (x_{n1}, x_{n2}) converge to the target dose levels (L_{p_1},L_{p_2}) and the MLEs $(\hat{\theta}_1^n,\hat{\theta}_2^n)$ converge to (θ_1^0,θ_2^0) provided $(\tilde{\theta}_1^n,\tilde{\theta}_2^n)\in S(\theta_1^0,\theta_2^0)$ for n>N.

In all of the procedures mentioned to this point, the observed design points have been dependent from one update to the next, but the responses at a given design point on the same update have been assumed to be conditionally independent. There are experimental situations, however, where the responses at a given dose level are not independent. For example, suppose the response curve $\mu(x)$ represents the proportion of subjects that have a toxic response to a drug at dose level x. Now suppose pregnant animals are administered a drug and each individual fetus within a litter is examined for a response. Responses among fetuses within a litter are expected to be more similar than responses across litters. Piegorsch (1993) discusses different ways of adjusting for this intralitter correlation. One possibility is to assume that responses within a litter at a particular dose level follow a beta-binomial distribution (Williams, 1975; Haseman and Kupper, 1979). Kupper *et al.* (1986) examined the influence of "litter effect" for beta-binomial data at fixed dose levels using a simulation study.

In comparing Kupper *et al.* to Moser and Faries note that both sets of authors introduce the complexities of dependent observations into their studies; however, these complexities materialize in different ways. Kupper *et al.* examine dependent (or correlated) observations at fixed dose levels using the beta-binomial distribution but assume that observations from one fixed dose level to the next are independent. Moser and Faries examine observations that are independent at given dose levels, but consider a sequential process that produces dependent observations from update to update.

The objective of this thesis is to estimate dose levels where toxicity occurs in specified proportions of subjects when the observed responses within a litter follow a beta-binomial distribution. This is done by combining the more complicated dependent aspects of Kupper *et al.*'s beta-binomial distribution with the complex dependent portions of Moser and Faries's sequential process. Specifically, the two-parameter logit SAM procedure is used to estimate dose levels. By applying a sequential process to beta-binomial data, dependencies within a litter as well as dependencies between observations from update to update are incurred. In this more complicated context, the following theoretical results are derived:

a) consistency of the MLEs,

b) asymptotic normality of the MLEs, and

c) consistency of the observed dose levels to the true dose level or to the closest value to the true dose level in the discrete set.

In Chapter 2 of this thesis the details for using the two parameter logit SAM procedure when the responses follow a beta-binomial distribution are given, followed by an example. The proofs of the asymptotic results are given in the Chapter 3. Chapter 4 describes the simulation results on the performance of the SAM procedure for small and medium samples. Conclusions are given in Chapter 5.

CHAPTER 2

SAM PROCEDURE FOR BETA-BINOMIAL DATA

2.1 *Notation and Updating Rule for SAM*

In a general context, consider a random variable Y whose expectation is an increasing function of a variable x. The exact form of the expectation of Y given x is unknown but can be denoted by the function $\mu(x)$. The objective is to sequentially observe values of Y at selected values of x and use the (x, y) pairs to estimate roots of $\mu(x)$ (a root of $\mu(x)$ is a value of x where $\mu(x)$ equals a specified constant).

In the specific context of a toxicity study, x is a dose level and Y is a binary response that takes on the value 1 if a fetus is malformed and 0 otherwise. The expectation of Y given x, $\mu(x) = P(Y=1|x)$, equals the probability of a malformed fetus at dose level x. The objective is to sequentially observe (x, y) pairs to estimate dose levels where pre-specified proportions of fetuses are malformed. To this end specific notation is established to identify the fetus, the litter, and the dose at a particular update of the sequential process.

Let p_1 and p_2 be two pre-specified proportions $0<p_1<p_2<1$. Let L_{p_j} be the value of x (the dose level) where $\mu(x)=p_j$ for $j=1, 2$. At the i^{th} update of the SAM process, observations are taken at dose levels x_{ij} for $i = 1, \ldots, n; j = 1, 2$. Suppose m_{ij} pregnant animals are tested at dose level x_{ij}, and each pregnant animal produces a litter of size r_{ijk} for $k=1,\ldots m_{ij}$. Further, suppose that the binary random variable $Y_{ijk\ell}$ takes the value 1 if the ℓ^{th} fetus within the k^{th} litter in the ij^{th} dose group gives a positive response (*e.g.*, is malformed) and takes the value 0 otherwise, for $\ell=1,\ldots,r_{ijk}$, $k=1,\ldots m_{ij}$. Hence, the random variable $Y_{ijk}=\sum_{\ell=1}^{r_{ijk}} Y_{ijk\ell}$ is the number of positive responses within the k^{th} litter in the ij^{th} dose group.

Williams (1975) and Haseman and Kupper (1979) suggest introducing the intralitter correlation by using the beta-binomial for the distribution of Y_{ijk} given x_{ij} :

$$P(Y_{ijk}=y_{ijk}\,|\,X_{ij}=x_{ij})=\binom{r_{ijk}}{y_{ijk}}\frac{\Gamma(\mu_{ij}/\gamma_{ij}+y_{ijk})\Gamma(r_{ijk}+(1-\mu_{ij})/\gamma_{ij}-y_{ijk})\Gamma(1/\gamma_{ij})}{\Gamma(\mu_{ij}/\gamma_{ij})\Gamma((1-\mu_{ij})/\gamma_{ij})\Gamma(1/\gamma_{ij}+r_{ijk})} \quad (1)$$

for $y_{ijk}=0,1,\ldots,r_{ijk}$. Under (1),

$$\mathrm{E}\left(\frac{Y_{ijk}}{r_{ijk}}\,\Big|\,X_{ij}=x_{ij}\right)=\mu_{ij}=\mu(x_{ij})$$

and

$$\mathrm{corr}(Y_{ijkl},Y_{i'j'k'l'})=\begin{cases}1 & \text{if } i=i',j=j',k=k',l=l'\\ \dfrac{\gamma_{ij}}{1+\gamma_{ij}} & \text{if } i=i',j=j',k=k',l\neq l'\\ 0 & \text{otherwise}\end{cases}$$

for $\gamma_{ij}>0$ where $\mu(x_{ij})=\mu_{ij}$ is the expected probability of a positive response at dose x_{ij} and $\gamma_{ij}/(1+\gamma_{ij})$ is the intralitter correlation in the ij^{th} dose group. The forms $\mu(x_{ij})$ and μ_{ij} will be used interchangeably in this paper, with the former used to stress that the expectation of the beta-binomial is a function of the dose level. Piegorsch (1993) refers to γ_{ij} as an overdispersion parameter measuring departure from binomial sampling. At $\gamma_{ij}/(1+\gamma_{ij})=0$, $Y_{ijk}|x_{ij}$ simplifies to a Binomial(r_{ijk}, $\mu(x_{ij})$) random variable. At $\gamma_{ij}/(1+\gamma_{ij})=1$, $r_{ijk}^{-1}Y_{ijk} \mid x_{ij}$ simpifies to a Bernoulli($\mu(x_{ij})$) random variable. At present time γ_{ij} will be restricted to a fixed unknown value, γ. In the appendix, the theoretical results will be extended, allowing this overdispersion parameter to vary across dose levels as γ_{ij}.

The two-parameter logit model, $\pi(x_{ij}, \theta_1, \theta_2)=[1+\exp\{-(\theta_1+\theta_2 x_{ij})\}]^{-1}$, will be used to represent the expected proportion of positive responses (the dose-response curve). However, $\mu(x_{ij})$ does not have to equal $\pi(x_{ij}, \theta_1, \theta_2)$. Rather, $\pi(x_{ij}, \theta_1, \theta_2)$ is introduced only so that SAM can utilize the maximum likelihood method to estimate the roots of $\mu(x)$.

The log-likelihood function where $\mu(x_{ij})$ is replaced by $\pi(x_{ij}, \theta_1, \theta_2)$ is given by

$$\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\left\{\sum_{s=0}^{y_{ijk}-1}\ln\left[\pi(x_{ij},\theta_1,\theta_2)+s\gamma\right]+\sum_{s=0}^{r_{ijk}-y_{ijk}-1}\ln\left[\left(1-\pi(x_{ij},\theta_1,\theta_2)\right)+s\gamma\right]-\sum_{s=0}^{r_{ijk}-1}\ln\left(1+s\gamma\right)\right\}, \quad (2)$$

where terms not involving $\pi(x_{ij},\theta_1,\theta_2)$ and γ have been ignored. Note that (2) is a function of θ_1, θ_2, and γ.

The updating rule for the two-parameter logit model SAM procedure when the responses within a litter follow a beta-binomial distribution is:

1. Based on the $\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}} r_{ijk}$ observations, calculate the values of θ_1, θ_2 and γ that maximize equation (2). Denote these MLEs by $(\hat{\theta}_1^n, \hat{\theta}_2^n, \hat{\gamma}^n)$.

2. Define the estimated expectation $\hat{\mu}_n(x) = \pi(x \mid \hat{\theta}_1^n, \hat{\theta}_2^n) = [1+\exp\{-(\hat{\theta}_1^n + \hat{\theta}_2^n x)\}]^{-1}$ and choose the next two dose levels, $x_{n+1,1}$, $x_{n+1,2}$ so that $\hat{\mu}_n(x_{n+1,j}) = p_j$ for j=1, 2. That is,

$$x_{n+1,1} = \frac{-\hat{\theta}_1^{(n)} + \ln\left(\frac{p_1}{1-p_1}\right)}{\hat{\theta}_2^{(n)}}$$

and

$$x_{n+1,2} = \frac{-\hat{\theta}_1^{(n)} + \ln\left(\frac{p_2}{1-p_2}\right)}{\hat{\theta}_2^{(n)}}. \tag{3}$$

After the n^{th} update, $x_{n+1,j}$ provides an estimate of L_{p_j} for j=1, 2.

2.2 *Example*

The following numerical example is used to demonstrate how SAM's root estimators are generated using the updating rule described above. Suppose that we are interested in finding the roots $L_{.2}$ and $L_{.8}$ when the responses, Y_{ijk}, have a beta-binomial distribution with $\gamma = 1$ and $\mu(x_{ij})$ set equal to the logit model with θ_1=-5.386 and θ_2=1.00, *i.e.*, $L_{.2}$=4.00 and $L_{.8}$=6.77. Our initial estimates of $L_{.2}$, $L_{.8}$, and $\gamma/(1+\gamma)$ are 4, 6, and 1/(1+1)=.5, respectively. That is, $x_{1,1}$=4, $x_{1,2}$=6, and the starting value of γ=1.

Two litters at each of the initial dose levels are observed, and the proportions of malformed fetuses for the four litters are recorded. The resulting set of observations is given below:

$$\left(x_{11},\frac{y_{111}}{r_{111}}\right)=\left(4,\frac{0}{15}\right),\ \left(x_{11},\frac{y_{112}}{r_{112}}\right)=\left(4,\frac{2}{12}\right),\left(x_{12},\frac{y_{121}}{r_{121}}\right)=\left(6,\frac{5}{10}\right),\ \left(x_{12},\frac{y_{122}}{r_{122}}\right)=\left(6,\frac{10}{10}\right).$$

From (2), the MLEs based on these observations are $\hat{\theta}_1^{(1)}$= -8.944, $\hat{\theta}_2^{(1)}$=1.716, and $\hat{\gamma}^{(1)}$=.341. Thus, from (3), the next two estimated dose levels are

$$x_{21}=4.404 \text{ and } x_{22}=6.020.$$

Responses for two more litters are now observed at both of these dose levels, and the MLEs $\hat{\theta}_1^{(2)}$, $\hat{\theta}_2^{(2)}$ $\hat{\gamma}^{(2)}$ are generated based on all eight pairs of observations. The dose levels $x_{3,1}$ and $x_{3,2}$ are then calculated from (3). The process continues in this manner. The results after 10 updates of the process are presented in Table 1. The final estimates of $L_{.2}$ and $L_{.8}$ are provided by $x_{11,1}$=4.35 and $x_{11,2}$=6.64, respectively. The estimator of any other root L_{p*} is the solution to $p* = [1+\exp\text{-}(\hat{\theta}_1^n+\hat{\theta}_2^n L_{p*})]^{-1}$. That is, $\hat{L}_{p*}^{(n)}=\{\ln[p*/(1-p*)]-\hat{\theta}_1^{(n)}\}/\hat{\theta}_2^{(n)}$. For example, the final estimates of $L_{.5}$ and $L_{.75}$ are $\hat{L}_{.5}^{(10)}=5.49$ and $\hat{L}_{.75}^{(10)}=6.40$.

Table 1. Example Data*

Update		Litter 1	Litter 2		Litter 1	Litter 2			
		at x_{i1}:	at x_{i1}:		at x_{i2}:	at x_{i2}:			
i	x_{i1}	y_{i11}/r_{i11}	y_{i12}/r_{i12}	x_{i2}	y_{i21}/r_{i21}	y_{i22}/r_{i22}	$\hat{\theta}_1^{(i)}$	$\hat{\theta}_2^{(i)}$	$\hat{\gamma}^{(i)}$
1	4.00	0/15	2/12	6.00	5/10	10/10	-8.94	1.72	.34
2	4.40	0/16	3/10	6.02	12/12	12/12	-11.71	2.35	.48
3	4.40	2/13	6/14	5.58	0/17	9/10	-8.07	1.59	.84
4	4.21	0/11	4/11	5.96	14/19	3/9	-7.39	1.41	.65
5	4.27	10/10	1/8	6.24	15/15	9/12	-5.92	1.16	.86
6	3.90	1/14	1/14	6.28	2/10	3/15	-4.67	.89	.82
7	3.71	0/10	0/16	6.83	11/11	15/15	-6.06	1.16	.85
8	4.03	0/8	0/17	6.41	0/18	13/14	-5.64	1.04	1.07
9	4.08	0/13	0/15	6.73	16/17	10/15	-5.94	1.07	.98
10	4.24	0/7	0/13	6.82	9/9	18/18	-6.65	1.21	1.02
	$\hat{L}_{.2}^{(10)}$ =			$\hat{L}_{.8}^{(10)}$ =					
	4.35			6.64					
	$L_{.2}$=			$L_{.8}$=			θ_1=	θ_2=	γ=
	4.00			6.77			-5.386	1.00	1.00

* All observations y_{ijk} are generated using a beta-binomial(r_{ijk}, $\mu(x_{ij})$, γ=1) where

$\mu(x_{ij})=[1+\exp-(\hat{\theta}_1^n+\hat{\theta}_2^n x_{ij})]^{-1}$ with θ_1=-5.386 and θ_2=1.00, i.e., $L_{.2}$=4.00 and $L_{.8}$=6.77.

CHAPTER 3

ASYMPTOTIC RESULTS

Let x_1, ..., x_d be d positive values where the x_{ij}'s can be observed. After the n^{th} update, the dose level $x_{n+1,1}$ and the dose level $x_{n+1,2}$ will be selected from a discrete list $x_1,\ldots,x_d$, such that $x_{n+1,1} < x_{n+1,2}$. Note that there are t=$d(d$-1)/2 possible pairs of dose levels such that $x_{n+1,1} < x_{n+1,2}$. A modified SAM procedure is used to generate new dose levels. This selection process is the same as that used in the appendix of Moser and Faries (1996). Specifically, after the n^{th} update, the procedure is defined as follows:

(a) Let $x_{n+1,1}$ be the value of x that minimizes $|\,\pi(x_{ij},\hat{\theta}_1^n,\hat{\theta}_2^n)-p_1\,|$ and let $x_{n+1,2}$ be the value of x that minimizes $|\,\pi(x_{ij},\hat{\theta}_1^n,\hat{\theta}_2^n)-p_2\,|$ where minimization is performed over all x= $x_1,\ldots,x_d$. The function $\pi(x_{ij},\theta_1,\theta_2)$ is the two-parameter logit model and $(\hat{\theta}_1^n,\hat{\theta}_2^n,\hat{\gamma}^n)$ are the maximum likelihood estimates calculated by maximizing (2) with respect to θ_1, θ_2 and γ.

(b) If $x_{n+1,1} \geq x_{n+1,2}$ from part (a), then reset $x_{n+1,1} = x_{n,1}$ and $x_{n+1,2} = x_{n,2}$.

Note, the consistency and asymptotic normality proofs below do not formally distinguish between rule (a) and rule (b) above. However, if convergence is to be attained then, asymptotically, $x_{n+1,1}$ and $x_{n+1,2}$ must take on unique values such that $x_{n+1,1} < x_{n+1,2}$. We have therefore included rule (b), which insures that $x_{n+1,1} < x_{n+1,2}$ for all n.

The first objective is to prove that $x_{n+1,1} \to x_1^0$ and $x_{n+1,2} \to x_2^0$ almost surely as $n \to \infty$ for the procedure defined in (a)-(b), where we initially assume that x_1^0 and x_2^0 $\in$

$x_1,\ldots,x_d$ are L_{p_1} and L_{p_2}, the values of x such that $\mu(L_{p_j}) = p_j$ for $j = 1, 2$. At the end of the consistency proof, we will broaden this definition of x_1^0 and x_2^0 slightly.

The following definitions and equations are required for the consistency proof. Let (θ_1^0, θ_2^0) be the value of (θ_1, θ_2) such that $\pi(x_j^0, \theta_1^0, \theta_2^0) = \mu(x_j^0)$ for j=1,2. Note that (θ_1^0, θ_2^0) depend on $\mu(x)$ and are therefore unknown parameters. Let γ^0 be the true value of γ.

After n updates of the (a), (b) process, the partial derivatives of the log-likelihood with respect to θ_1, θ_2 and γ (divided by $K = \sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}} r_{ijk}$) are

$$I_K^1(\theta_1,\theta_2,\gamma) =$$

$$K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\pi(x_{ij},\theta_1,\theta_2)[1-\pi(x_{ij},\theta_1,\theta_2)]\left\{\sum_{s=0}^{y_{ijk}-1}\left[\pi(x_{ij},\theta_1,\theta_2)+s\gamma\right]^{-1} - \sum_{s=0}^{r_{ijk}-y_{ijk}-1}\left([1-\pi(x_{ij},\theta_1,\theta_2)]+s\gamma\right)^{-1}\right\} \quad (4)$$

$$I_K^2(\theta_1,\theta_2,\gamma) =$$

$$K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}x_{ij}\pi(x_{ij},\theta_1,\theta_2)[1-\pi(x_{ij},\theta_1,\theta_2)]\left\{\sum_{s=0}^{y_{ijk}-1}\left[\pi(x_{ij},\theta_1,\theta_2)+s\gamma\right]^{-1} - \sum_{s=0}^{r_{ijk}-y_{ijk}-1}\left([1-\pi(x_{ij},\theta_1,\theta_2)]+s\gamma\right)^{-1}\right\} \quad (5)$$

and

$$I_K^3(\theta_1,\theta_2,\gamma)=$$

$$K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\left\{\sum_{s=0}^{y_{ijk}-1}\frac{s}{[\pi(x_{ij},\theta_1,\theta_2)+s\gamma]}+\sum_{s=0}^{r_{ijk}-y_{ijk}-1}\frac{s}{[1-\pi(x_{ij},\theta_1,\theta_2)]+s\gamma}-\sum_{s=0}^{r_{ijk}-1}\frac{s}{1=s\gamma}\right\}. \tag{6}$$

The following three equations are the expectation of I_K^1, I_K^2 and I_K^3 where the observed y_{ijk}'s are replaced by the random variable Y_{ijk}.

$$\widetilde{I}_K^1(\theta_1,\theta_2,\gamma)=\mathrm{E}(I_k^1(\theta_1,\theta_2,\gamma))=$$

$$K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\pi(x_{ij},\theta_1,\theta_2)[1-\pi(x_{ij},\theta_1,\theta_2)]\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}[\pi(x_{ij},\theta_1,\theta_2)+s\gamma]^{-1}\right.$$

$$\left.-\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}\left([1-\pi(x_{ij},\theta_1,\theta_2)]+s\gamma\right)^{-1}\right\} \tag{7}$$

$$\widetilde{I}_K^2(\theta_1,\theta_2,\gamma)=\mathrm{E}(I_k^2(\theta_1,\theta_2,\gamma))=$$

$$K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}x_{ij}\pi(x_{ij},\theta_1,\theta_2)[1-\pi(x_{ij},\theta_1,\theta_2)]\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}\left[\pi(x_{ij},\theta_1,\theta_2)+s\gamma\right]^{-1}\right.$$

$$\left.-\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}\left([1-\pi(x_{ij},\theta_1,\theta_2]+s\gamma\right)^{-1}\right\} \tag{8}$$

$$\widetilde{I}_K^3(\theta_1,\theta_2,\gamma)=\mathrm{E}(I_k^3(\theta_1,\theta_2,\gamma))$$

$$K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}\frac{s}{[\pi(x_{ij},\theta_1,\theta_2)+s\gamma]}+\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}\frac{s}{[1-\pi(x_{ij},\theta_1,\theta_2)]+s\gamma}\right.$$

$$\left.-\sum_{s=0}^{r_{ijk}-1}\frac{s}{1+s\gamma}\right\}. \tag{9}$$

Denote $(\widetilde{\theta}_1^{\,n},\widetilde{\theta}_2^{\,n},\widetilde{\gamma}^{\,n})$ as the solution to the equations $\widetilde{I}_K^1(\theta_1,\theta_2,\gamma)=$ $\widetilde{I}_K^2(\theta_1,\theta_2,\gamma)=$ $\widetilde{I}_K^3(\theta_1,\theta_2,\gamma)=0.$

Equations (7), (8) and (9) can be rewritten as

$$\widetilde{I}_K^1(\theta_1,\theta_2,\gamma)=$$

$$\sum_{i=1}^{t}\hat{f}_i\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\pi(x_{ij},\theta_1,\theta_2)(1-\pi(x_{ij},\theta_1,\theta_2))\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}[\pi(x_{ij},\theta_1,\theta_2)+s\gamma]^{-1}\right.$$

$$\left.-\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}((1-\pi(x_{ij},\theta_1,\theta_2))+s\gamma)^{-1}\right\} \tag{10}$$

$$\widetilde{I}_K^2(\theta_1,\theta_2,\gamma)=$$

$$\sum_{i=1}^{t}\hat{f}_i\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}x_{ij}\pi(x_{ij},\theta_1,\theta_2)(1-\pi(x_{ij},\theta_1,\theta_2))\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}[\pi(x_{ij},\theta_1,\theta_2)+s\gamma]^{-1}\right.$$

$$\left.-\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}((1-\pi(x_{ij},\theta_1,\theta_2))+s\gamma)^{-1}\right\} \tag{11}$$

$\tilde{I}_K^3(\theta_1,\theta_2,\gamma)=$

$$\sum_{i=1}^{t}\hat{f}_i\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\mathrm{E}\left\{\sum_{s=0}^{y_{ijk}-1}\frac{s}{[\pi(x_{ij},\theta_1,\theta_2)+s\gamma]}+\sum_{s=0}^{r_{ijk}-y_{ijk}-1}\frac{s}{[1-\pi(x_{ij},\theta_1,\theta_2)]+s\gamma}\right.$$

$$\left.-\sum_{s=0}^{r_{ijk}-1}\frac{s}{1+s\gamma}\right\} \tag{12}$$

where $\hat{f}_i$ is the relative frequency of quadruplets (x_{i1}, m_{i1}, x_{i2}, m_{i2}) after the first n updates ($m_{ij} < \infty$).

In order to establish consistency, $\mu(x)$ cannot differ too much from $\pi(x,\theta_1,\theta_2)$. To characterize their difference, define the set

$$S(\theta_1^0,\theta_2^0)=\{(\theta_1,\theta_2):\forall i=1,\ldots,t\ [x_{ij}:\forall j=1,2\ \mathrm{x}_{ij}\neq \mathrm{x}_j^0 \Rightarrow$$

$$|\,\pi(x_j^0,\theta_1,\theta_2)-p_j\,|<|\,\pi(x_{ij},\theta_1,\theta_2)-p_j\,|]\}. \tag{13}$$

The following condition, which is required for the consistency proof, dictates how close $\mu(x)$ is to $\pi(x,\theta_1,\theta_2)$.

Condition C.1. For $n > N$, $(\tilde{\theta}_1^{\,n},\tilde{\theta}_2^{\,n})\in S(\theta_1^0,\theta_2^0)$ where N is a finite positive value.

The following two lemmas will be needed in the proof of Theorem 1:

Lemma 1. *Let Y ~Beta-binomial(p,γ,r). Then*

$$\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{1}{p+s\gamma}\right)=\mathrm{E}\left(\sum_{s=0}^{r-Y-1}\frac{1}{1-p+s\gamma}\right).$$

The proof of Lemma 1 is given in Appendix A.

Lemma 2. *Let Y~Beta-binomial(p,γ,r). Then*

$$\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{s}{p+s\gamma}+\sum_{s=0}^{r-Y-1}\frac{s}{1-p+s\gamma}\right)=\sum_{s=0}^{r-1}\frac{s}{1+s\gamma}.$$

The proof of Lemma 2 is given in Appendix B.

Theorem 1. *Assume that condition C.1 is satisfied. Let* $\left(\hat{\theta}_1^n,\hat{\theta}_2^n,\hat{\gamma}^n\right)$ *be the maximum likelihood estimators of* $(\theta_1, \theta_2, \gamma)$ *and let* $(x_{n+1,1}, x_{n+1,2})$ *be the selected dose levels after the* n^{th} *update of the procedure; then almost surely,* $(x_{n+1,1}, x_{n+1,2}) \to \left(x_1^0, x_2^0\right)$ *and* $\left(\hat{\theta}_1^n,\hat{\theta}_2^n,\hat{\gamma}^n\right)\to\left(\theta_1^0,\theta_2^0,\gamma^0\right)$ *as* $n \to \infty$ *where* $\left(\hat{\theta}_1^n,\hat{\theta}_2^n,\hat{\gamma}^n\right)$ *and* $(x_{n+1,1}, x_{n+1,2})$ *are defined in part a) of the procedure.*

Proof. First, observe that for each dose level x_{ij}, the functions

$$\pi(\theta_1,\theta_2)[1-\pi(\theta_1,\theta_2)]\left[\pi(x_{ij},\theta_1,\theta_2)+s\gamma\right]^{-1}$$

and

$$\pi(\theta_1,\theta_2)[1-\pi(\theta_1,\theta_2)]\left[1-\pi(x_{ij},\theta_1,\theta_2)+s\gamma\right]^{-1}$$

are uniformly continuous in θ_1 and θ_2 over the finite rectangle [A, B] × [C, D]. Then for any ε > 0 and for each x_{ij}, there exists a partition A= $t_0 < t_1 < \ldots < t_p$ = B and C= $s_0 < s_1 < \ldots < s_q$ = D such that for any $(\theta_1, \theta_2) \in$ [A, B] × [C, D] there exists a $(t_{p_0}, s_{q_0}) \in$ [A, B] × [C, D] such that

$$\Big|\pi\left(x_{ij},t_{p_0},s_{q_0}\right)\left(1-\pi\left(x_{ij},t_{p_0},s_{q_0}\right)\right)\left[\pi\left(x_{ij},t_{p_0},s_{q_0}\right)+s\gamma\right]^{-1}$$

$$-\pi\left(x_{ij},\theta_1,\theta_2\right)\left(1-\pi\left(x_{ij},\theta_1,\theta_2\right)\right)\left[\pi\left(x_{ij},\theta_1,\theta_2\right)+s\gamma\right]^{-1}\Big|<\varepsilon \qquad (14)$$

and

$$\Big|\pi\left(x_{ij},t_{p_0},s_{q_0}\right)\left(1-\pi\left(x_{ij},t_{p_0},s_{q_0}\right)\right)\left[1-\pi\left(x_{ij},t_{p_0},s_{q_0}\right)+s\gamma\right]^{-1}$$

$$-\pi\left(x_{ij},\theta_1,\theta_2\right)\left(1-\pi\left(x_{ij},\theta_1,\theta_2\right)\right)\left[1-\pi\left(x_{ij},\theta_1,\theta_2\right)+s\gamma\right]^{-1}\Big|<\varepsilon \qquad (15)$$

Because there are only d possible dose levels and because y_{ijk} is a discrete random variable with only $(r_{ijk}+1)$ possible values, the partition may be chosen so that restrictions (14) and (15) are valid for all x_{ij}, y_{ijk} pairs.

Next, separate $I_K^1(\theta_1,\theta_2,\gamma)-\widetilde{I}_K^1(\theta_1,\theta_2,\gamma)$ into the sum of three pieces: $I_{K1}^1(\theta_1,\theta_2,\gamma)$, I_{K2}^1, $I_{K3}^1(\theta_1,\theta_2,\gamma)$, where

$$I_{K1}^1\left(\theta_1,\theta_2,\gamma\right)=$$

$$K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\pi\left(x_{ij},\theta_1,\theta_2\right)\left(1-\pi\left(x_{ij},\theta_1,\theta_2\right)\right)\left\{\sum_{s=0}^{y_{ijk}-1}\left[\pi\left(x_{ij},\theta_1,\theta_2\right)+s\gamma\right]^{-1}\right.$$

$$\left.-\sum_{s=0}^{r_{ijk}-y_{ijk}-1}\left(\left(1-\pi\left(x_{ij},\theta_1,\theta_2\right)\right)+s\gamma\right)^{-1}\right\}$$

$$-K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\left(1-\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\right)\left\{\sum_{s=0}^{y_{ijk}-1}\left[\pi\left(x_{ij},t_{p_0},s_{q_0}\right)+s\gamma\right]^{-1}\right.$$

$$\left.-\sum_{s=0}^{r_{ijk}-y_{ijk}-1}\left(\left(1-\pi\left(x_{ij},t_{p_0},s_{q_0}\right)\right)+s\gamma\right)^{-1}\right\}$$

$I_{K2}^{1} =$

$$K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\left(1-\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\right)\left\{\sum_{s=0}^{y_{ijk}-1}\left[\pi\left(x_{ij},t_{p_0},s_{q_0}\right)+s\gamma\right]^{-1}\right.$$

$$\left.-\sum_{s=0}^{r_{ijk}-y_{ijk}-1}\left(\left(1-\pi\left(x_{ij},t_{p_0},s_{q_0}\right)\right)+s\gamma\right)^{-1}\right\}$$

$$-K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\left(1-\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\right)\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}\left[\pi\left(x_{ij},t_{p_0},s_{q_0}\right)+s\gamma\right]^{-1}\right.$$

$$\left.-\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}\left(\left(1-\pi\left(x_{ij},t_{p_0},s_{q_0}\right)\right)+s\gamma\right)^{-1}\right\}$$

and

$I_{K3}^{1}\left(\theta_1,\theta_2,\gamma\right) =$

$$K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\left(1-\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\right)\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}\left[\pi\left(x_{ij},t_{p_0},s_{q_0}\right)+s\gamma\right]^{-1}\right.$$

$$\left.-\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}\left(\left(1-\pi\left(x_{ij},t_{p_0},s_{q_0}\right)\right)+s\gamma\right)^{-1}\right\}$$

$$-K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\pi\left(x_{ij},\theta_1,\theta_2\right)\left(1-\pi\left(x_{ij},\theta_1,\theta_2\right)\right)\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}\left[\pi\left(x_{ij},\theta_1,\theta_2\right)+s\gamma\right]^{-1}\right.$$

$$\left.-\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}\left(\left(1-\pi\left(x_{ij},\theta_1,\theta_2\right)\right)+s\gamma\right)^{-1}\right\}$$

From (14) and (15) it follows that for any sequence y_{ijk}, i=1,...,n; j=1,2; k=1,...,m_{ij},

$$\sup_{t_{p0},s_{q0}\in[A,B]\times[C,D]}\left|I_{K1}^{1}(\theta_1,\theta_2)\right| \le K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\sum_{s=0}^{y_{ijk}-1}\varepsilon + K^{-1}\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\sum_{s=0}^{r_{ijk}-y_{ijk}-1}\varepsilon < \varepsilon.$$

Since $I_{K3}^{1}(\theta_1,\theta_2,\gamma)$= -$\mathrm{E}_{Y_{111},...,Y_{n2m_{n2}}}\left[I_{K1}^{1}(\theta_1,\theta_2,\gamma)\right]$, $I_{K3}^{1}(\theta_1,\theta_2,\gamma)$ is a convex combination of $\mathrm{I}_{K1}^{1}(\theta_1,\theta_2)$ evaluated over all possible sequences of $Y_{111,...,}Y_{n2m_{n2}}$; therefore,

$$\sup_{t_{p0},s_{q0}\in[A,B]\times[C,D]}\left|I_{K3}^{1}(\theta_1,\theta_2,\gamma)\right| \le \varepsilon\,.$$

Now it must be shown that for each $\left(t_{p_0},s_{q_0}\right)$, I_{K2}^{1} tends to zero almost surely. Let

$$\mathrm{g}(Y_{ijk}) = \pi\left(x_{ij},t_{p_0},s_{p_0}\right)\left(1-\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\right)\left\{\sum_{s=0}^{Y_{ijk}-1}\left[\pi\left(x_{ij},t_{p_0},s_{p_0}\right)+s\gamma\right]^{-1} - \sum_{s=0}^{r_{ijk}-Y_{ijk}-1}\left(\left(1-\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\right)+s\gamma\right)^{-1}\right\}$$

The expected value of $g(Y_{ijk})$ is:

$$\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\left(1-\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\right)\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}\left[\pi\left(x_{ij},t_{p_0},s_{p_0}\right)+s\gamma\right]^{-1} - \sum_{s=0}^{r_{ijk}-Y_{ijk}-1}\left(\left(1-\pi\left(x_{ij},t_{p_0},s_{p_0}\right)\right)+s\gamma\right)^{-1}\right\}.$$

Let $Z_{ijk} = g(Y_{ijk})$-E$[g(Y_{ijk})]$. Define the σ-field $\Im_n = \Im(Z_{111},\ldots,Z_{n2m_{n2}})$ and $I_{K2}^1 = K^{-1}\, T_n$

where $T_n = \sum_{i=1}^{n} \sum_{j=1}^{2} \sum_{k=1}^{m_{ij}} Z_{ijk}$. Then

$$
\begin{aligned}
\mathrm{E}\left(T_{n+1} \middle| \Im_n\right) &= \mathrm{E}\left(T_n + \sum_{j=1}^{2} \sum_{k=1}^{m_{n+1,j}} Z_{n+1,j,k} \middle| \Im_n \right) \\
&= T_n + \mathrm{E}\left(\sum_{j=1}^{2} \sum_{k=1}^{m_{n+1,j}} Z_{n+1,j,k} \middle| \Im_n \right) \\
&= T_n + \sum_{j=1}^{2} \sum_{k=1}^{m_{n+1,j}} \mathrm{E}\left(Z_{n+1,j,k}\right) \\
&= T_n \qquad \text{(since } \mathrm{E}\left(Z_{n+1,j,k}\right) = 0 \\
&\qquad\qquad \text{for all } j=1,2;\ k=1,\ldots,m_{ij})
\end{aligned}
$$

So T_n forms a martingale for fixed (t_{p_0}, s_{q_0}). Note that the terms in the summation in I_{K2}^1 are bounded. The limit theorem for martingales (Corollary 2 in Section 7.3 of Shiryayev 1984) implies that I_{K2}^1 converges to zero almost surely.

We have now established that

$$
\sup_{\theta_1,\theta_2 \in [A,B]\times[C,D]} \left| I_K^1(\theta_1,\theta_2,\gamma) - \tilde{I}_K^1(\theta_1,\theta_2,\gamma) \right| \to 0, a.s.. \tag{16}
$$

Using a similar argument we can also conclude that

$$
\sup_{\theta_1,\theta_2 \in [A,B]\times[C,D]} \left| I_K^2(\theta_1,\theta_2,\gamma) - \tilde{I}_K^2(\theta_1,\theta_2,\gamma) \right| \to 0, a.s. \tag{17}
$$

and

$$\sup_{\theta_1,\theta_2\in[A,B]\times[C,D]}\left|I_K^3(\theta_1,\theta_2,\gamma)-\widetilde{I}_K^3(\theta_1,\theta_2,\gamma)\right|\to 0, a.s. \qquad (18)$$

Since $(\hat{\theta}_1^n,\hat{\theta}_2^n,\hat{\gamma}^n)$ is the solution to (4), (5), and (6), Condition C.1, (16), (17) and (18) insure that almost surely, $(\hat{\theta}_1^n,\hat{\theta}_2^n)\in S(\theta_1^0,\theta_2^0)$ eventually. Hence, $(\hat{\theta}_1^n,\hat{\theta}_2^n)$ satisfies

$$\forall i=1,\ldots,t\ [\forall j=1,2\ \ x_{ij}\neq x_j^0 \Rightarrow |\pi(x_j^0,\hat{\theta}_1^n,\hat{\theta}_2^n)-p_j|<|\pi(x_{ij},\hat{\theta}_1^n,\hat{\theta}_2^n)-p_j|].$$

Thus, for large enough n, $(x_{n+1,1},x_{n+1,2})\equiv(x_1^0,x_2^0)$.

To establish the consistency of $(\hat{\theta}_1^n,\hat{\theta}_2^n,\hat{\gamma}^n)$, observe that as n tends to infinity all the $\hat{f}_i$'s in (10), (11) and (12) tend to zero, except for the one corresponding to (x_1^0,x_2^0), the sum of which tends to 1. Thus, $(\widetilde{\theta}_1^n,\widetilde{\theta}_2^n,\widetilde{\gamma}^n)$, the solution to $\widetilde{I}_K^1(\theta_1,\theta_2,\gamma)$ = $\widetilde{I}_K^2(\theta_1,\theta_2,\gamma)$ = $\widetilde{I}_K^3(\theta_1,\theta_2,\gamma)$=0 will be close to the solution of

$$\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\pi(x_j^0,\theta_1,\theta_2)(1-\pi(x_j^0,\theta_1,\theta_2))\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}\left[\pi(x_j^0,\theta_1,\theta_2)+s\gamma\right]^{-1}\right.$$
$$\left.-\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}\left((1-\pi(x_j^0,\theta_1,\theta_2))+s\gamma\right)^{-1}\right\}=0 \qquad (19)$$

$$\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}x_j^0\pi(x_j^0,\theta_1,\theta_2)(1-\pi(x_j^0,\theta_1,\theta_2))\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}\left[\pi(x_j^0,\theta_1,\theta_2)+s\gamma\right]^{-1}\right.$$
$$\left.-\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}\left((1-\pi(x_j^0,\theta_1,\theta_2))+s\gamma\right)^{-1}\right\}=0 \qquad (20)$$

and

$$\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}} \mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1} \frac{s}{\left[\pi\left(x_j^0,\theta_1,\theta_2\right)+s\gamma\right]} + \sum_{s=0}^{r_{ijk}-Y_{ijk}-1} \frac{s}{\left(1-\pi\left(x_j^0,\theta_1,\theta_2\right)\right)+s\gamma} - \sum_{s=0}^{r_{ijk}-1} \frac{s}{1+s\gamma}\right\}=0 \qquad (21)$$

By Lemma 1,

$$\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}\left[\mu\left(x_j^0\right)+s\gamma^0\right]^{-1} - \sum_{s=0}^{Y_{ijk}-r_{ijk}-1}\left[\left(1-\mu\left(x_j^0\right)\right)+s\gamma^0\right]^{-1}\right\}=0$$

for j=1,2 since $Y_{ijk} \sim \mathrm{BB}\left(\mu(x_j^0),\gamma^0,r_{ijk}\right)$ for any i=1,..n and any k=1,...,m_{ij}. In addition, by Lemma 2,

$$\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1} \frac{s}{[\mu(x_j^0)+s\gamma^0]} + \sum_{s=0}^{r_{ijk}-Y_{ijk}-1} \frac{s}{[1-\mu(x_j^0)]+s\gamma^0} - \sum_{s=0}^{r_{ijk}-1} \frac{s}{1+s\gamma^0}\right\}=0.$$

Since $\pi\left(x_j^0,\theta_1^0,\theta_2^0\right)= \mu(x_j^0)$, the solution to (19), (20), and (21) is $\left(\theta_1^0,\theta_2^0,\gamma^0\right)$. Applying (16), (17), and (18) we obtain the consistency of $\left(\hat{\theta}_1^n,\hat{\theta}_2^n,\hat{\gamma}^n\right)$. This finishes the proof.

Remark on the value of x_j^0: The proof of Theorem 1 does not depend on the assumption that $\mu(x_j^0)=p_j$ for j=1,2. Therefore, (x_{n1}, x_{n2}) converges to $\left(x_1^0,x_2^0\right)$ almost surely as long as $\mu(x_j^0)$ is closest to p_j among all possible $x_1,\ldots,x_d$ for j=1,2.

Remark on Condition C.1. Suppose that the true model is the logit model. Therefore, $\pi\left(x_{ij},\theta_1^0,\theta_2^0\right)=\mu(x_{ij})$, and the solution to (7), (8), (9) is $\left(\theta_1^0,\theta_2^0,\gamma^0\right)$ for all n. That is, $\left(\tilde{\theta}_1^n,\tilde{\theta}_2^n,\tilde{\gamma}^n\right)=\left(\theta_1^0,\theta_2^0,\gamma^0\right)$ for all n. In particular, $p_j=\mu(x_j^0)=\pi\left(L_{p_j},\theta_1^0,\theta_2^0\right)=$ $\pi\left(L_{p_j},\tilde{\theta}_1^n,\tilde{\theta}_2^n\right)$, which implies

$$\left|\pi\left(L_{p_j},\tilde{\theta}_1^n,\tilde{\theta}_2^n\right)-p_j\right|=0<\left|\pi\left(x_{ij},\tilde{\theta}_1^n,\tilde{\theta}_2^n\right)-p_j\right|$$

for j=1,2 and any $x_{ij}\neq L_{p_j}$. Therefore, when the true model is the logit model, $\left(\tilde{\theta}_1^n,\tilde{\theta}_2^n\right)\in S\left(\theta_1^0,\theta_2^0\right)$ for any n and condition C.1 holds.

Theorem 2. *Suppose the conditions of Theorem 1 are satisfied. Then*

$$\textit{asymptotically}\begin{bmatrix}\hat{\theta}_1^n-\theta_1^0\\ \hat{\theta}_2^n-\theta_2^0\\ \hat{\gamma}^n-\gamma^0\end{bmatrix}\sim N_3\left\{\begin{bmatrix}0\\0\\0\end{bmatrix},\mathbf{I}^{-1}\right\} \tag{22}$$

where $\mathbf{I}$ *is the information matrix. Define* $\pi_j=\pi(x_j^0,\theta_1^0,\theta_2^0)$. *Then*

$$\mathbf{I}=-\begin{bmatrix}i_{11}&i_{12}&i_{13}\\ i_{21}&i_{22}&i_{23}\\ i_{31}&i_{32}&i_{33}\end{bmatrix}\ \textit{where}$$

$$i_{11}=\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}\left(\frac{\pi_j^2(1-\pi_j)(\pi_j+s\gamma^0)+\pi_j(1-\pi_j)^2(\pi_j+s\gamma^0)-\pi_j^2(1-\pi_j)^2}{\left(\pi_j+s\gamma^0\right)^2}\right)\right\}$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\mathrm{E}\left\{\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}\left(\frac{\pi_j^2(1-\pi_j)(1-\pi_j+s\gamma^0)+\pi_j(1-\pi_j)^2(1-\pi_j+s\gamma^0)+\pi_j^2(1-\pi_j)^2}{(1-\pi_j+s\gamma^0)^2}\right)\right\}$$

$$i_{12}=\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}x_j^0\left(\frac{\pi_j^2(1-\pi_j)(\pi_j+s\gamma^0)+\pi_j(1-\pi_j)^2(\pi_j+s\gamma^0)-\pi_j^2(1-\pi_j)^2}{(\pi_j+s\gamma^0)^2}\right)\right\}$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\mathrm{E}\left\{\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}x_j^0\left(\frac{\pi_j^2(1-\pi_j)(1-\pi_j+s\gamma^0)+\pi_j(1-\pi_j)^2(1-\pi_j+s\gamma^0)+\pi_j^2(1-\pi_j)^2}{(1-\pi_j+s\gamma^0)^2}\right)\right\}$$

$i_{22}=$

$$\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\mathrm{E}\left\{\sum_{s=0}^{Y_{ijk}-1}(x_j^0)^2\left(\frac{\pi_j^2(1-\pi_j)(\pi_j+s\gamma^0)+\pi_j(1-\pi_j)^2(\pi_j+s\gamma^0)-\pi_j^2(1-\pi_j)^2}{(\pi_j+s\gamma^0)^2}\right)\right\}$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\mathrm{E}\left\{\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}(x_j^0)^2\left(\frac{\pi_j^2(1-\pi_j)(1-\pi_j+s\gamma^0)+\pi_j(1-\pi_j)^2(1-\pi_j+s\gamma^0)+\pi_j^2(1-\pi_j)^2}{(1-\pi_j+s\gamma^0)^2}\right)\right\}$$

$$i_{13}=\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\pi_j(1-\pi_j)\mathrm{E}\left\{-\sum_{s=0}^{Y_{ijk}-1}\frac{s}{(\pi_j+s\gamma^0)^2}+\sum_{s=0}^{r_{ijk}-Y_{ijk}-1}\frac{s}{\{(1-\pi_j)+s\gamma^0\}^2}\right\}$$

$$i_{23} = \sum_{i=1}^{n} \sum_{j=1}^{2}\sum_{k=1}^{m_{ij}} x_j^0 \pi_j (1-\pi_j) \mathrm{E}\left\{ -\sum_{s=0}^{Y_{ijk}-1} \frac{s}{(\pi_j + s\gamma^0)^2} + \sum_{s=0}^{r_{ijk}-Y_{ijk}-1} \frac{s}{\{(1-\pi_j)+s\gamma^0\}^2} \right\}$$

$$\text{and } i_{33} = \sum_{i=1}^{n} \sum_{j=1}^{2}\sum_{k=1}^{m_{ij}} \mathrm{E}\left\{ -\sum_{s=0}^{Y_{ijk}-1} \frac{s^2}{(\pi_j + s\gamma^0)^2} - \sum_{s=0}^{r_{ijk}-Y_{ijk}-1} \frac{s^2}{\{(1-\pi_j)+s\gamma^0\}^2} + \sum_{s=0}^{r_{ijk}-1} \frac{s^2}{\left(1+s\gamma^0\right)^2} \right\}.$$

Proof. From Theorem 1, $(x_{n+1,1}, x_{n+1,2})$ tends to $\left(x_1^0, x_2^0\right)$ almost surely. Since there are only t possible pairs of design points, it follows that $(x_{n+1,1}, x_{n+1,2}) \equiv \left(x_1^0, x_2^0\right)$ for large enough n. Therefore, asymptotically, the distribution of $\left(\hat{\theta}_1^n, \hat{\theta}_2^n, \hat{\gamma}^n\right)$is the same as the distribution of the solution to

$$\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}} \pi\left(x_j^0,\theta_1,\theta_2\right)\left(1-\pi\left(x_j^0,\theta_1,\theta_2\right)\right)\left\{\sum_{s=0}^{y_{ijk}-1}\left[\pi\left(x_j^0,\theta_1,\theta_2\right)+s\gamma\right]^{-1} - \sum_{s=0}^{r_{ijk}-y_{ijk}-1}\left(\left(1-\pi\left(x_j^0,\theta_1,\theta_2\right)\right)+s\gamma\right)^{-1}\right\}=0,$$

x_j^0

$$\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}} x_j^0\pi\left(x_j^0,\theta_1,\theta_2\right)\left(1-\pi\left(x_j^0,\theta_1,\theta_2\right)\right)\left\{\sum_{s=0}^{y_{ijk}-1}\left[\pi\left(x_j^0,\theta_1,\theta_2\right)+s\gamma\right]^{-1} - \sum_{s=0}^{r_{ijk}-y_{ijk}-1}\left(\left(1-\pi\left(x_j^0,\theta_1,\theta_2\right)\right)+s\gamma\right)^{-1}\right\} = 0$$

and

$$\sum_{i=1}^{n}\sum_{j=1}^{2}\sum_{k=1}^{m_{ij}}\left\{\sum_{s=0}^{y_{ijk}-1}\frac{s}{[\pi(x_j^0,\theta_1,\theta_2)+s\gamma]}+\sum_{s=0}^{r_{ijk}-y_{ijk}-1}\frac{s}{[1-\pi(x_j^0,\theta_1,\theta_2)]+s\gamma}-\sum_{s=0}^{r_{ijk}-1}\frac{s}{1+s\gamma}\right\}=0.$$

Since $(x_1^0, y_{111}), (x_2^0, y_{121}), \ldots, (x_1^0, y_{n1m_{n1}}), (x_2^0, y_{n2m_{n2}})$ are independently distributed samples, the asymptotic normality of $(\hat{\theta}_1^n, \hat{\theta}_2^n, \hat{\gamma}^n)$ follows the standard maximum likelihood estimator approach. The covariance matrix is the inverse of the information matrix and takes the form given in (22). This completes the proof.

Remark on γ: The proof of Theorem 1 assumes that γ is a fixed value for all dose levels. However, suppose γ is allowed to take on $g \leq d$ different values over the range of possible dose levels. The proof of Theorem 1 remains substantially unchanged except that instead of having three partial derivatives of the log-likelihood function (one with respect to each parameter: (θ_1, θ_2 and γ), 2+g partial derivatives are required (one for θ_1, one for θ_2 and g for $\gamma_1, \ldots\gamma_g$). Since $(x_{n,1}, x_{n,2}) \rightarrow (x_1^0, x_2^0)$ for all g, asymptotically at most 4 parameters will remain: θ_1, θ_2, γ_i, and γ_j.

CHAPTER 4

SIMULATION STUDY

4.1 *The Setup*

A simulation study was performed to examine the effect of intralitter correlation on the root estimates from the SAM procedure. Since SAM requires the existence of MLEs at each update, some other initial procedure is required to start SAM. The Moser-Fei (1991) procedure was used to generate the initial dose levels until the MLEs existed. Even when the MLEs exist, for small n, the estimates of $\hat{\theta}_1^{(n)}$ and $\hat{\theta}_2^{(n)}$ can be unreasonably small or large, causing unnecessarily large steps to be taken early in the procedure. When this happens, the procedure may not recover if the total number of updates is small. To avoid this problem, bounds have been set on $\hat{\theta}_1^{(n)}$ and $\hat{\theta}_2^{(n)}$. When the MLEs do not exist, the Moser-Fei procedure is used to obtain the next dose levels instead of SAM. Typically, Moser-Fei is used for the first one or two updates, and then SAM is used for the remaining updates. However, on occasion, the Moser-Fei procedure is used for all n updates. The average number of updates after which only SAM is used (*i.e.*, Moser-Fei is no longer used), and the percentage of times that the Moser-Fei procedure is used for all n updates have been reported.

Ten thousand simulation runs were performed for each of the 80 combinations of the following factors:

a) true model $\mu(x_{ij})$: logit or complementary log-log,

b) starting values: $(L_{.2}, L_{.8})$ or $(L_{.75}, L_{.95})$,

c) number of litters per dose group: $m_{ij} = 2$ or $m_{ij} = 5$,

d) number of updates: 10 or 20, and

e) intralitter correlation: .1, .3, .5, .7, or .9.

The specific form of the two models in a) are: the logit, $[1+\exp-(\theta_1+\theta_2 x_{ij})]^{-1}$, with θ_1=0 and θ_2=1; and the complementary log-log, $\{1-\exp[-\exp(\theta_1+\theta_2 x_{ij})]\}$, with θ_1=-.51 and θ_2=.71. Regardless of the true model, $\mu(x_{ij})$, the MLEs were calculated using the two-parameter logit model.

To begin each simulation run, two initial dose levels, x_{11} and x_{21} had to be selected. From b) note that these starting values were set at either 1) $x_{11}=L_{.2}$ and $x_{21}=L_{.8}$ or 2) $x_{11}=L_{.75}$ and $x_{21}=L_{.9}$. The numerical values of $L_{.2}$, $L_{.8}$, $L_{.75}$, and $L_{.9}$ depend on which model was used in a).

The litter size, r_{ijk}, is a random value from a truncated Poisson distribution with mean 12 and sample space {1, ..., 20}. According to Kupper *et al.* (1986), this distribution of litter sizes is generally representative of that encountered in some actual experimental situations.

After completion of the 10,000 runs for each of the 80 combinations of factors, estimates of $L_{.10}$, $L_{.25}$, $L_{.50}$, $L_{.75}$, and $L_{.90}$ and mean square errors (MSEs) were calculated.

4.2 *Results*

The results of the simulation are presented in Tables 2-5. Tables 2 and 3 display MSEs for estimates of $L_{.10}$, $L_{.25}$, $L_{.50}$, $L_{.75}$, and $L_{.90}$ when the true model is the logit. Starting values of $L_{.20}$ and $L_{.80}$ were used to obtain the values in Table 2, while starting values of $L_{.75}$ and $L_{.95}$ were used in Table 3. Tables 4 and 5 display the corresponding results when the true model is the complementary log-log.

For all of the tables (2-5), the following effects are noticed:

a) For a fixed number of litters per dose and a fixed number of updates, the MSEs increase as the intralitter correlation increases.

b) For a fixed number of litters per dose and a fixed correlation value, the MSEs decrease as the number of updates increase.

c) For a fixed number of updates and a fixed correlation value, the MSEs decrease as the number of litters per dose increase.

d) The MSEs tend to be smaller for the logit model than for the complementary log-log.

Result a) is expected. Note that $\mathrm{Var}(Y_{ijk}|x_{ij})=[1+\{(n-1)\gamma/(1+\gamma)\}]n\mu_{ij}(1-\mu_{ij})$ which is an increasing function of the correlation, $\gamma/(1+\gamma)$. Now if asymptotically the variance of $\hat{L}_{p*}$ for this SAM procedure operates like the variance of $\hat{L}_{p*}$ for the optimal two dimensional Robbins-Monro procedure, then

$$\mathrm{Var}(\hat{L}_{p*})=\mathrm{Var}(k\hat{L}_{p_1}+(1\text{-}k)\,\hat{L}_{p_2})$$

where

$$\mathrm{Var}(\hat{L}_{p_j})=c_j^{-2}\,\mathrm{Var}(Y_{ijk}\mid x_{ij})/n(2c_j^{-1}\mu'(L_{p_j})-1),\qquad j=1,2$$

where k and c_j do not depend on $\mathrm{Var}(Y_{ijk}|x_{ij})$. Therefore, as the variance of Y_{ijk} given x_{ij} increases, the variance of $\hat{L}_{p*}$ increases for $0<p^*<1$; as the variance of $\hat{L}_{p*}$ increases, the MSE increases since the $\mathrm{MSE}(\hat{L}_{p*})=\mathrm{Var}(\hat{L}_{p*})+[\mathrm{Bias}(\hat{L}_{p*})]^2$. Thus, asymptotically it is expected that the MSE increases linearly as a function of the intralitter correlation. It is not surprising, therefore, that the effect is evident for smaller n. Figures 1 and 2 display evidence of the linear increase in the MSEs of $L_{.50}$ and as a function of the correlation for both the logit and the complementary log-log when n=10 and n=20. Similar patterns are evident for the MSEs of $L_{.10}$, $L_{.25}$, $L_{.75}$, and $L_{.90}$.

The reason the MSEs decrease in b) and c) is that more data is being used to predict the roots. Effect d) occurs because the logit model is being used to calculate the MLEs. Therefore, when the true model *is* the logit model, the bias in the MSEs is zero. When the true model is actually the complementary log-log, the squared bias in the MSEs is non-zero, resulting in larger MSEs for the log-log relative to the logit.

For a starting set of values of $L_{.20}$ and $L_{.80}$, the average number of updates for the SAM procedure to take over the Moser-Fei procedure was between 1 and 3.5. For a starting set of values of $L_{.75}$ and $L_{.95}$, the average number of updates before the SAM procedure took over was slightly larger (between 1 and 5.4). Moser-Fei is used for more updates when the starting values are high because for high dose levels all of the responses tend to be positive, and the MLEs do not exist until some of the responses are not positive.

The number of times that Moser-Fei was used for all n updates is less than 1% in all but one case. The exception was when the correlation was .9, the number of updates was 10, and the starting values were $L_{.75}$ and $L_{.95}$. In this worst-case scenario (correlation=.9, number of updates=10 and starting values of $L_{.75}$ and $L_{.95}$), the procedure stayed in the Moser-Fei procedure for all 10 updates 7.34% and 5.67% of the time for the logit and complementary log-log model, respectively. The percentages were higher for this set of conditions because when the correlation is close to one, the response for a given litter tends to be either zero or r_{ijk}. Since the MLEs could not be calculated until at least one of the responses was equal to a value other than zero or r_{ijk}, the Moser-Fei procedure was used for relatively more updates.

Table 2

Summary of simulations for the Logit Model with starting values $L_{.20}$ and $L_{.80}$

# of litters	# of Up-dates	corre-lation	MSE for $L_{.10}$	MSE for $L_{.25}$	MSE for $L_{.50}$	MSE for $L_{.75}$	MSE for $L_{.90}$	Avg starting update number for SAM	% times Moser-Fei is used for all *n* up-dates
2	10	.1	.10158	.04603	.02707	.04470	.09891	1.08	0.00
		.3	.20263	.08908	.05059	.08716	.19879	1.31	0.00
		.5	.29106	.12747	.07152	.12322	.28255	1.63	0.00
		.7	.39695	.17241	.09938	.17785	.40784	2.08	0.00
		.9	.56539	.25129	.15076	.26380	.59040	3.23	2.23
2	20	.1	.04883	.02206	.01317	.02214	.04898	1.08	0.00
		.3	.09949	.04363	.02505	.04375	.09973	1.31	0.00
		.5	.14664	.06339	.03529	.06235	.14457	1.64	0.00
		.7	.19332	.08413	.04828	.08579	.19665	2.10	0.00
		.9	.26593	.11673	.06715	.11719	.26685	3.49	0.50
5	10	.1	.04063	.01832	.01067	.01766	.03932	1.00	0.00
		.3	.07803	.03407	.01942	.03407	.07802	1.01	0.00
		.5	.11508	.04979	.02748	.04815	.11181	1.09	0.00
		.7	.15148	.06688	.03847	.06625	.15021	1.25	0.00
		.9	.20451	.09168	.05329	.08934	.19984	1.85	0.10
5	20	.1	.01976	.00884	.00522	.00888	.01984	1.00	0.00
		.3	.03899	.01713	.00976	.01689	.03851	1.02	0.00
		.5	.05549	.02377	.01335	.02425	.05647	1.08	0.00
		.7	.07691	.03349	.01868	.03246	.07485	1.26	0.00
		.9	.09818	.04332	.02529	.04411	.09975	1.88	0.00

Table 3

Summary of simulations for the Logit Model with starting values $L_{.75}$ and $L_{.95}$

# of litters	# of Up-dates	corre-lation	MSE for $L_{.10}$	MSE for $L_{.25}$	MSE for $L_{.50}$	MSE for $L_{.75}$	MSE for $L_{.90}$	Avg starting update number for SAM	% times Moser-Fei is used for all *n* up-dates
2	10	.1	.12598	.05907	.03015	.03924	.08633	1.81	0.00
		.3	.26276	.12279	.05959	.07315	.16347	2.51	0.00
		.5	.41802	.19711	.09164	.10159	.22699	3.13	0.03
		.7	.62643	.29987	.13724	.13855	.30378	3.74	0.39
		.9	1.11743	.56059	.26764	.23858	.47341	4.54	7.34
2	20	.1	.05355	.02476	.01375	.02050	.04501	1.83	0.00
		.3	.10763	.04836	.02532	.03851	.08792	2.48	0.00
		.5	.16167	.07290	.03691	.05372	.12332	3.11	0.00
		.7	.22893	.10414	.05286	.07507	.17079	3.81	0.00
		.9	.33502	.15464	.07454	.09473	.21520	5.34	0.28
5	10	.1	.04620	.02124	.01126	.01626	.03625	1.17	0.00
		.3	.08864	.04031	.02094	.03055	.06913	1.46	0.00
		.5	.13740	.06303	.03126	.04209	.09551	1.83	0.00
		.7	.20396	.09565	.04662	.05685	.12636	2.37	0.00
		.9	.29832	.14140	.06708	.07534	.16619	3.24	0.16
5	20	.1	.02091	.00948	.00535	.00852	.01899	1.16	0.00
		.3	.04038	.01823	.01006	.01586	.03562	1.46	0.00
		.5	.05926	.02629	.01392	.02214	.05095	1.81	0.00
		.7	.08160	.03670	.01920	.02913	.06647	2.33	0.00
		.9	.11578	.05300	.02710	.03806	.08590	3.29	0.00

Table 4

Summary of simulations for the Complementary Log-log Model with starting values $L_{.20}$ and $L_{.80}$

# of litters	# of Up-dates	corre-lation	MSE for $L_{.10}$	MSE for $L_{.25}$	MSE for $L_{.50}$	MSE for $L_{.75}$	MSE for $L_{.90}$	Avg starting update number for SAM	% times Moser-Fei is used for all updates
2	10	.1	.22755	.06862	.06842	.03566	.13384	1.08	0.00
		.3	.37015	.12084	.08986	.06396	.17828	1.32	0.00
		.5	.50783	.17329	.11256	.09115	.21715	1.65	0.00
		.7	.65154	.22893	.13688	.12545	.28276	2.10	0.00
		.9	.85867	.32230	.18383	.18759	.41223	3.22	2.46
2	20	.1	.14972	.03596	.05286	.01971	.11283	1.08	0.00
		.3	.22957	.06675	.06513	.03305	.13287	1.31	0.00
		.5	.30708	.09333	.07501	.04572	.14842	1.64	0.00
		.7	.38842	.12240	.08600	.06076	.17393	2.09	0.00
		.9	.47959	.16120	.10352	.08109	.21108	3.48	0.80
5	10	.1	.12651	.03044	.05210	.01728	.11155	1.00	0.00
		.3	.18853	.05111	.05987	.02779	.12359	1.02	0.00
		.5	.25186	.07297	.06730	.03818	.14103	1.08	0.00
		.7	.31056	.09455	.07621	.04790	.15085	1.25	0.00
		.9	.37228	.12669	.09295	.06376	.18136	1.85	0.20
5	20	.1	.09496	.01757	.04676	.01165	.10232	1.00	0.00
		.3	.12801	.02892	.05082	.01662	.10809	1.01	0.00
		.5	.16055	.03906	.05367	.02104	.11416	1.08	0.00
		.7	.19088	.05257	.05943	.02607	.12299	1.25	0.00
		.9	.23380	.06694	.06558	.03451	.13113	1.86	0.00

Table 5

Summary of simulations for the Complementary Log-log Model with starting values $L_{.75}$ and $L_{.95}$

# of litters	# of Up-dates	corre-lation	MSE for $L_{.10}$	MSE for $L_{.25}$	MSE for $L_{.50}$	MSE for $L_{.75}$	MSE for $L_{.90}$	Avg starting update number for SAM	% times Moser-Fei is used for all *n* up-dates
2	10	.1	.35804	.09710	.06676	.03543	.10971	1.81	0.00
		.3	.60300	.19088	.10068	.06470	.14298	2.52	0.00
		.5	.86060	.29544	.13644	.08809	.17308	3.13	0.00
		.7	1.14512	.41607	.17820	.11392	.21516	3.77	0.40
		.9	1.66031	.67960	.30557	.21032	.37166	4.71	5.67
2	20	.1	.19100	.04345	.05227	.02060	.10251	1.83	0.00
		.3	.29416	.07994	.06639	.03518	.11304	2.50	0.00
		.5	.39974	.11339	.07634	.04752	.12288	3.12	0.00
		.7	.52118	.16380	.09389	.06030	.14596	3.81	0.00
		.9	.68405	.23054	.11684	.07684	.17252	5.35	0.13
5	10	.1	.19220	.03558	.04554	.01739	.09297	1.17	0.00
		.3	.28947	.06742	.05533	.02832	.10176	1.46	0.00
		.5	.38012	.10333	.06912	.03964	.11384	1.83	0.00
		.7	.48544	.14587	.08446	.05131	.12967	2.36	0.00
		.9	.64651	.21302	.10790	.06461	.14341	3.25	0.10
5	20	.1	.12017	.01714	.04236	.01152	.09476	1.16	0.00
		.3	.16474	.03103	.04666	.01706	.09862	1.45	0.00
		.5	.20097	.04402	.05176	.02204	.10191	1.81	0.00
		.7	.24670	.06052	.05837	.02838	.10524	2.35	0.00
		.9	.30735	.08443	.06701	.03514	.11255	3.27	0.00

Figure 1

MSEs for Estimating $L_{.50}$ when $\mu(x)$ is the Logit Model
Starting values $L_{.75}$ and $L_{.95}$

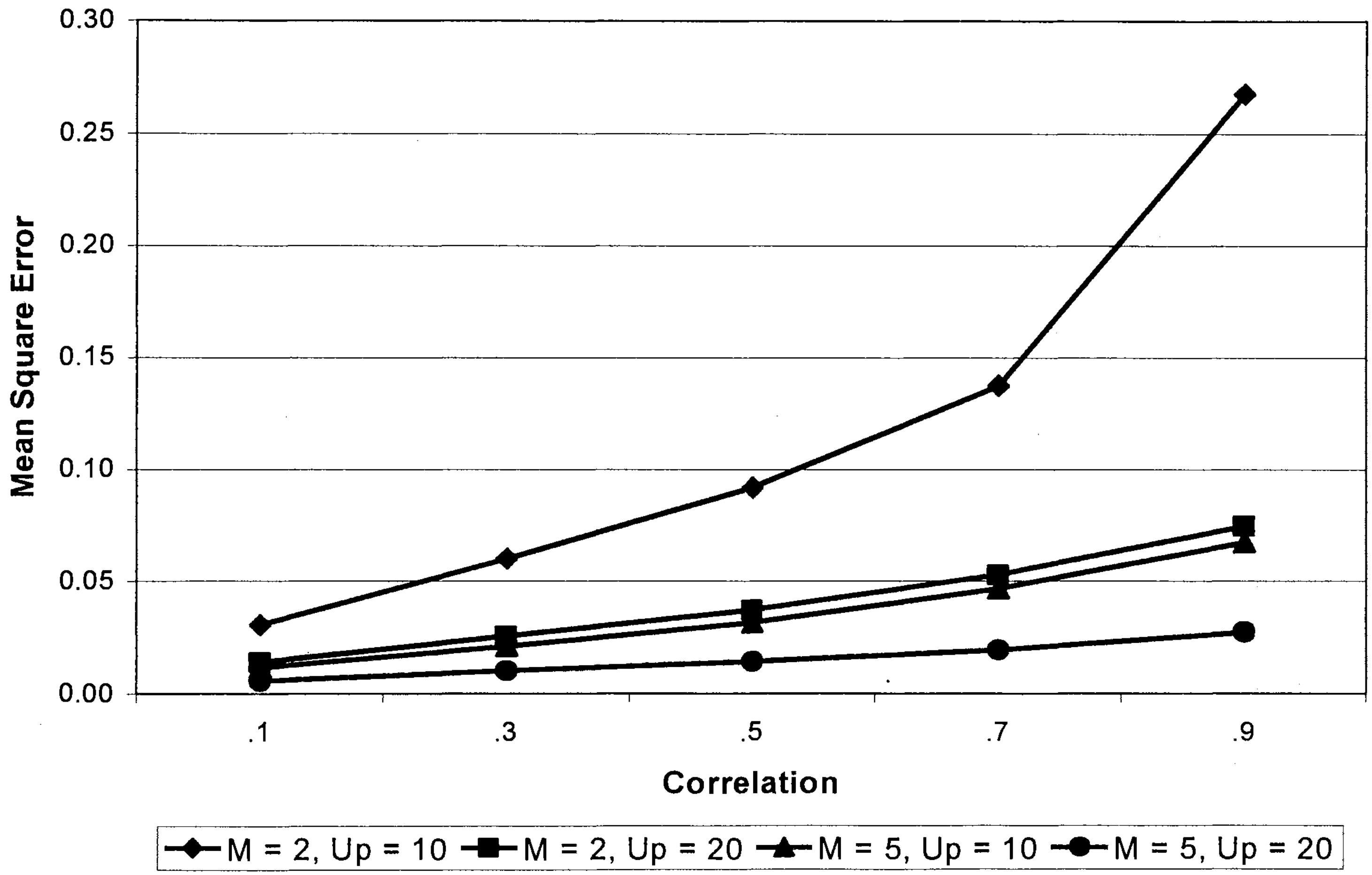
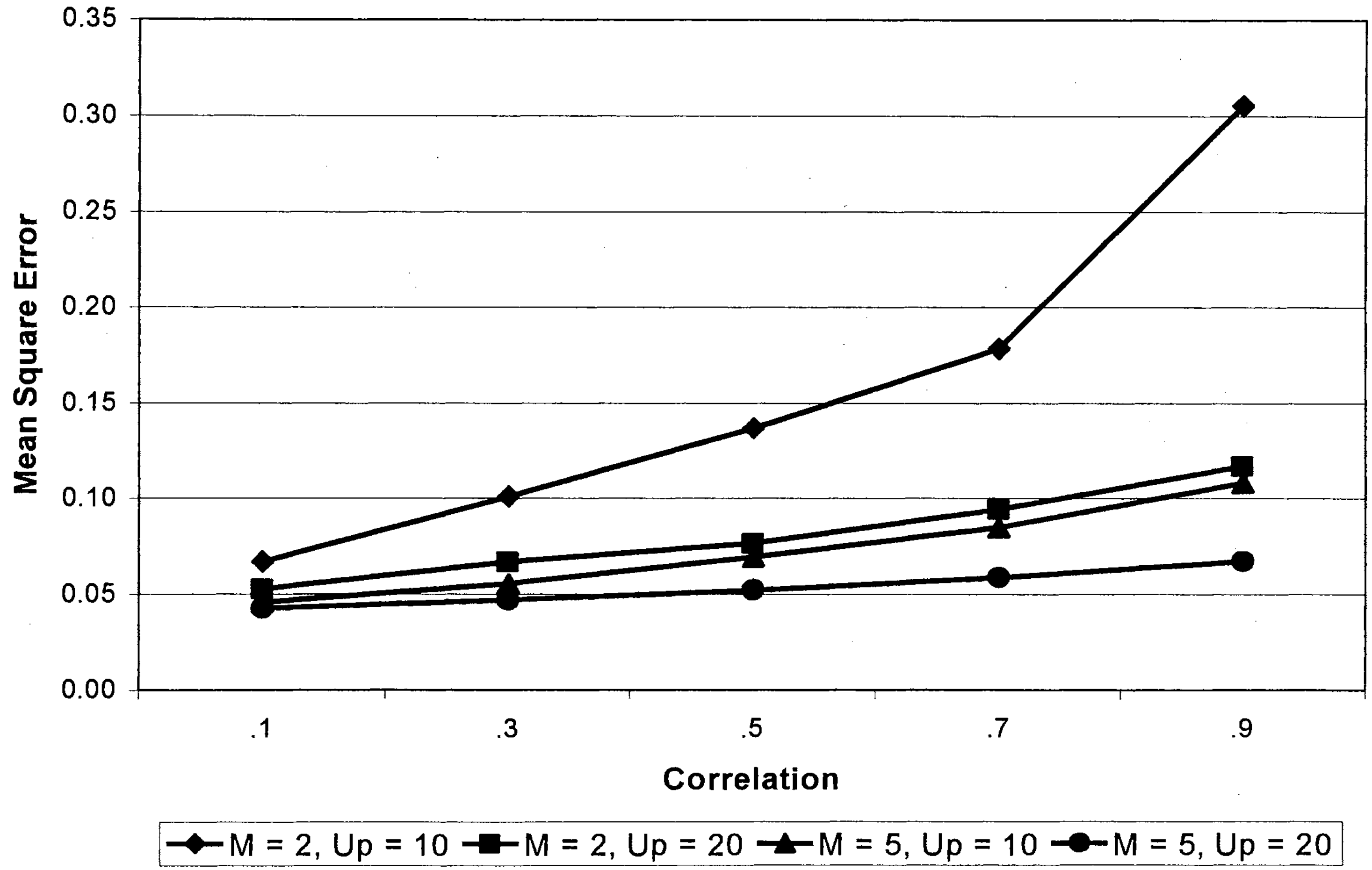


Figure 2

MSEs for Estimating $L_{.50}$ when $\mu(x)$ is the Complementary Log-log Model
Starting values $L_{.75}$ and $L_{.95}$



CHAPTER 5

CONCLUSION

If the SAM procedure for the two-parameter logit model is used to estimate dose levels when the responses within a litter follow a beta-binomial distribution, the MLEs of the three parameters (θ_1, θ_2, and γ) are consistent and asymptotically normal. The observed dose levels are consistent for a discrete set of dose levels. Furthermore, as noted in the appendix, the results can be extended to situations in which the overdispersion parameter is allowed to vary across dose levels. The simulation results support the asymptotic results, indicating that the effective amount of information decreases (*i.e.*, the MSE increases) as the correlation within a litter increases.

Since the assumption that the responses within a litter follow a beta-binomial distribution may not be appropriate, further research should be directed at examining a broader class of distributions to model the response variable. Possible alternatives include the Probit-normal-binomial and the Logistic-normal-binomial.

BIBLIOGRAPHY

Anbar, D. (1978), "A Stochastic Newton-Raphson Method," *Journal of Statistical Planning and Inference*,2,153-164.

Blum, J.R. (1954), "Approximation Methods Which Converge With Probability One," *Annals of Mathematical Statistics*, 25, 382-386.

Chung, K.L. (1954), "On a Stochastic Approximation Method," *Annals of Mathematical Statistics*, 25, 463-483.

Dixon, W. J. and Mood, A. M. (1948), "A Method for Obtaining and Analyzing Sensitivity Data," *Journal of the American Statistical Association*, 43, 109-126.

Haseman, J. K., and Kupper, L.L. (1979), "Analysis of Dichotomous Response Data from Certain Toxicological Experiments," *Biometrics*, 35, 281-293.

Kupper, L.L., Portier, C.J., Hogan, M.D., and Yamamoto, E. (1986), "The impact of litter effects on dose-response modeling in teratology," *Biometrics*, 42, 85-98.

Moser, B.K., and Faries, D. (1996), "A New Sequential Approximation Method," *Journal of Statistical Planning and Inference*, 55, 175-189.

Moser, B.K. and Fei, W. (1991), "A New Sequential Design Based on the Robbins-Monro Procedure," *Metrika*, 38, 321-333.

Piegorsch, W.W. (1993), "Biometrical Methods for Testing Dose Effects of Environmental Stimuli in Laboratory Studies," *Environmetrics*, 4, 483-505.

Robbins, H. and Monro, S. (1951), "A Stochastic Approximation Method," *Annals of Mathematical Statistics*, 38, 181-190.

Shen, L. Z. and O'Quigley, J. (1996), "Consistency of Continual Reassessment Method Under Model Misspecification," *Biometrika*, 83, 2, 395-405.

Shiryayev, A.N. (1984), *Probability*, (Trans, R.P. Boas), New York: Spriger-Verlag.

Storer, B.E. (1989), "Design and analysis of Phase I clinical trials," *Biometrics*, 45, 925-937.

Williams, D. A. (1975), "The analysis of binary responses from toxicological experiments involving reproduction and teratogenicity," *Biometrics*, 31,457-462.

Wu, C. F. J. (1985), "Efficient Sequential Designs With Binary Data," *Journal of the American Statistical Society*, 80,974-983.

APPENDIX A

Lemma 1. *If Y~BB(r,p,γ) then*

$$\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{1}{p+s\gamma}\right)=\mathrm{E}\left(\sum_{s=0}^{r-Y-1}\frac{1}{1-p+s\gamma}\right).$$

Proof.

Prove $\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{1}{p+s\gamma}\right)=\sum_{s=0}^{r-1}\frac{1}{1+s\gamma}$ by induction.

a) Show for r=2.

$$\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{1}{p+s\gamma}\right)=\left(\frac{1}{p}\right)\mathrm{P}(Y=1|r=2)+\left(\frac{1}{p}+\frac{1}{p+\gamma}\right)\mathrm{P}(Y=2|r=2)$$

$$=\frac{\left(\frac{1}{p}\right)\binom{2}{1}\left(\frac{p}{\gamma}\right)\left(\frac{1-p}{\gamma}\right)}{\left(1+\frac{1}{\gamma}\right)\frac{1}{\gamma}}+\frac{\left(\frac{1}{p}+\frac{1}{p+\gamma}\right)\left(\frac{p}{\gamma}+1\right)\frac{p}{\gamma}}{\left(1+\frac{1}{\gamma}\right)\frac{1}{\gamma}}$$

$$=\frac{2(1-p)}{1+\gamma}+\frac{2p+\gamma}{1+\gamma}=\frac{2+\gamma}{1+\gamma}=1+\frac{1}{\gamma}$$

$$=\sum_{s=0}^{2-1}\frac{1}{1+s\gamma}$$

b) Assume $\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{1}{p+s\gamma}\right)=\sum_{s=0}^{r-1}\frac{1}{1+s\gamma}$ when Y~$BB(r,p,\gamma)$.

c) Show $E\left(\sum_{s=0}^{Y-1}\frac{1}{p+s\gamma}\right)=\sum_{s=0}^{r}\frac{1}{1+s\gamma}$ when $Y\sim BB(r+1,p,\gamma)$.

First note that $P(Y=y|r+1,p,\gamma)=\left\{\frac{r-y+\frac{1-p}{\gamma}}{r+\frac{1}{\gamma}}\right\}P(Y=y|r,p,\gamma)$

$$=\left\{1+\frac{y(1-\gamma)-p(1+r)}{(1+r\gamma)(r+1-y)}\right\}P(Y=y|r,p,\gamma)$$

where y=0,…r.

So for $Y\sim BB(r+1,p,\gamma)$

$$E\left(\sum_{s=0}^{Y-1}\frac{1}{p+s\gamma}\right)=\left(\frac{1}{p}\right)P(Y=1|r+1,p,\gamma)+\left(\frac{1}{p}+\frac{1}{p+\gamma}\right)P(Y=2|r+1,p,\gamma)+$$

$$\left(\frac{1}{p}+\frac{1}{p+\gamma}+\frac{1}{p+2\gamma}\right)P(Y=3|r+1,p,\gamma)+\ldots+$$

$$\left(\frac{1}{p}+\frac{1}{p+\gamma}+\frac{1}{p+2\gamma}+\cdots+\frac{1}{p+(r-1)\gamma}\right)P(Y=r|r+1,p,\gamma)+$$

$$\left(\frac{1}{p}+\frac{1}{p+\gamma}+\frac{1}{p+2\gamma}+\cdots+\frac{1}{p+r\gamma}\right)P(Y=r+1|r+1,p,\gamma)$$

$$=\frac{1}{p}\left\{1+\frac{(1-\gamma-p(1+r)}{(1+r\gamma)r}\right\}\mathrm{P}(Y=1|r,p,\gamma)+$$

$$\left(\frac{1}{p}+\frac{1}{p+\gamma}\right)\left\{1+\frac{(2(1-\gamma)-p(1+r)}{(1+r\gamma)(r-1)}\right\}\mathrm{P}(Y=2|r,p,\gamma)+$$

$$\left(\frac{1}{p}+\frac{1}{p+\gamma}+\frac{1}{p+2\gamma}+\frac{1}{p+3\gamma}\right)\left\{1+\frac{(3(1-\gamma)-p(1+r)}{(1+r\gamma)(r-2)}\right\}\mathrm{P}(Y=3|r,p,\gamma)+\cdots+$$

$$\left(\frac{1}{p}+\frac{1}{p+\gamma}+\cdots+\frac{1}{p+(r-1)\gamma}\right)\left\{1+\frac{(r(1-\gamma)-p(1+r)}{(1+r\gamma)}\right\}\mathrm{P}(Y=r|r,p,\gamma)+$$

$$\left(\frac{1}{p}+\frac{1}{p+\gamma}+\cdots+\frac{1}{p+r\gamma}\right)\mathrm{P}(Y=r+1|r+1,p,\gamma).$$

from b)

$$=\sum_{s=0}^{r-1}\frac{1}{1+s\gamma}+\left(\frac{1}{1+r\gamma}\right)\left[\frac{1}{p}\left\{\frac{(1-\gamma-p(1+r)}{r}\right\}\mathrm{P}(Y=1|r,p,\gamma)+\right.$$

$$\left(\frac{1}{p}+\frac{1}{p+\gamma}\right)\left\{\frac{(2(1-\gamma)-p(1+r)}{(r-1)}\right\}\mathrm{P}(Y=2|r,p,\gamma)+$$

$$\left(\frac{1}{p}+\frac{1}{p+\gamma}+\frac{1}{p+2\gamma}+\frac{1}{p+3\gamma}\right)\left\{\frac{(3(1-\gamma)-p(1+r)}{(r-2)}\right\}\mathrm{P}(Y=3|r,p,\gamma)+\cdots+$$

$$\left(\frac{1}{p}+\frac{1}{p+\gamma}+\cdots+\frac{1}{p+(r-1)\gamma}\right)\left\{\frac{(r(1-\gamma)-p(1+r)}{1}\right\}\mathrm{P}(Y=r|r,p,\gamma)+$$

$$\left.\left(\frac{1}{p}+\frac{1}{p+\gamma}+\cdots+\frac{1}{p+r\gamma}\right)\frac{(p+r\gamma)(p+(r-1)\gamma)\cdots p}{(1+(r-1)\gamma)(1+(r-2)\gamma)\cdots(1+\gamma)}\right]. \tag{1}$$

It remains to show that the term in square brackets is equal to one. To this end, we will first prove the following:

$$\mathrm{P}(Y=s|r)=\frac{1}{p+s\gamma}\left\{\sum_{y=s+1}^{r}\left[\frac{y(1-\gamma)-p(1+r)}{1+r-y}\mathrm{P}(Y=y|r,p,\gamma)\right]+(p+s\gamma)\mathrm{P}(Y=r|r,p,\gamma)\right\}. \tag{2}$$

The term on the right hand side can be rewritten in the following way:

$$\frac{1}{p+s\gamma}\left\{\sum_{y=s+1}^{r}\left[\frac{y(1-\gamma)-p(1+r)}{1+r-y}\mathrm{P}(Y=y|r,p,\gamma)\right]+(p+s\gamma)\mathrm{P}(Y=r|r,p,\gamma)\right\}$$

$$=\frac{1}{p+s\gamma}\left\{\sum_{y=s+1}^{r-1}\left[\frac{y(1-\gamma)-p(1+r)}{1+r-y}\mathrm{P}(Y=y|r,p,\gamma)\right]+r(1-p)\mathrm{P}(Y=r|r,p,\gamma)\right\}$$

$$=\frac{1}{p+s\gamma}\left\{\sum_{y=s+1}^{r-1}\left[(1+\mathrm{r}\gamma)(\mathrm{P}(Y=y|r+1,p,\gamma)-\mathrm{P}(Y=y|r,p,\gamma))\right]+r(1-p)\mathrm{P}(Y=r|r,p,\gamma)\right\}$$

$$=\frac{1+r\gamma}{p+s\gamma}\left\{1-\left[\sum_{y=0}^{s}(\mathrm{P}(Y=y|r+1)+\mathrm{P}(Y=r|r+1)+\mathrm{P}(Y=r+1|r+1))\right]-\right.$$
$$\left.\left[1-\left[\sum_{y=0}^{s}\mathrm{P}(Y=y|r)+\mathrm{P}(Y=r|r)\right]\right]+\frac{r(1-p)}{1+r\gamma}\mathrm{P}(Y=r|r,p,\gamma)\right\}$$

$$=\frac{1+r\gamma}{p+s\gamma}\left\{\begin{array}{l}-\sum_{y=0}^{s}(\mathrm{P}(Y=y|r+1)-\mathrm{P}(Y=r|r+1)-\mathrm{P}(Y=r+1|r+1))\\+\sum_{y=0}^{s}\mathrm{P}(Y=y|r)+\mathrm{P}(Y=r|r)\left[1+\frac{r(1-p)}{1+r\gamma}\right]\end{array}\right\}$$

$$=\frac{1+r\gamma}{p+s\gamma}\left\{\begin{array}{l}-\sum_{y=0}^{s}\left[\mathrm{P}(Y=y|r)\left(1+\frac{y(1-\gamma)-p(1+r)}{(1+r-y)(1+r\gamma)}\right)\right]-\mathrm{P}(Y=r|r)\left(1+\frac{r(1-\gamma)-p(1+r)}{(1+r\gamma)}\right)\\-\mathrm{P}(Y=r|r)\left(\frac{p+r\gamma)}{1+r\gamma}\right)+\sum_{y=0}^{s}\mathrm{P}(Y=y|r)+\mathrm{P}(Y=r|r)\left[1+\frac{r(1-p)}{1+r\gamma}\right]\end{array}\right\}$$

$$=\frac{1+r\gamma}{p+s\gamma}\left\{-\sum_{y=0}^{s}\mathrm{P}(Y=y|r)\frac{y(1-\gamma)-p(1+r)}{(1+r-y)(1+r\gamma)}\right\}$$

$$=\frac{1}{p+s\gamma}\left\{-\sum_{y=0}^{s}\mathrm{P}(Y=y|r)\frac{y(1-\gamma)-p(1+r)}{(1+r-y)}\right\}$$

Induction will be used to show that the last term is equal to $\mathrm{P}(Y=s|r)$. That is, we will prove that:

$$\frac{1}{p+s\gamma}\left\{-\sum_{y=0}^{s}\mathrm{P}(Y=y|r)\frac{y(1-\gamma)-p(1+r)}{(1+r-y)}\right\}=\mathrm{P}(Y=s|r).$$

i) Show for s=0.

$$\frac{1}{p}\left(-\sum_{y=0}^{0}\mathrm{P}(Y=y|r)\frac{y(1-\gamma)-p(1+r)}{(r+1-y)}\right)$$
$$=\frac{1}{p}\frac{\mathrm{P}(Y=0|r)p(1+r)}{r+1}$$
$$=\mathrm{P}(Y=0|r)$$

ii) Assume $\frac{1}{p+s\gamma}\left\{-\sum_{y=0}^{s}\mathrm{P}(Y=y|r)\frac{y(1-\gamma)-p(1+r)}{(1+r-y)}\right\}=\mathrm{P}(Y=s|r)$.

iii) Show $\frac{1}{p+(s+1)\gamma}\left\{-\sum_{y=0}^{s+1}\mathrm{P}(Y=y|r)\frac{y(1-\gamma)-p(1+r)}{(1+r-y)}\right\}=\mathrm{P}(Y=s+1|r)$.

$$\frac{1}{p+(s+1)\gamma}\left\{-\sum_{y=0}^{s+1}\mathrm{P}(Y=y|r)\frac{y(1-\gamma)-p(1+r)}{(1+r-y)}\right\}$$

by ii) is equal to

$$\frac{1}{p+(s+1)\gamma}\left\{\mathrm{P}(Y=s|r)(p+s\gamma)-\mathrm{P}(Y=s+1|r))\frac{(s+1)(1-\gamma)-p(1+r)}{r-s}\right\}$$

$$= \frac{1}{p+(s+1)\gamma}\left\{\begin{array}{l}\dfrac{\binom{r}{s+1}\left(\dfrac{s+1}{r-s}\right)\Gamma\left(\dfrac{p}{\gamma}+s\right)\Gamma\left(r-s+\dfrac{1-p}{\gamma}\right)\Gamma\left(\dfrac{1}{\gamma}\right)}{\Gamma\left(\dfrac{p}{\gamma}\right)\Gamma\left(\dfrac{1-p}{\gamma}\right)\Gamma\left(r+\dfrac{1}{\gamma}\right)}(p+s\gamma)- \\ \dfrac{\binom{r}{s+1}\Gamma\left(\dfrac{p}{\gamma}+s+1\right)\Gamma\left(r-(s+1)+\dfrac{1-p}{\gamma}\right)\Gamma\left(\dfrac{1}{\gamma}\right)}{\Gamma\left(\dfrac{p}{\gamma}\right)\Gamma\left(\dfrac{1-p}{\gamma}\right)\Gamma\left(r+\dfrac{1}{\gamma}\right)}\dfrac{(s+1)(1-\gamma)-p(1+r)}{r-s}\end{array}\right\}$$

$$= \frac{1}{p+(s+1)\gamma}\left\{\begin{array}{l}\dfrac{\binom{r}{s+1}\Gamma\left(\dfrac{p}{\gamma}+s\right)\Gamma\left(r-(s+1)+\dfrac{1-p}{\gamma}\right)\Gamma\left(\dfrac{1}{\gamma}\right)}{\Gamma\left(\dfrac{p}{\gamma}\right)\Gamma\left(\dfrac{1-p}{\gamma}\right)\Gamma\left(r+\dfrac{1}{\gamma}\right)}(p+s\gamma) \\ \left[\dfrac{\left(r-(s+1)+\dfrac{1-p}{\gamma}\right)(p+s\gamma)(s+1)}{r-s}-\dfrac{\left(\dfrac{p}{\gamma+s}\right)(r-s)[p+(s+1)\gamma]}{r-s}\right]\end{array}\right\}$$

$$= \frac{1}{p+(s+1)\gamma}\left\{\frac{\binom{r}{s+1}\Gamma\left(\frac{p}{\gamma}+s\right)\Gamma\left(r-(s+1)+\frac{1-p}{\gamma}\right)\Gamma\left(\frac{1}{\gamma}\right)}{\Gamma\left(\frac{p}{\gamma}\right)\Gamma\left(\frac{1-p}{\gamma}\right)\Gamma\left(r+\frac{1}{\gamma}\right)}\left[\frac{\left(\frac{p}{\gamma}+s\right)(r-s)[p+(s+1)\gamma]}{r-s}\right]\right\}$$

$$= \frac{\binom{r}{s+1}\Gamma\left(\frac{p}{\gamma}+s+1\right)\Gamma\left(r-(s+1)+\frac{1-p}{\gamma}\right)\Gamma\left(\frac{1}{\gamma}\right)}{\Gamma\left(\frac{p}{\gamma}\right)\Gamma\left(\frac{1-p}{\gamma}\right)\Gamma\left(r+\frac{1}{\gamma}\right)}$$

$$= \mathrm{P}(Y = s+1 \mid r).$$

This proves (2). Applying this equality to (1) results in the following equation:

$$\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{1}{p+s\gamma}\right) = \sum_{s=0}^{r-1}\frac{1}{1+s\gamma}+\left(\frac{1}{1+r\gamma}\right)\sum_{s=0}^{r}\mathrm{P}(Y=s|r)$$

$$=\sum_{s=0}^{r-1}\frac{1}{1+s\gamma}+\left(\frac{1}{1+r\gamma}\right)=\sum_{s=0}^{(r+1)-1}\frac{1}{1+s\gamma}.$$

Therefore, by induction, $\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{1}{p+s\gamma}\right)=\sum_{s=0}^{r-1}\frac{1}{1+s\gamma}$ when $Y\sim BB(r,p,\gamma)$. Now let $Z=r-Y$.

Then $Z\sim BB(r,1\text{-}p,\gamma)$ and $\mathrm{E}\left(\sum_{s=0}^{Z-1}\frac{1}{1-p+s\gamma}\right)=\sum_{s=0}^{r-1}\frac{1}{1+s\gamma}$.

APPENDIX B

Lemma 2. *If Y~BB(r,p,γ) then*

$$\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{s}{p+s\gamma}+\sum_{s=0}^{r-Y-1}\frac{s}{1-p+s\gamma}\right)=\sum_{s=0}^{r-1}\frac{s}{1+s\gamma}.$$

Proof.

By induction:

a) Show for r=2.

$$\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{s}{p+s\gamma}\right)=0\cdot\mathrm{P}(Y=1\mid r=2)+(0+\frac{1}{p+\gamma})\mathrm{P}(Y=2\mid r=2)$$

$$=\frac{1}{p+\gamma}\left(\frac{(p+\gamma)p}{1+\gamma}\right)$$

$$=\frac{p}{1+\gamma}$$

$$\mathrm{E}\left(\sum_{s=0}^{r-Y-1}\frac{s}{1-p+s\gamma}\right)=(0+\frac{1}{1-p+\gamma})\mathrm{P}(Y=0\mid r=2)+0\cdot\mathrm{P}(Y=1\mid r=2)$$

$$=\frac{1}{1-p+\gamma}\left(\frac{(1-p+\gamma)(1-p)}{1+\gamma}\right)$$

$$=\frac{1-p}{1+\gamma}$$

so $\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{s}{p+s\gamma}+\sum_{s=0}^{r-Y-1}\frac{s}{1-p+s\gamma}\right)=\frac{p}{1+\gamma}+\frac{1-p}{1+\gamma}=\frac{1}{1+\gamma}=\sum_{s=0}^{2-1}\frac{s}{1+s\gamma}.$

b) Assume $\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{s}{p+s\gamma}+\sum_{s=0}^{r-Y-1}\frac{s}{1-p+s\gamma}\right)=\sum_{s=0}^{r-1}\frac{s}{1+s\gamma}$ when *Y~BB(r,p,γ)*.

c) Show $\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{s}{p+s\gamma}+\sum_{s=0}^{r-Y}\frac{s}{1-p+s\gamma}\right)=\sum_{s=0}^{r}\frac{s}{1+s\gamma}$ when *Y~BB(r+*1*,p,γ)*.

First note that $\mathrm{P}(Y=y|r+1,p,\gamma)=\left\{\dfrac{r-y+\dfrac{1-p}{\gamma}}{r+\dfrac{1}{\gamma}}\right\}\mathrm{P}(Y=y|r,p,\gamma)$

$$=\left\{1+\frac{y(1-\gamma)-p(1+r)}{(1+r\gamma)(r+1-y)}\right\}\mathrm{P}(Y=y|r,p,\gamma)$$

where y=0,…r.

So for $Y\sim BB(r+1,p,\gamma)$

$$\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{s}{p+s\gamma}+\sum_{s=0}^{r-Y}\frac{s}{1-p+s\gamma}\right)$$

$$=\left(\frac{1}{1-p+\gamma}+\frac{2}{1-p+2\gamma}+\cdots+\frac{r}{1-p+r\gamma}\right)\mathrm{P}(Y=0\,|\,r+1,p,\gamma)$$
$$+\left(\frac{1}{1-p+\gamma}+\frac{2}{1-p+2\gamma}+\cdots+\frac{r-1}{1-p+(r-1)\gamma}\right)\mathrm{P}(Y=1\,|\,r+1,p,\gamma)$$
$$+\left(\frac{1}{p+\gamma}+\frac{1}{1-p+\gamma}+\cdots+\frac{r-2}{1-p+(r-2)\gamma}\right)\mathrm{P}(Y=2\,|\,r+1,p,\gamma)$$
$$\vdots$$
$$+\left(\frac{1}{p+\gamma}+\cdots+\frac{r-2}{p+(r-2)\gamma}+\frac{1}{1-p+\gamma}\right)\mathrm{P}(Y=r-1\,|\,r+1,p,\gamma)$$
$$+\left(\frac{1}{p+\gamma}+\cdots+\frac{r-1}{p+(r-1)\gamma}\right)\mathrm{P}(Y=r\,|\,r+1,p,\gamma)$$
$$+\left(\frac{1}{p+\gamma}+\cdots+\frac{r}{p+r\gamma}\right)\mathrm{P}(Y=r+1\,|\,r+1,p,\gamma)$$

$$=\left(\frac{1}{1-p+\gamma}+\frac{2}{1-p+2\gamma}+\cdots+\frac{r}{1-p+r\gamma}\right)\left(1+\frac{-p(1+r)}{(1+r\gamma)(r+1)}\right)\mathrm{P}(Y=0\,|\,r,p,\gamma)$$
$$+\left(\frac{1}{1-p+\gamma}+\frac{2}{1-p+2\gamma}+\cdots+\frac{r-1}{1-p+(r-1)\gamma}\right)\left(1+\frac{(1-\gamma)-\mathrm{p}(1+\mathrm{r})}{(1+r\gamma)r}\right)\mathrm{P}(Y=1\,|\,r,p,\gamma)$$
$$+\left(\frac{1}{p+\gamma}+\frac{1}{1-p+\gamma}+\cdots+\frac{r-2}{1-p+(r-2)\gamma}\right)\left(1+\frac{2(1-\gamma)-p(1+r)}{(1+r\gamma)(r+1-2)}\right)\mathrm{P}(Y=2\,|\,r,p,\gamma)$$
$$\vdots$$
$$+\left(\frac{1}{p+\gamma}+\cdots+\frac{r-2}{p+(r-2)\gamma}+\frac{1}{1-p+\gamma}\right)\left(1+\frac{(r\text{-}1)(1\text{-}\gamma)-p(1+r)}{(1+r\gamma)2}\right)\mathrm{P}(Y=r-1\,|\,r,p,\gamma)$$

$$+\left(\frac{1}{p+\gamma}+\cdots+\frac{r-1}{p+(r-1)\gamma}\right)\left(1+\frac{r(1-\gamma)-p(1+r)}{1+r\gamma}\right)\mathrm{P}(Y=r\mid r,p,\gamma)$$
$$+\left(\frac{1}{p+\gamma}+\cdots+\frac{r}{p+r\gamma}\right)\mathrm{P}(Y=r+1\mid r+1,p,\gamma).$$

By the induction assumption the above term is equal to

$$\sum_{s=0}^{r-1}\frac{s}{1+s\gamma}+\frac{r}{1+r\gamma}\left[\frac{r}{1-p+r\gamma}\mathrm{P}(Y=0\mid r,p,\gamma)\left(\frac{1+r\gamma}{r}\right)+\cdots+\frac{1}{1-p+\gamma}\mathrm{P}(Y=r-1\mid r,p,\gamma)\left(\frac{1+r\gamma}{r}\right)\right.$$

$$+\left(\frac{1}{1-p+\gamma}+\frac{2}{1-p+2\gamma}+\cdots+\frac{r}{1-p+r\gamma}\right)\left(\frac{-p}{r}\right)\mathrm{P}(Y=0\mid r,p,\gamma)$$
$$+\left(\frac{1}{1-p+\gamma}+\frac{2}{1-p+2\gamma}+\cdots+\frac{r-1}{1-p+(r-1)\gamma}\right)\left(\frac{(1-\gamma)-\mathrm{p}(1+\mathrm{r})}{r^2}\right)\mathrm{P}(Y=1\mid r,p,\gamma)$$
$$+\left(\frac{1}{p+\gamma}+\frac{1}{1-p+\gamma}+\cdots+\frac{r-2}{1-p+(r-2)\gamma}\right)\left(\frac{2(1-\gamma)-p(1+r)}{(r-1)r}\right)\mathrm{P}(Y=2\mid r,p,\gamma)$$
$$\vdots$$
$$+\left(\frac{1}{p+\gamma}+\cdots+\frac{r-2}{p+(r-2)\gamma}+\frac{1}{1-p+\gamma}\right)\left(\frac{(r-1)(1-\gamma)-p(1+r)}{r2}\right)\mathrm{P}(Y=r-1\mid r,p,\gamma)$$

$$+\left(\frac{1}{p+\gamma}+\cdots+\frac{r-1}{p+(r-1)\gamma}\right)\left(\frac{r(1-\gamma)-p(1+r)}{r}\right)\mathrm{P}(Y=r\mid r,p,\gamma)$$
$$\left.+\left(\frac{1}{p+\gamma}+\cdots+\frac{r}{p+r\gamma}\right)\frac{p+r\gamma}{r}\mathrm{P}(Y=r\mid r,p,\gamma)\right]. \qquad (1)$$

It remains to show that the term in square brackets is equal to one. To show this we will first prove that

$$\mathrm{P}(Y=s\mid r)=\frac{s}{p+s\gamma}\left(\sum_{y=s+1}^{r}\frac{y(1-\gamma)-p(1+r)}{(1+r-y)r}\mathrm{P}(Y=y\mid r,p,\gamma)+\frac{p+r\gamma}{r}\mathrm{P}(Y=r\mid r,p,\gamma)\right)$$

$$+\frac{r-s}{1-p+(r-s)\gamma}\left(\sum_{y=0}^{s}\frac{y(1-\gamma)-p(1+r)}{(1+r-y)r}\mathrm{P}(Y=y\mid r,p,\gamma)+\frac{1+r\gamma}{r}\mathrm{P}(Y=s\mid r,p,\gamma\right).$$

From 2) and 3) in Lemma 1, the above equality is equivalent to:

$$\mathrm{P}(Y=s\mid r)=\frac{s}{p+s\gamma}\left(-\sum_{y=0}^{s}\frac{y(1-\gamma)-p(1+r)}{(1+r-y)r}\mathrm{P}(Y=y\mid r,p,\gamma)\right)$$

$$+\frac{r-s}{1-p+(r-s)\gamma}\left(\sum_{y=0}^{s}\frac{y(1-\gamma)-p(1+r)}{(1+r-y)r}\mathrm{P}(Y=y\mid r,p,\gamma)+\frac{1+r\gamma}{r}\mathrm{P}(Y=s\mid r,p,\gamma\right).$$

Combining terms appropriately, the above equality is equivalent to:

$$\mathrm{P}(Y=s|r)=\frac{1}{p+s\gamma}\left\{-\sum_{y=0}^{s}\mathrm{P}(Y=y|r)\frac{y(1-\gamma)-p(1+r)}{(1+r-y)}\right\}.$$

But this equality was proven by induction in Lemma 1. Applying this equality to (1) results in the following equation:

$$\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{s}{p+s\gamma}+\sum_{s=0}^{r-Y}\frac{s}{1-p+s\gamma}\right)=\sum_{s=0}^{r-1}\frac{s}{1+s\gamma}+\left(\frac{r}{1+r\gamma}\right)\sum_{s=0}^{r}\mathrm{P}(Y=s|r)$$

$$=\sum_{s=0}^{r-1}\frac{s}{1+s\gamma}+\left(\frac{r}{1+r\gamma}\right)$$

$$=\sum_{s=0}^{(r+1)-1}\frac{s}{1+s\gamma}.$$

Therefore, by induction,

$$\mathrm{E}\left(\sum_{s=0}^{Y-1}\frac{s}{p+s\gamma}+\sum_{s=0}^{r-Y-1}\frac{s}{1-p+s\gamma}\right)=\sum_{s=0}^{r-1}\frac{s}{1+s\gamma}.$$

APPENDIX C

FORTRAN PROGRAM

```
use msimsl
        use portlib
        integer up, m, np, dim, it
        integer iparam(7), stayatmf
        real lambda, phi
        parameter (it=10000)
        parameter (np=3)
        parameter (dim=210)
        data up/10/
        data m/2/
        data phi/.3/
        data lambda/12/
        data iseed/478387/
        double precision x (dim)
        double precision u(dim)
        double precision uu(dim)
        double precision rchy(21)
        double precision py(21)
        double precision cy(21)
        double precision betaa(21), betaad, alpha, delta
        double precision poi(20)
        double precision cpoi(20)
        double precision theta1
        double precision theta2
        double precision gamhat
        double precision t(np),tguess(np),tlb(np),tub(np)
        double precision tscale(np),fscale,rparam(7),fvalue
        double precision g(np), h(1:3,1:3), lkhd
        double precision x0min, x0max, xrmin, xrmax
        double precision q1
        double precision l10hat, l25hat, l50hat, l75hat, l90hat
        double precision mse10, mse25, mse50, mse75, mse90
        double precision rn10, rn25, rn50, rn75, rn90

        integer y(dim), r(dim)
        integer n, strt, chk
        integer count, hitbnd, stybnd, xbnd
        integer iseed
        integer toty1, toty2, totr1, totr2

        logical flag, flag2, flag3
```

```
        common y,r,x,m,n

        external likelhd, grad, hess

        open (unit=2, access='append', file='d:\thesis\simulations\loglog1.txt')
        open (unit=4, access='append', file='d:\thesis\simulations\xbndlglg.txt')
        call rnset (iseed)

        stayatmf=0
        hitbnd=0
        stybnd=0
        xbnd=0
        do 400 l=1,it
        flag2=.true.
        flag3=.false.
        n =1
        strt=0
        count=0
        x=0d0
        y=0d0
        r=0d0
        u=0d0
        uu=0d0
        theta1=0D0
        theta2=0D0
        gamhat=0D0
        tlb(1)=-5d0
        tub(1)=5d0
        tlb(2)=.2d0
        tub(2)=5d0
        tlb(3)=0d0
        tub(3)=9999d0
        fscale=1d0
        tscale(1)=1d0
        tscale(2)=1d0
        tscale(3)=1d0

SET STARTING VALUES

        do 12 i=1,m
                x(i)=-1.386D0
                x(m+i)=1.386D0
12      continue
```

GENERATE UNIFORM RANDOM VARIABLES FOR DETERMINING LITTER SIZES AND RESPONSES (Y'S)

```
        do 13 i=1,2*m*up
                u(i)=drnunf()
                uu(i)=drnunf()
```

GENERATE LITTER SIZES

```
13      continue
        do 15 j=1, 2*m*up
         do 14 i= 1, 20
          poi(i)=(poidf(i,lambda)-poidf(i-1,lambda))/.98840
                cpoi(1)=poi(1)
                if (i .ne. 1) then
                        cpoi(i)=cpoi(i-1)+poi(i)
                endif
                if (uu(j) .lt. cpoi(i) .and. r(j) .eq. 0) then
                        r(j)=i
                endif
                if (i .eq. 20 .and. r(j) .eq. 0) then
                        r(j)=20
                endif
14       continue
15      continue
```

GENERATE Y'S

```
140     do 100 i=1, 2*m
         count=count+1
         delta=(1d0-phi)/phi*(exp(-exp(-.51+.71*x(count))))
         alpha=(1d0-phi)/phi*(1d0-exp(-exp(-.51+.71*x(count))))
         betaad=dgamma(alpha)*dgamma(delta)/dgamma(alpha+delta)
         cy=0d0
         do 110 j=0, r(count)
           rchy(j+1)=fac(r(count))/(fac(j)*fac(r(count)-j))
           betaa(j+1)=dgamma(alpha+j)*dgamma(r(count)+delta-j)
                      /dgamma(alpha+r(count)+delta)
           py(j+1)=rchy(j+1)*betaa(j+1)/betaad
           cy(1)=py(1)
                if (j .ne. 0) then
                  cy(j+1)=cy(j)+py(j+1)
                endif
```

```
             if (u(count) .lt. cy(j+1)) then
              y(count)=j
              goto 100
             endif
110       continue
100      continue
```

CHECK EXISTENCE OF MLE'S

```
         chk=0
         x0min=1000
         xrmin=1000
         x0max=-1000
         xrmax=-1000
         do 120 i=1, 2*n
          do 121 j=1,m
           if (y(m*(i-1)+j) .ne. 0) then
                if (x(m*(i-1)+j) .lt. x0min) then
                 x0min=x(m*(i-1)+j)
                endif
                if (x(m*(i-1)+j) .gt. x0max) then
                 x0max=x(m*(i-1)+j)
                endif
               endif
121       continue
          do 122 j=1,m
           if (y(m*(i-1)+j) .ne. r(m*(i-1)+j)) then
                 if (x(m*(i-1)+j) .gt. xrmax) then
                  xrmax=x(m*(i-1)+j)
           endif
                 if (x(m*(i-1)+j) .lt. xrmin) then
                  xrmin=x(m*(i-1)+j)
                 endif
               endif
122       continue
120      continue

         if (x0min .lt. x0max .and. xrmax .gt. xrmin .and. x0min .lt. xrmax) then
          chk=1
         endif
         if (x0min .eq. x0max .and. xrmin .lt. x0min .and. xrmax .gt. x0max) then
          chk=1
         endif
         if (xrmin .eq. xrmax .and. x0min .lt. xrmin .and. x0max .gt. xrmax) then
          chk=1
```

```
      endif
      flag=.false.
      do 123 i=1,2*m*n
        if (y(i) .ne. 0 .and. y(i) .ne. r(i)) then
          flag=.true.
              goto 124
        endif
123   continue
124   if (chk .eq. 1 .and. flag) then
        strt=n
        goto 200
      endif
```

MOSER-FEI PROCEDURE (MLE'S DO NOT EXIST)

```
      toty1=0
      toty2=0
      totr1=0
      totr2=0
      do 130 i=1,m
        toty1=toty1+y(2*m*n-2*m+i)
        totr1=totr1+r(2*m*n-2*m+i)
        toty2=toty2+y(2*m*n-m+i)
        totr2=totr2+r(2*m*n-m+i)
130 continue
      q1=(1/dble(n))*((1/(.022*totr1))*(toty1-totr1*.2))
      q2=((1/dble(n))*((1/(.022*totr2))*(toty2-totr2*.8)))
      x(2*m*n+1)=x(2*m*n-2*m+1)-q1
      x(2*m*n+m+1)=x(2*m*n-m+1)-q2
      if (x(2*m*n+1) .lt. -4.0 .or. x(2*m*n+m+1) .lt. -4.0) then
        x(2*m*n+1)=x(1)
        x(2*m*n+m+1)=x(m+1)
      endif
      if (x(2*m*n+m+1) .gt. 4.0 .or. x(2*m*n+1) .gt. 4.0) then
        x(2*m*n+1)=x(1)
        x(2*m*n+m+1)=x(m+1)
      endif
      if (x(2*m*n+1)-x(2*m*n-2*m+1) .gt. 5.0) then
        x(2*m*n+1)=x(2*m*n-2*m+1)+5.0
      endif
      if (x(2*m*n+1)-x(2*m*n-2*m+1) .lt. -5.0) then
        x(2*m*n+1)=x(2*m*n-2*m+1)-5.0
      endif
      if (x(2*m*n+m+1)-x(2*m*n-m+1) .gt. 5.0) then
        x(2*m*n+m+1)=x(2*m*n-m+1)+5.0
```

```
        endif
        if (x(2*m*n+m+1)-x(2*m*n-m+1) .lt. -5.0) then
         x(2*m*n+m+1)=x(2*m*n-m+1)-5.0
        endif
        if (x(2*m*n+1) .gt. x(2*m*n+m+1) .and. x(2*m*n+1) .gt. x(2*m*(n-1)+1)) then
         x(2*m*n+1)=x(2*m*(n-1)+1)+.25*(x(2*m*(n-1)+m+1)-x(2*m*(n-1)+1))
        endif
        if (x(2*m*n+1) .gt. x(2*m*n+m+1) .and. x(2*m*n+m+1) .lt. x(2*m*(n-1)+m+1))
then
         x(2*m*n+m+1)=x(2*m*(n-1)+m+1)-.25*(x(2*m*(n-1)+m+1)-x(2*m*(n-1)+1))
        endif
        do 131 i=2,m
         x(2*m*n+i)=x(2*m*n+1)
         x(2*m*n+m+i)=x(2*m*n+m+1)
131     continue
        n=n+1
        if (n .eq. up) then
          stayatmf=stayatmf+1
          rn10=(-log(1.0/4.0)+log(9.0))/(-log(1.0/4.0)+log(4.0))
          rn25=(-log(1.0/4.0)+log(3.0))/(-log(1.0/4.0)+log(4.0))
          rn50=(-log(1.0/4.0))/(-log(1.0/4.0)+log(4.0))
          rn75=(-log(1.0/4.0)+log(1.0/3.0))/(-log(1.0/4.0)+log(4.0))
          rn90=(-log(1.0/4.0)+log(1.0/9.0))/(-log(1.0/4.0)+log(4.0))
          l10hat=rn10*x(2*m*(up-1)+1) + (1-rn10)*x(2*m*(up-1)+m+1)
          l25hat=rn25*x(2*m*(up-1)+1) + (1-rn25)*x(2*m*(up-1)+m+1)
          l50hat=rn50*x(2*m*(up-1)+1) + (1-rn50)*x(2*m*(up-1)+m+1)
          l75hat=rn75*x(2*m*(up-1)+1) + (1-rn75)*x(2*m*(up-1)+m+1)
          l90hat=rn90*x(2*m*(up-1)+1) + (1-rn90)*x(2*m*(up-1)+m+1)
          goto 390
        endif
        goto 140

CALCULATE MLE'S

GENERATE Y'S

200     do 220 i=1, (up-strt+1)
         if (n .gt. strt) then
          do 230 k=1, 2*m
           count=count+1
           delta=(1d0-phi)/phi*(exp(-exp(-.51+.71*x(count))))
           alpha=(1d0-phi)/phi*(1d0-exp(-exp(-.51+.71*x(count))))
           if (delta .lt. .0001d0 .or. alpha .lt. .0001d0) then
                      write (4,*) alpha, delta, x(count), l, strt, n, count
             endif
```

```
        betaad=dgamma(alpha)*dgamma(delta)/dgamma(alpha+delta)
        cy=0d0
     do 240 j=0, r(count)
       rchy(j+1)=fac(r(count))/(fac(j)*fac(r(count)-j))
       betaa(j+1)=dgamma(alpha+j)*dgamma(r(count)+delta-j)
                 /dgamma(alpha+r(count)+delta)
       py(j+1)=rchy(j+1)*betaa(j+1)/betaad
       cy(1)=py(1)
           if (j .ne. 0) then
             cy(j+1)=cy(j)+py(j+1)
           endif
           if (u(count) .lt. cy(j+1)) then
             y(count)=j
             goto 230
           endif
240     continue
230    continue
      endif
      tguess(1)=theta1
      tguess(2)=theta2
      tguess(3)=gamhat
      iparam(1)=0
```

CALL IMSL SUBROUTINE TO CALCULATE MLE'S

```
      call dbcoah (likelhd,grad,hess,np,tguess,0,tlb,tub,tscale,fscale,
                   iparam,rparam,t,fvalue)
      if ((t(1) .eq. -5d0 .or. t(1) .eq. 5d0 .or. t(2) .eq. .2d0 .or. t(2) .eq. 5d0 .or. t(3)
         .eq. 0d0 .or. t(3) .eq. 9999d0) .and. flag2) then
            hitbnd=hitbnd+1
            flag2=.false.
      endif
      if ((t(1) .eq. -5d0 .or. t(1) .eq. 5d0 .or. t(2) .eq. .2d0 .or. t(2) .eq. 5d0 .or. t(3)
        .eq. 0d0 .or. t(3) .eq. 9999d0) .and. n .eq. 10) then
         stybnd=stybnd+1
         write (3,*) n,stybnd,l,t
      endif
      theta1=t(1)
      theta2=t(2)
      gamhat=t(3)
```

CALCULATE NEXT X'S

```
      x(2*m*n+1)=(log(.25)-theta1)/theta2
      x(2*m*n+m+1)=(log(4.0)-theta1)/theta2
```

```
        if (x(2*m*n+1) .lt. -5d0) then
         x(2*m*n+1)=-5d0
         flag3=.true.
        endif
        if (x(2*m*n+1) .gt. 5d0) then
         x(2*m*n+1)=5d0
          flag3=.true.
        endif
        if (x(2*m*n+m+1) .lt. -5d0) then
         x(2*m*n+m+1)=-5d0
         flag3=.true.
        endif
        if (x(2*m*n+m+1) .gt. 5d0) then
         x(2*m*n+m+1)=5d0
         flag3=.true.
        endif
        if (n .eq. 10 .and. flag3) then
         xbnd=xbnd+1
        endif
        do 250 ii=2, m
         x(2*m*n+ii)=x(2*m*n+1)
         x(2*m*n+m+ii)=x(2*m*n+m+1)
250     continue
        n=n+1
220 continue
       l10hat=(-theta1+log(1.0/9.0))/theta2
       l25hat=(-theta1+log(1.0/3.0))/theta2
       l50hat=(-theta1)/theta2
       l75hat=(-theta1+log(3.0))/theta2
       l90hat=(-theta1+log(9.0))/theta2
       avgstrt=avgstrt+strt
390    mse10=mse10+(l10hat-(log(log(10d0/9d0))/.71+.51/.71))**2
       mse25=mse25+(l25hat-(log(log(4d0/3d0))/.71+.51/.71))**2
       mse50=mse50+(l50hat-(log(log(2d0))/.71+.51/.71))**2
       mse75=mse75+(l75hat-(log(log(4d0))/.71+.51/.71))**2
       mse90=mse90+(l90hat-(log(log(10d0))/.71+.51/.71))**2
       bias10=bias10+(l10hat-(log(log(10d0/9d0))/.71+.51/.71))
       bias25=bias25+(l25hat-(log(log(4d0/3d0))/.71+.51/.71))
       bias50=bias50+(l50hat-(log(log(2d0))/.71+.51/.71))
       bias75=bias75+(l75hat-(log(log(4d0))/.71+.51/.71))
       bias90=bias90+(l90hat-(log(log(10d0))/.71+.51/.71))
       avgl10=avgl10+l10hat
       avgl25=avgl25+l25hat
       avgl50=avgl50+l50hat
```

```
      avgl75=avgl75+l75hat
      avgl90=avgl90+l90hat
400 continue
      mse10=mse10/(l-1)
      mse25=mse25/(l-1)
      mse50=mse50/(l-1)
      mse75=mse75/(l-1)
      mse90=mse90/(l-1)
      bias10=bias10/(l-1)
      bias25=bias25/(l-1)
      bias50=bias50/(l-1)
      bias75=bias75/(l-1)
      bias90=bias90/(l-1)
      avgl10=avgl10/(l-1)
      avgl25=avgl25/(l-1)
      avgl50=avgl50/(l-1)
      avgl75=avgl75/(l-1)
      avgl90=avgl90/(l-1)
      avgstrt=avgstrt/(l-1)
      write(2,*)iseed,phi,m,up,x(1),avgstrt,stayatmf,hitbnd,stybnd,mse10,mse25,mse50,
           mse75,mse90,bias10,bias25,bias50,bias75,bias90,avgl10,avgl25,avgl50,
           avgl75,avgl90
      write(4,*) iseed,phi,m,up,x(1),xbnd
300   end
```

SUBROUTINE TO CALCULATE GRADIENT

```
      subroutine grad (np,t,g)

      integer k,np,m,n
      integer y(210), r(210)

      real s

      double precision theta1, theta2, gamhat
      double precision sum1, sum2, pred
      double precision suma, sumb, sumc
      double precision g(np), t(np), x(210)

      common y,r,x,m,n

      theta1=t(1)
      theta2=t(2)
      gamhat=t(3)
      g(1)=0d0
```

```
      g(2)=0d0
      g(3)=0d0
      sum1=0d0
      sum2=0d0
      sum3=0d0
      suma=0d0
      sumb=0d0
      sumc=0d0
      do 20 k=1,2*m*n
            pred=1d0/(1d0+dexp(-(theta1+theta2*x(k))))
            sum1=0d0
            suma=0d0
            if (y(k)-1 .ge. 0) then
              do 21 s=0, y(k)-1
                  sum1=sum1+1/(pred+s*gamhat)
                  suma=suma+s/(pred+s*gamhat)
21            continue
            endif
            sum2=0d0
            sumb=0d0
            if (r(k)-y(k)-1 .ge. 0) then
              do 22 s=0, r(k)-y(k)-1
                  sum2=sum2+1/(1-pred+s*gamhat)
                  sumb=sumb+s/(1-pred+s*gamhat)
22            continue
            endif
            sumc=0d0
            do 23 s=0, r(k)-1
              sumc=sumc+s/(1+s*gamhat)
23          continue
            g(1)=g(1)+pred*(1-pred)*(sum1-sum2)
            g(2)=g(2)+x(k)*pred*(1-pred)*(sum1-sum2)
            g(3)=g(3)+suma+sumb-sumc
20    continue
      g(1)=-1d0*g(1)
      g(2)=-1d0*g(2)
      g(3)=-1d0*g(3)
      return
      end
```

SUBROUTINE TO CALCULATE LIKELIHOOD FUNCTION

```
 subroutine likelhd (np,t,lkhd)

      integer k,np,m,n
```

```
      integer y(210), r(210)

      double precision s

      double precision theta1, theta2, gamhat
      double precision sum1, sum2, sum3, pred
      double precision lkhd, t(np), x(210)

      common y,r,x,m,n

      theta1=t(1)
      theta2=t(2)
      gamhat=t(3)
      lkhd=0d0
      sum1=0d0
      sum2=0d0
      sum3=0d0
      do 16 k=1,2*m*n
            pred=1d0/(1d0+dexp(-(theta1+theta2*x(k))))
            if (y(k)-1 .ge. 0) then
              do 17 s=0, y(k)-1
                    sum1=sum1+dlog(pred+s*gamhat)
17            continue
            endif
            if (r(k)-y(k)-1 .ge. 0) then
              do 18 s=0, r(k)-y(k)-1
                    sum2=sum2+dlog(1-pred+s*gamhat)
18            continue
            endif
            do 19 s=0, r(k)-1
              sum3=sum3+dlog(1+s*gamhat)
19          continue
            lkhd=lkhd+sum1+sum2-sum3
            sum1=0d0
            sum2=0d0
            sum3=0d0
16    continue
      lkhd=-1d0*lkhd
      return
      end
```

SUBROUTINE TO CALCULTE HESSIAN

```
      subroutine hess (np,t,h,ldh)
```

```
integer k,np,m,n
integer y(210), r(210)

real s

double precision theta1, theta2, gamhat
double precision sum1, sum2, pred
double precision pc1, pc2, pc3
double precision sum3a, sum3b
double precision sumc1, sumc2, sumc3
double precision t(np), x(210), h(1:3,1:3)
common y,r,x,m,n

theta1=t(1)
theta2=t(2)
gamhat=t(3)
h(1,1)=0d0
h(1,2)=0d0
h(1,3)=0d0
h(2,1)=0d0
h(2,2)=0d0
h(2,3)=0d0
h(3,1)=0d0
h(3,2)=0d0
h(3,3)=0d0
sum1=0d0
sum2=0d0
sum3a=0d0
sum3b=0d0
sumc1=0d0
sumc2=0d0
sumc3=0d0
do 24 k=1,2*m*n
        pred=1d0/(1d0+dexp(-(theta1+theta2*x(k))))
        sum1=0d0
        pc1=0d0
        pc2=0d0
        pc3=0d0
        sum3a=0d0
        sumc1=0d0
        if (y(k)-1 .ge. 0) then
          do 25 s=0, y(k)-1
                pc1=((pred**2)*(1-pred)*(pred+s*gamhat)/(pred+s*gamhat)**2)
                pc2=(pred*(1-pred)**2*(pred+s*gamhat)/(pred+s*gamhat)**2)
                pc3=((pred**2)*(1-pred)**2/(pred+s*gamhat)**2)
```

```
                sum1=sum1-pc1+pc2-pc3
                sum3a=sum3a-s/(pred+s*gamhat)**2
                sumc1=sumc1-s**2/(pred+s*gamhat)**2
25           continue
          endif
          sum2=0d0
          sum3b=0d0
          sumc2=0d0
          pc1=0d0
          pc2=0d0
          pc3=0d0
          if (r(k)-y(k)-1 .ge. 0) then
           do 26 s=0, r(k)-y(k)-1
                pc1=((pred**2)*(1-pred)*(1-pred+s*gamhat)/
     &                                    (1-pred+s*gamhat)**2)
                pc2=(pred*(1-pred)**2*(1-pred+s*gamhat)/
     &                                    (1-pred+s*gamhat)**2)
                pc3=((pred**2)*(1-pred)**2/(1-pred+s*gamhat)**2)
                sum2=sum2-pc1+pc2+pc3
                sum3b=sum3b+s/(1-pred+s*gamhat)**2
                sumc2=sumc2-s**2/((1-pred)+s*gamhat)**2
26           continue
          endif
          sumc3=0d0
          do 27 s=0, r(k)-1
           sumc3=sumc3+s**2/(1+s*gamhat)**2
27        continue
          h(1,1)=h(1,1)+(sum1-sum2)
          h(1,2)=h(1,2)+x(k)*(sum1-sum2)
          h(2,2)=h(2,2)+x(k)**2*(sum1-sum2)
          h(1,3)=h(1,3)+pred*(1-pred)*(sum3a+sum3b)
          h(2,3)=h(2,3)+x(k)*pred*(1-pred)*(sum3a+sum3b)
          h(3,3)=h(3,3)+(sumc1+sumc2+sumc3)
24     continue
       h(1,1)=-1d0*h(1,1)
       h(1,2)=-1d0*h(1,2)
       h(1,3)=-1d0*h(1,3)
       h(2,2)=-1d0*h(2,2)
       h(2,3)=-1d0*h(2,3)
       h(3,3)=-1d0*h(3,3)
       h(2,1)=h(1,2)
       h(3,1)=h(1,3)
       h(3,2)=h(2,3)
       return
       end
```

VITA

Laura Price Coombs

Candidate for the Degree of

Doctor of Philosophy

Thesis: SEQUENTIAL ROOT ESTIMATION FOR THE BETA-BINOMIAL DISTRIBUTION

Major Field: Statistics

Biographical:

Personal Data: Born in Tallahassee, Florida, on December 22, 1965, the daughter Tom and Nancy Price. Married William Thomas Coombs of Knoxville, Tennessee on September 29, 1990.

Education: Graduated from Archbishop Curley Notre-Dame High School, Miami, Florida in June 1984; received Bachelor of Science degree in Mathematics from Florida Southern College, Lakeland, Florida in May, 1987; received Master of Statisics degree from University of Florida, Gainesville, Florida in May 1989. Completed the requirements for the Doctor of Philosophy degree with a major in Statistics at Oklahoma State University in May 1999.

Experience: Teaching Assistant, Department of Statistics, University of Florida, 1987 to 1989; Teaching Assistant, Department of Foundations of Education, University of Florida, 1989 to 1990; Programmer/Analyst, Infotech, Inc., Gainesville, Florida, 1990 to 1992; Teaching Assistant, Department of Curriculum and Instruction, Oklahoma State University, 1993 to 1995; Research Assistant, Department of Statistics, Oklahoma State University and Arkansas Foundation for Medical Care, Fort Smith, Arkansas, 1997 to 1998; Teaching Assistant, Department of Statistics, Oklahoma State University, 1996 to present.

Professional Memberships: American Statistical Association; Sigma Xi, Scientific Research Society