TWO STAGE ESTIMATION OF THE QUANTILES OF THE LOGIT MODEL

By

TESFAYE LIRANSO

Bachelor of Science
Addis Ababa University
Addis Ababa, Ethiopia
1982

Master of science
Addis Ababa University
Addis Ababa, Ethiopia
1987

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
December, 1993

TWO STAGE ESTIMATION OF THE QUANTILES OF

THE LOGIT MODEL

Thesis Approved:

Barry Kurt Moser
Thesis Adviser

R Leroy Folks

David L. Weeks

Mark E Payton

Kevin M. Currier

Thomas C. Collins
Dean of the Graduate College

ACKNOWLEDGEMENTS

A special thank you to my dissertation adviser, Dr. Barry K. Moser, for suggesting this topic and for his valuable guidance and timely help. I would also like to thank Dr. Leroy Folks, Head of the Statistics Department, for his help in using Chi Writer. The support, advice and information gained from committee members Dr. David Weeks, Dr. Mark Payton and Dr. Kevin Currier is also appreciated. A special thanks goes to all the faculty and graduate students in the Department of Statistics. Their friendship and help made my Ph.D. work much easier.

I would like to thank the Rotary Foundation for providing financial support for the last three years of my study. Thank you to Dr. Don Peters, for his friendship and help in communicating with the Rotary Foundation. I am also grateful to the Statistics Department for providing financial assistance for this semester which made the completion of my dissertation work possible.

I would like to express my appreciation to my colleagues at Awassa College of Agriculture and my former professors in Addis Ababa for their support in my effort to secure a study leave from my employer.

My wife, Gesegessu Yemane, deserves a special recognition for her moral and material support in facilitating my work. Her willingness to work double shifts was one of the primary sources of meeting my financial needs. Thank you to my children who rarely complained about the lack of time available to them from me, and their frequent support for my efforts towards a Ph.D.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE PAGE

CHAPTER I

INTRODUCTION

In various research areas, the outcomes of experiments are assumed to be dichotomous, response or non-response. For instance, consider the following experiments.

(1) Some five-inch lengths of plastic conduit pipe are subjected to impacts of various energies. The response to an impact is either brittle " fail " or " no fail ".

(2) Weights are dropped onto explosive materials from various heights. The response is " explode " or " not explode".

Both of these trials can be characterized by the following statistical model. A stimulus is applied at various levels x, and the response Y(x) is a random variable taking on two values, one and zero, the probability of a response being P(x). For increasing x, P(x) increases gradually from zero to one, following some *S*-shaped curve. A typical *S*-shaped type curve is that of the logit model,

$$(1.1) \quad P(Y(x) = 1)=\exp(\theta_1+\theta_2 x)/(1+\exp(\theta_1+\theta_2 x))$$

where θ_1 and θ_2 are unknown parameters and x is the level at which the stimulus is applied.

For such response curves, an experimenter must first decide the region or parameter of interest. For experiment (1) above, one of the most important features is the location of the lower end of the response curve. This can be stated in terms of the energy level, L_p, for which the probability of the positive response is p,

$$P(Y(L_p)=1) = p \tag{1.2}$$

where p may be .01 or .10. L_p is often referred to as a p^{th} *quantile* of the response curve.

Fixed and sequential methods of estimation of L_p already exist in the literature. Both methods have advantages and disadvantages. The advantages of fixed sample estimation is that it is faster to apply and less expensive than the sequential method. But the variance of the estimated roots can be quite high for certain arrangements of the x levels. On the other hand, the benefit of the sequential method is that it provides estimates with smaller variance. However, it can have the drawback of being more expensive and time consuming. The objective of this research paper is to develop a new procedure that takes advantage of the good aspects of the fixed and sequential procedures, while down-playing their poorer qualities.

The material in this thesis is organized in the following manner. Chapter II presents a literature review of fixed and sequential estimation methods for single and multiple roots. In Chapter III, a new method is developed which is aimed at providing point estimates of the quantiles of the *logit model*. The procedure involves two stages in which a sequential estimation method in the first stage is combined with a fixed sample size procedure in the second stage to produce estimates of the quantiles of the logit model with certain desirable properties. This method is called TWO STAGE ESTIMATION (TSE) procedure. The results of simulation studies are presented in Chapter IV. Appendix A discusses the properties of the TSE procedure when it is applied to two parameter dichotomous models other than the logit model.

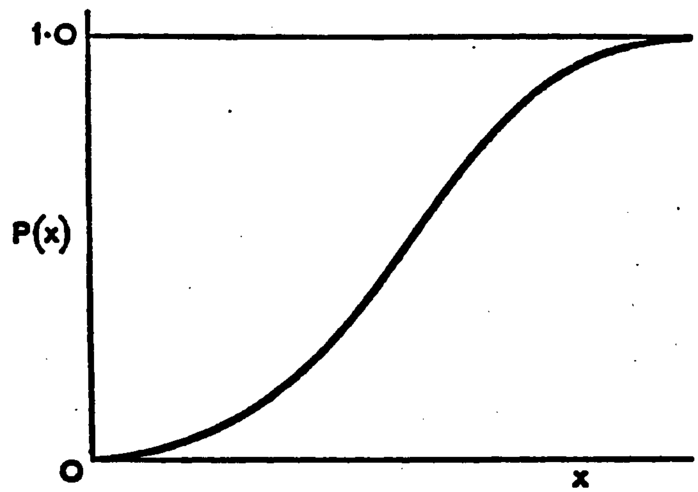


FIGURE 1 GRAPH OF THE LOGIT RESPONSE CURVE

CHAPTER II

LITERATURE REVIEW

Sequential Estimation Methods for Single Roots

An experimenter observes random variables, $Y(x)$, which have a distribution depending on the level of x. Let the expectation of $Y(x)$ be $M(x)$,

$$(2.1.1) \qquad E\{Y(x)\} = M(x) = \int_{-\infty}^{+\infty} y(x)\,dH(y;\underset{\sim}{\theta}|x),$$

where $H(y;\underset{\sim}{\theta}|x) = Pr\{Y(x) \leq y\}$ and $\underset{\sim}{\theta}$ is a vector of unknown parameters.

The experimenter is interested in estimating a single value of x, say L_p, at which the expected response is

$$(2.1.2) \qquad M(L_p) = p.$$

Robbins and Monro (1951) proposed the following sequential procedure to estimate a single value L_p. Starting with an initial guess x_1, successive observations $Y_j(x_j)$ are taken at levels x_j chosen by the formula

$$(2.1.3) \qquad x_{n+1} = x_n - a_n\{y_n(x_n) - p\}$$

where the sequence $\{a_j\}$, $j = 1, 2, \ldots$ is decreasing and $\lim_{j\to\infty} a_j = 0$. After n observations, x_{n+1} is taken as an estimate of L_p. Robbins and Monro assumed that:

(1) ∃ a positive number c such that:

$$Pr\{|Y(x)| \le c\} = \int_{-c}^{+c} dH(y|x) \quad \forall x.$$

(2) $M(x)$ is an increasing function of x, $M(L_p) = p$, and $M'(L_p) > 0$.

(3) $\{a_n\}$ is a sequence of positive constants such that

$$0 < \sum_{1}^{\infty} a_n^2 = A < \infty \text{ and } \sum_{2}^{\infty} \frac{a_n}{(a_1 + a_2 + \ldots a_{n-1})} = \infty.$$

Under conditions (1) - (3), Robbins and Monro showed that x_{n+1} converges to L_p in L^2, that is,

$$\text{(2.1.4)} \qquad \lim_{n\to\infty} E[x_{n+1} - L_p]^2 = 0.$$

Blum (1954), Dvoretzky (1956), and Robbins and Sigmund (1971) showed that $x_n \xrightarrow{a.s} L_p$ under certain conditions with a proper choice of $\{a_n\}$. Chung (1954) and later Sacks (1958) found conditions under which the Robbins-Monro estimator is asymptotically normal. Sacks (1958) defined $a_n = n^{-1}A$, $n = 1, 2, 3, \ldots$ where A is a positive constant. Therefore Sacks' form of (2.1.3) is

$$x_{n+1} = x_n - n^{-1}A\{y_n(x_n) - p\}. \tag{2.1.5}$$

He showed that under certain general conditions the sequence

$$n^{\frac{1}{2}}\left[x_n - L_p\right] \xrightarrow{L} N(0, A^2\sigma^2/(2A\beta-1)) \text{ where} \tag{2.1.6}$$

$$\sigma^2 = \lim_{x \to L_p} E[Y(x) - p]^2 \tag{2.1.7}$$

$$= \lim_{n\to\infty} \mathrm{var}(\, Y(x_n))$$

and

$$\beta = M'(x)\big|_{x=L_p}. \tag{2.1.8}$$

The asymptotic variance is minimized if $A = \beta^{-1}$. But M and L_p are unknown. Thus, to estimate the minimal variance, it is necessary to estimate β.

The problem of estimating β was first considered by Albert and Gardner (1967) and Sakrison (1965). Both Albert and Gardner and Sakrison replaced the constant A by a stochastic sequence estimating β^{-1}. Because in both cases the estimating sequence depends on M, their methods are useful only when M is a known function.

Ventor (1967) studied the case of unknown M. His procedure requires that at the n^{th} stage (n = 1, 2, ...) the experimenter takes two observations, Y_1 and Y_2, at $x_n - c_n$ and $x_n + c_n$, where x_n is the n^{th} approximation and $\{c_n\}$, n = 1, 2, ... is a sequence of positive numbers converging to zero.

Anbar (1977) suggested estimating β by the least square estimator

(2.1.9)
$$b_{m,n} = \frac{\sum_{i=m}^{n} (x_i - \bar{x}_{m,n}) y_i}{\sum_{i=m}^{n} (x_i - \bar{x}_{m,n})^2}$$

where $\bar{x}_{m,n} = (n-m)^{-1} \sum_{i=m+1}^{n} x_i$ and $m = m(n) = o((\log n)^{\frac{1}{2}+\varepsilon}$ $\forall \varepsilon > 0$.

He recommended the stochastic approximation method given by

(2.1.10)
$$x_{n+1} = x_n - (y_n - p)/nb_{m(n),n-1}.$$

Anbar developed a truncated version of (2.1.10) given by

(2.1.11)
$$x_{n+1} = x_n - A_{m,n} n^{-1}(y_n - p),$$

$n = 1, 2, 3, \ldots$ where

(2.1.12)
$$A_{m,n}^{-1} = \begin{cases} \alpha_1 & \text{if } b_{m,n-1} < \alpha_1 \\ b_{m,n} & \text{if } \alpha_1 \le b_{m,n-1} \le \alpha_2 , \\ \alpha_2 & \text{if } \alpha_2 \le b_{m,n-1} \end{cases}$$

$b_{m,n}$ is as in (2.1.9), $0 < \alpha_1 < \alpha_2 < \infty$ and Y_n is a random variable with conditional distribution given $(x_1, y_1, \ldots y_{n-1})$. Under certain regularity conditions, Anbar showed that:

(2.1.13) (a) $b_{m,n} \xrightarrow{a.s} \beta$

(b) $x_n \xrightarrow{a.s} L_p$

$$\text{(c)}\ \sqrt{n}(x_n - L_p) \xrightarrow{d} N\ (\ 0,\ \sigma^2 / \beta^2)$$

where σ^2 and β are as in (2.1.7) and (2.1.8), respectively.

Wu (1985) suggested estimating the root L_p from an estimate of the entire function M. He used a parametric model

$$F(x) = H(x|\lambda_1, \lambda_2, \ldots \lambda_k)$$

where H is continuous in x,

$$\lim_{x \to -\infty} H(x|\lambda_1, \lambda_2, \ldots \lambda_k) = 0, \quad \lim_{x \to \infty} H(x|\lambda_1, \lambda_2, \ldots \lambda_k) = 1.$$

The general method of this sequential design procedure for estimating L_p is as follows:

(1) Find an efficient estimator $\hat{\lambda}_n = \hat{\lambda}\ [(x_i, y_i)_1^n\ ]$ of $\lambda = (\lambda_1, \lambda_2, \ldots \lambda_k)'$.

(2) Define the estimated quantal response curve $\hat{F}_n(x) = H(x|\hat{\lambda}_n)$ and choose the next design level x_{n+1} such that $\hat{F}_n(x_{n+1}) = p$. After n updates, x_{n+1} is taken as the estimator of L_p.

Sequential Estimation Method for Multiple Roots

Beginning with the paper by Robbins and Monro, much work has been done in stochastic approximation methods with the purpose of estimating a single root of $M(x, \underset{\sim}{\theta})$. The problem of estimating multiple roots was addressed by Moser and Fei (1991) through estimating the entire curve $M(x, \underset{\sim}{\theta})$.

Their procedure was based on a k-dimensional Robbin-Monro model summarized as follows.

Let p_j, $j = 1, 2, \ldots k$ be such that $M(x,\underset{\sim}{\theta}) = p_j$, for given p_j. If $\underset{\sim}{\theta}' = (\theta_1,\theta_2,\ldots,\theta_k)$ then the k roots L_{p_j} characterize $M(x,\underset{\sim}{\theta})$. The estimates $\hat{L}_{p_j}$ then induce estimates of the parameters $\hat{\underset{\sim}{\theta}}' = (\hat{\theta}_1, \hat{\theta}_2, \ldots\hat{\theta}_k)$ which will in turn produce an estimate of the curve $M(x,\underset{\sim}{\theta})$ through $M(x,\hat{\underset{\sim}{\theta}})$. Then for any p of interest, its root L_p is estimated by $\hat{L}_p$ as a solution of the equation $M(x,\hat{\underset{\sim}{\theta}}) = p$. The estimates of the j^{th} tangent slope β_j, $j = 1, 2,\ldots k$ of $\underset{\sim}{\beta}' = (\beta_1, \beta_2,\ldots\beta_k)$, given in (2.1.8) are computed as $\hat{\beta}_j = M'(x,\hat{\theta})\big|_{x=\hat{L}_{p_j}}$.

Fixed Sample Estimation Method for Multiple Roots

Multiple roots can also be estimated by using a fixed sample approach. Again let $M(x,\underset{\sim}{\theta})$ represent the curve where $\underset{\sim}{\theta}' = (\theta_1, \theta_2, \ldots,\theta_k)$ is a k-dimensional vector of unknown parameters. Fix r different values of x_j , $j = 1, 2, \ldots r$ where $r \geq k$. At each value of x_j, observe n_j values of Y_j, $j = 1, 2, \ldots ,r$. Using all $n = \sum_{j=1}^{r} n_j$ observations, generate the maximum likelihood estimates $\hat{\underset{\sim}{\theta}}' = (\hat{\theta}_1, \hat{\theta}_2, \ldots,\hat{\theta}_k)$. Then, for any value of p, the root estimate $\hat{L}_p$ is the value of x that satisfies the equation $M(x,\hat{\underset{\sim}{\theta}}) = p$.

Sequential and Fixed Estimation Using the Logit Model for Dichotomous Data

For the logit model the problem is to estimate L_p such that $M(L_p;\theta_1,\theta_2) = p$ where

$$M(x;\theta_1,\theta_2) = \exp(\theta_1+\theta_2 x)/(1+\exp(\theta_1+\theta_2 x)). \tag{2.4.1}$$

Note that (2.4.1) is the expectation of Y(x) following a binary logit model since

$$M(x;\theta_1,\theta_2) = 0(1-p) + 1p = P(Y(x)=1|x), \tag{2.4.2}$$

where $P(Y(x)=1|x)$ is the right side of (2.4.1). If an estimate of a single root of the logit model, L_p, is of interest, we can use equation (2.1.5); that is,

$$x_{n+1} = x_n - n^{-1}A(y(x_n) - p), \tag{2.4.3}$$

and after n observations, x_{n+1} is used to estimate L_p.

It should be noted, however, that estimates of a single root L_p through (2.4.3) will not provide estimates of the two parameters θ_1 and θ_2 from the logit equation (2.4.1), and therefore will not provide estimates of the entire curve

$M(x,\underset{\sim}{\theta})$. However, as described below, estimates of θ_1 and θ_2, estimates of the entire curve $M(x,\underset{\sim}{\theta})$ and estimates of any root can be generated by a Moser-Fei sequential or by a fixed sample procedure.

Now it will be demonstrated how the Moser-Fei process can be used to get estimates of θ_1, θ_2, $M(x,\underset{\sim}{\theta})$, and $M'(x,\underset{\sim}{\theta})$ for any L_p and for any p. From (2.4.1), for any $p_1 \neq p_2$,

$$p_1 = \exp(\theta_1+\theta_2 L_{p_1})/(1+\exp(\theta_1+\theta_2 L_{p_1}))$$

and

$$p_2 = \exp(\theta_1+\theta_2 L_{p_2})/(1+\exp(\theta_1+\theta_2 L_{p_2})).$$

Solving these two equations for θ_1 and θ_2 we obtain

$$\text{(2.4.4)} \qquad \theta_1 = \frac{L_{p_1}\log(p_2/(1-p_2)) - L_{p_2}\log(p_1/(1-p_1))}{L_{p_1} - L_{p_2}}$$

$$\text{(2.4.5)} \qquad \theta_2 = \frac{\log(p_1/(1-p_1)) - \log(p_2/(1-p_2))}{L_{p_1} - L_{p_2}}$$

Estimates of θ_1, θ_2 are obtained by replacing L_{p_1}, L_{p_2} by $x_{1,n}$, $x_{2,n}$, respectively. The Moser-Fei procedure provides $x_{1,n}$ and $x_{2,n}$ as follows.

$$\text{(2.4.6)} \qquad \begin{bmatrix} x_{1,n+1} \\ X_{2,n+1} \end{bmatrix} = \begin{bmatrix} x_{1,n} \\ X_{2,n} \end{bmatrix} - n^{-1} \begin{bmatrix} (A_{1,n})^{-1}(y_{1,n} - p_1) \\ (A_{2,n})^{-1}(y_{2,n} - P_2) \end{bmatrix}$$

where $A_{1,n}$, $A_{2,n}$ are estimates of the tangent slopes $M'(L_{p_1})$, $M'(L_{p_2})$, respectively. Note that $M'(L_{p_i})$, $i = 1, 2$, is estimated by $M'(x_{i,n})$ which is obtained from (2.4.6). Moreover, for any constant p in the range of $M(x,\underset{\sim}{\theta})$ the root L_p is estimated by

$$\hat{L}_p = rx_{1,n} + (1-r)x_{2,n} \tag{2.4.7}$$

where

$$r = (G^{-1}(p_2) - G^{-1}(p))/(G^{-1}(p_2) - G^{-1}(p_1)) \tag{2.4.8}$$

with $G^{-1}(p) = \log(p/(1-p))$. The asymptotic variance of (2.4.7) is given by

$$\sigma^2_{p_1} r^2/[M'(L_{p_1})]^2 + \sigma^2_{p_2}(1-r)^2/[M'(L_{p_2})]^2 \tag{2.4.9}$$

where $\sigma^2_{p_i}$, $i = 1, 2$ is the variance of the estimate of L_{p_i}.

The fixed sample approach for the logit model is as follows. For given p_j ($j=1,2,...,k$), let x_j be the root of the equation $M(x,\theta) = p_j$. Suppose $\hat{\theta}_1$ and $\hat{\theta}_2$ are the maximum likelihood estimators of θ_1 and θ_2 based on n_j observations of Y_j taken at each fixed x_j, which are assumed to be known. Then for any $p\epsilon(0,1)$, the root estimate of $M(L_p,\underset{\sim}{\theta}) = p$ is

(2.4.9) $$\hat{L}_p = (-\hat{\theta}_1 + \log(p/(1-p)))/\hat{\theta}_2.$$

The asymptotic variance of $\hat{L}_p$ has the same form as (3.2.19) given by

(2.4.10) $$\mathrm{Var}(\hat{L}_p) = (aL_p^2 - 2bL_p + c)/(ac-b^2)$$

where $$a = \sum_{j=1}^{r} n_j p_j (1-p_j),$$

$$b = \sum_{j=1}^{r} n_j x_j p_j (1-p_j)$$

and

$$c = \sum_{j=1}^{r} n_j x_j^2 p_j (1-p_j).$$

We now illustrate how the variances of the root estimators for the fixed sample procedure are affected by the arrangement of the x's. Table 1 provides asymptotic variance of root estimators for the fixed sample method for various arrangements of x's. The first four rows consider the case of estimating L_p(p = .10, .25, .50, .75, .90, .95, .99) when two x's are sampled in such a way that half of the observations are at one x and the remaining half at the other x. The next three rows describe cases where five x-values are sampled in such a way that 1/5 of the observations are taken at each of the x's. In the first column various arrangements of x's are described. The parameters to be estimated L_p (p = .10, .25, .50, .75, .95, .99) are shown at the top part of the table and the corresponding asymptotic variances are given in the body of the table. The last row of the table provides the asymptotic variance when the roots are estimated by Moser-Fei procedure with p_1=.2 and p_2=.8. This selection of p_1, p_2 provides the smallest root variances for the Moser-Fei procedure.

Comparing the fixed and sequential variances, we observe that the variance of the estimated roots are smaller if the x's are chosen by sequential estimation at $L_{.2}$ and $L_{.8}$. Moreover, the fixed root variances when the x's are arranged at $L_{.2}$ and $L_{.8}$ are the same as the sequential root variances. Therefore, running a fixed sample procedure where the x_i's are arranged near $L_{.2}$ and $L_{.8}$ would provide a fast, inexpensive method of producing root estimates with small variance. But apriori, it is not known where $L_{.2}$ and $L_{.8}$ are located. So one approach is to use sequential procedure first to get estimates of $L_{.2}$ and $L_{.8}$, then run a fixed approach at the two estimated roots.

TABLE 1

Variance of L_p for Various Arrangements of x's

Arrangt. Of the x_i's (Fixed)	$L_{.10}$	$L_{.25}$	$L_{.50}$	$L_{.75}$	$L_{.90}$	$L_{.95}$	$L_{.99}$
$L_{.20}, L_{.80}$	$\frac{21.95}{N}$	$\frac{10.18}{N}$	$\frac{6.25}{N}$	$\frac{10.18}{N}$	$\frac{21.95}{N}$	$\frac{34.45}{N}$	$\frac{74.92}{N}$
$L_{.33}, L_{.67}$	$\frac{48.06}{N}$	$\frac{15.41}{N}$	$\frac{4.52}{N}$	$\frac{15.41}{N}$	$\frac{48.06}{N}$	$\frac{82.71}{N}$	$\frac{194.94}{N}$
$L_{.50}, L_{.93}$	$\frac{49.53}{N}$	$\frac{21.78}{N}$	$\frac{8.00}{N}$	$\frac{8.19}{N}$	$\frac{22.35}{N}$	$\frac{39.96}{N}$	$\frac{101.77}{N}$
$L_{.80}, L_{.93}$	$\frac{472.32}{N}$	$\frac{249.47}{N}$	$\frac{99.02}{N}$	$\frac{20.97}{N}$	$\frac{15.34}{N}$	$\frac{52.87}{N}$	$\frac{254.52}{N}$
$L_{.01}, L_{.20}, L_{.50}, L_{.80}, L_{.93}$	$\frac{12.33}{N}$	$\frac{14.47}{N}$	$\frac{7.90}{N}$	$\frac{11.09}{N}$	$\frac{24.04}{N}$	$\frac{38.43}{N}$	$\frac{86.22}{N}$
$L_{.88}, L_{.83}, L_{.50}, L_{.67}, L_{.80}$	$\frac{116.45}{N}$	$\frac{52.23}{N}$	$\frac{15.43}{N}$	$\frac{6.04}{N}$	$\frac{24.06}{N}$	$\frac{51.98}{N}$	$\frac{158.62}{N}$
$L_{.20}, L_{.33}, L_{.50}, L_{.67}, L_{.93}$	$\frac{29.24}{N}$	$\frac{11.08}{N}$	$\frac{5.47}{N}$	$\frac{12.41}{N}$	$\frac{31.90}{N}$	$\frac{52.32}{N}$	$\frac{118.03}{N}$
(Sequential case)	$L_{.10}$	$L_{.25}$	$L_{.50}$	$L_{.75}$	$L_{.90}$	$L_{.95}$	$L_{.99}$
$L_{.20}, L_{.80}$	$\frac{21.95}{N}$	$\frac{10.18}{N}$	$\frac{6.25}{N}$	$\frac{10.18}{N}$	$\frac{21.95}{N}$	$\frac{34.45}{N}$	$\frac{74.92}{N}$

Roots to be Estimated: $L_{.10}$ through $L_{.99}$ (column heads).

CHAPTER III

PROPERTIES OF THE TWO STAGE ESTIMATION PROCEDURE

The New Procedure

Consider sampling from the logit model

$$(3.1.1)\quad \Pr\{Y_{ij}(x_{ij})=1|x_{ij}\} = \exp(\theta_1+\theta_2 x_{ij})/(1+\exp(\theta_1+\theta_2 x_{ij}))$$

in the following way.

Stage 1.(Sequential Stage) Let (x_{ij},y_{ij}) for i = 1, 2 and j = 1, 2, ..., m where y_{ij} is a value of a dichotomous random variable and x_{ij} is the value of the j^{th} sequential estimator of L_{p_i} using the Moser - Fei method. Let the m^{th} estimate of L_{p_i} be $x_{im} = x_i^*$.

Stage 2.(Fixed Stage) For each i = 1, 2, fix x_i^*, and take a sample of size n_i independent observations (x_i^*,y_{ij}), j = m+1, m+2, ..., $m+n_i$ at x_i^*.

Let p be any value $0 < p < 1$. Solving (2.4.1) for x, we obtain $L_p = (-\theta_1+ \log(p/(1-p)))/ \theta_2$, where L_p is the value of x such that $\Pr\{Y(x)=1|x\} = p$. Hence the TSE of L_p is the maximum likelihood estimator of L_p based on $(2m+n_1+n_2)$ samples:

$$(3.1.2) \qquad \hat{L}_p = (-\hat{\theta}_1 + \log(p/(1-p)))/\ \hat{\theta}_2$$

where $\hat{\theta}_1$ and $\hat{\theta}_2$ are the maximum likelihood estimators of θ_1 and θ_2, respectively.

Asymptotic Variance of the TSE of L_p

The likelihood function of $(2m+n_1+n_2)$ observations from a two parameter binary model using the TSE process is

$$(3.2.1) \qquad L(\theta_1,\theta_2) = \prod_{i=1}^{2}\prod_{j=1}^{m} g(y_{ij}|x_{ij}) \prod_{i=1}^{2}\prod_{j=m+1}^{m+n_i} g(y_{ij}|x_i^*)$$

where $g(y_{ij}|x_{ij}) = [\Pr\{y_{ij}=1|x_{ij}\}]^{y_{ij}}[1-\Pr\{y_{ij}=1|x_{ij}]^{1-y_{ij}}$.

To obtain the above result, note that

$$L(\theta_1,\theta_2) = g(x_{11},y_{11},x_{21},y_{21},\ldots,x_{1m},y_{1m},x_{2m},y_{2m},x_1^*,y_{1,m+1}, \ldots y_{1,m+n_1},x_2^*,y_{2,m+1}, \ldots,y_{2,m+n_2})$$

$$= g(y_{1,m+1},y_{1,m+2},\ldots,y_{1,m+n_1},y_{2,m+1},\ldots,y_{2,m+n_2}|x_1^*, x_2^*,x_{11},y_{11},x_{21},y_{21},\ldots,x_{1m},y_{1m},x_{2m},y_{2m})(x_1^*,x_2^*, x_{11},y_{11},x_{21},y_{21},\ldots,x_{1m},y_{1m},x_{2m},y_{2m})$$

$$(3.2.1a) \quad = g(y_{1,m+1}, y_{2,m+2}, \ldots, y_{1,m+n_1}, y_{2,m+1}, y_{2,m+2}, \ldots, y_{2,m+n_2} | x_1^*, x_2^*) g(x_{11}, y_{11}, \ldots, x_{1m}, y_{1m}, x_{2m}, y_{2m}).$$

The above equality holds since $y_{1,m+1}$, $y_{2,m+2}$, . . ., $y_{1,m+n_1}$, $y_{2,m+1}$, $y_{2,m+2}$, ..., $y_{2,m+n_2}$ depend on x_1^*, x_2^*, x_{11}, y_{11}, x_{21}, y_{21}, ..., x_{1m}, y_{1m}, x_{2m}, y_{2m} only through x_1^* and x_2^*. Also, x_1^* and x_2^* are completely determined given x_{11}, y_{11}, x_{21}, y_{21}, ..., x_{1m}, y_{1m}, x_{2m}, y_{2m}. Now, note that given x_1^*, x_2^* the y_{ij}'s, $i = 1, 2$, $j = m+1, m+2, \ldots, m+n_i$ are independent with each y_{1j} dependending only on x_1^* and each y_{2j} dependending only on x_2^*. Therefore, the first term on the right hand side of (3.2.1a) becomes

$$g(y_{1,m+1}, \ldots, y_{1,m+n_1}, y_{2,m+1}, \ldots, y_{2,m+1}, \ldots, y_{2,m+n_2} | x_1^*, x_2^*)$$

$$(3.2.1b) \qquad = \prod_{i=1}^{2} \prod_{j=m+1}^{m+n_i} g(y_{ij} | x_i^*)$$

Furthermore, the second term on the right hand side of (3.2.1a) can be rewritten as

$$g(x_{11}, y_{11}, x_{21}, y_{21}, \ldots, x_{1m}, y_{1m}, x_{2m}, y_{2m})$$
$$= g(y_{1m}, y_{2m} | x_{1m}, x_{2m}, x_{11}, y_{11}, \ldots, x_{1,m-1}, y_{1,m-1}, x_{2,m-1}, y_{2,m-1}) g(x_{1m}, x_{2m}, x_{11}, y_{11}, \ldots, x_{1,m-1}, y_{1,m-1}, x_{2,m-1}, y_{2,m-1})$$

$$=g(y_{1m},y_{2m}|x_{1m},x_{2m})g(x_{11},y_{11},\ldots,x_{1,m-1},$$
$$y_{1,m-1},x_{2,m-1},y_{2,m-1})$$

$$= g(y_{1m}|x_{1m})g(y_{2m}|x_{2m})g(x_{11},y_{11},\ldots,$$
$$x_{1,m-1},y_{1,m-1},x_{2,m-1},y_{2,m-1}).$$

Again the above inequality holds since y_{1m} and y_{2m} depend on x_{1m}, x_{2m}, x_{11}, y_{11}, $x_{1,m-1}$, $y_{1,m-1}$, $x_{2,m-1}$, $y_{2,m-1}$ only through x_{1m} and x_{2m}. Also, x_{1m} and x_{2m} are completely determined given x_{11}, y_{11}, ..., $x_{1,m-1}$, $y_{1,m-1}$, $x_{2,m-1}$, $y_{2,m-1}$. Now note that given x_{1m}, x_{2m} the variables y_{1m}, y_{2m} are independent, with y_{1m} dependent on x_{1m} and y_{2m} dependent only on x_{2m}. Therefore,

(3.2.1c) $$g(y_{1m},y_{2m}|x_{1m},x_{2m})=g(y_{1m}|x_{1m})g(y_{2m}|x_{2m}).$$

Now, simply apply (3.2.1c) iteratively to obtain

(3.2.1d) $$g(x_{11},y_{11},x_{21},y_{21},\ldots,x_{2m},y_{2m})=\prod_{i=1}^{2}\prod_{j=1}^{m}g(y_{ij}|x_{ij}).$$

The likelihood function (3.2.1) is obtained by substituting (3.2.1b) and (3.2.1d) into (3.2.1a).

For the logit model

$$y_{ij}(x_{ij}) = \begin{cases} 1 & \text{with probability } P_{ij} = \Pr\{y_{ij}=1|x_{ij}\}. \\ 0 & \text{with probability } 1-P_{ij}. \end{cases}$$

The variables $y_{ij}|x_{ij}$ are distributed as independent

Bernouli (P_{ij}). Hence (3.2.1) becomes

$$L(\theta_1,\theta_2)=\prod_{i=1}^{2}\prod_{j=1}^{m}[P_{ij}]^{y_{ij}}[1-P_{ij}]^{1-y_{ij}}\prod_{i=1}^{2}\prod_{j=m+1}^{m+n_i}[P_i]^{y_{ij}}[1-P_i]^{1-y_{ij}}$$

where $P_{ij}=Pr\{y_{ij}=1|x_{ij}\}=\exp(\theta_1+\theta_2x_{ij})/(1+\exp(\theta_1+\theta_2x_{ij}))$ and

$P_i = Pr\{y_{ij} = 1| x_i^*\} = \exp(\theta_1+\theta_2x_i^*)/(1+\exp(\theta_1+\theta_2x_i^*)$.

$$(3.2.2)\quad L(\theta_1,\theta_2)=\prod_{i=1}^{2}\prod_{j=1}^{m}[P_{ij}]^{y_{ij}}[1-P_{ij}]^{1-y_{ij}}\prod_{i=1}^{2}\left\{[P_i]^{\sum_{j=m+1}^{m+n_i}y_{ij}}[1-P_i]^{n_i-\sum_{j=m+1}^{m+n_i}y_{ij}}\right\}.$$

The logarithm of (3.2.2) is

$$(3.2.3)\quad \ell = \sum_{i=1}^{2}\sum_{j=1}^{m}\{ y_{ij}\log P_{ij} + (1-y_{ij})\log(1-P_{ij})\} + \sum_{i=1}^{2}\left\{\sum_{j=m+1}^{m+n_i} y_{ij}(\log P_i) + (n_i-\sum_{j=m+1}^{m+n_i}y_{ij})\log(1-P_i)\right\}$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{m}\{y_{ij}(\theta_1+\theta_2x_{ij}) - \log[1+\exp(\theta_1+\theta_2x_{ij})]\} + \sum_{i=1}^{2}\{\sum_{j=m+1}^{m+n_i}y_{ij}\}(\theta_1+\theta_2x_i^*) - \sum_{i=1}^{2}n_i\log(1+\exp(\theta_1+\theta_2x_i^*))$$

To find the value of $\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix}$ that maximizes ℓ, we differentiate ℓ with respect to θ_1 and θ_2 and set the resulting expressions equal to zero.

(3.2.4) $$\frac{\partial \ell}{\partial \theta_1} = \sum_{i=1}^{2}\sum_{j=1}^{m}\{ y_{ij} - \frac{\exp(\theta_1+\theta_2 x_{ij})}{(1+\text{ex}(\theta_1+\theta_2 x_{ij}))}\} + \sum_{i=1}^{2}\sum_{j=m+1}^{m+n_i} y_{ij}$$

$$-\sum n_i\{\exp(\theta_1+ \theta_2 x_i^*)/(1+\exp(\theta_1+\theta_2 x_i^*))\}$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{m+n_i} y_{ij} - \sum_{i=1}^{2}\sum_{j=1}^{m} P_{ij} - \sum_{i=1}^{2} n_i P_i.$$

(3.2.5) $$\frac{\partial \ell}{\partial \theta_2} = \sum_{i=1}^{2}\sum_{j=1}^{m}\{ y_{ij}x_{ij} - \frac{\exp(\theta_1+\theta_2 x_{ij})}{(1+\exp(\theta_1+\theta_2 x_{ij}))}x_{ij}\} +$$

$$\sum_{i=1}^{2}\{\sum_{j=m+1}^{m+n_i} y_{ij}\}x_i^* - \sum_{i=1}^{2} n_i \frac{\exp(\theta_1+\theta_2 x_i^*)}{(1+\exp(\theta_1+\theta_2 x_i^*))}x_i^*$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{m}\{ y_{ij}x_{ij} - P_{ij}x_{ij}\} + \sum_{i=1}^{2}\sum_{j=m+1}^{m+n_i} y_{ij}x_i^* - \sum_{i=1}^{2} n_i P_i x_i^*.$$

The maximum likelihood estimates of θ_1 and θ_2 are the values $\hat{\theta}_1$ and $\hat{\theta}_2$ that satisfy the following conditions.

(3.2.6) $$\sum_{i=1}^{2}\sum_{j=1}^{m+n_i} y_{ij} = \sum_{i=1}^{2}\sum_{j=1}^{m} P_{ij} + \sum_{i=1}^{2} n_i P_i$$

(3.2.7) $$\sum_{i=1}^{2}\sum_{j=1}^{m} y_{ij}x_{ij} + \sum_{i=1}^{2} x_i^* \sum_{j=m+1}^{m+n_i} y_{ij} = \sum_{i=1}^{2}\sum_{j=1}^{m} P_{ij}x_{ij} + \sum_{i=1}^{2} n_i P_i x_i^*.$$

The above equations are *non-linear* and they have to be solved by numerical iteration. If the Newton-Raphson method (Dobson, 1990, P.40) is used, then the kth iterative estimate of $\theta' = (\theta_1, \theta_2)$ is given by

$$\hat{\underset{\sim}{\theta}}^{(k)} = \hat{\underset{\sim}{\theta}}^{(k-1)} - \left[\frac{\partial^2 \ell}{\partial\theta_i \partial\theta_j}\right]^{-1}_{\underset{\sim}{\theta}=\hat{\theta}^{(k-1)}} U^{(k-1)} \tag{3.2.8a}$$

where $\left[\frac{\partial^2 \ell}{\partial\theta_i \partial\theta_j}\right]_{\underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}^{(k-1)}}$ (i, j = 1, 2) is the matrix of second derivatives of ℓ evaluated at $\theta=\hat{\theta}^{(k-1)}$ and $U^{(k-1)} = \begin{bmatrix} \frac{\partial \ell}{\partial\theta_1} \\ \frac{\partial \ell}{\partial\theta_2} \end{bmatrix}$ evaluated at $\underset{\sim}{\theta} = \hat{\underset{\sim}{\theta}}^{(k-1)} = \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix}^{(k-1)}$. An alternative procedure which is sometimes simpler than the Newton-Raphson method is called the *Method of Scoring* (Dobson, 1990, p.40). It involves replacing the matrix of second derivatives in (3.2.8a) by the matrix of expected values $E[\partial\ell^2/\partial\theta_1\partial\theta_2]$. Thus equation (3.2.8a) is replaced by

$$\hat{\underset{\sim}{\theta}}^{(k)} = \hat{\underset{\sim}{\theta}}^{(k-1)} + [I^{(k-1)}]^{-1} U^{(k-1)} \tag{3.2.8b}$$

where $I^{(k-1)}$ denotes the information matrix evaluated at $\hat{\underset{\sim}{\theta}}^{(k-1)}$. The second derivatives of ℓ are given by

$$\frac{\partial^2 \ell}{\partial\theta_1^2} = -\{\sum_{i=1}^{2}\sum_{j=1}^{m} P_{ij}(1-P_{ij}) + \sum_{i=1}^{2} n_i P_i (1-P_i)\} \tag{3.2.9}$$

(3.2.10) $$\frac{\partial^2 \ell}{\partial\theta_2 \partial\theta_1} = -\{\sum_{i=1}^{2}\sum_{j=1}^{m} x_{ij}P_{ij}(1-P_{ij}) + \sum_{i=1}^{2} n_i P_i (1-P_i) x_i^*\}$$

(3.2.11) $$\frac{\partial^2 \ell}{\partial\theta^2} = -\{\sum_{i=1}^{2}\sum_{j=1}^{m} P_{ij}(1-P_{ij})x_{ij}^2 + \sum_{i=1} n_i P_i (1-P_i)(x_i^*)^2$$

The *information matrix* denoted by I is given by

(3.2.12) $$I = -\begin{bmatrix} E\,\dfrac{\partial^2 \ell}{\partial\theta_1^2} & E\,\dfrac{\partial^2 \ell}{\partial\theta_1 \partial\theta_2} \\ E\,\dfrac{\partial^2 \ell}{\partial\theta_1 \partial\theta_2} & E\,\dfrac{\partial^2 \ell}{\partial\theta_2^2} \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

where

(3.2.13) $$a = -E\frac{\partial^2 \ell}{\partial\theta^2} = \sum_{i=i}^{2}\sum_{j=1}^{m} P_{ij}(1-P_{ij}) + \sum_{i=1}^{2} n_i P_i (1-P_i)$$

(3.2.14) $$b = -E\frac{\partial^2 \ell}{\partial\theta_1 \partial\theta_2} = \sum_{i=1}^{2}\sum_{j=1}^{m} x_{ij}P_{ij}(1-P_{ij}) + \sum_{i=1}^{2} n_i P_i (1-P_i) x_i^*$$

(3.2.15) $$c = -E\frac{\partial^2 \ell}{\partial\theta_2^2} = \sum_{i=1}^{2}\sum_{j=1}^{m} x_{ij}^2 P_{ij}(1-P_{ij}) + \sum_{i=1}^{2} n_i (x_i^*)^2 P_i (1-P_i)$$

Hence for large n_i the asymptotic variance-covariance matrix of two stage estimators of θ_1 and θ_2 is given by

$$(3.2.16) \qquad \mathrm{Var}\begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} = \frac{1}{ac - b^2} \begin{bmatrix} c & -b \\ -b & a \end{bmatrix}$$

where a, b, and c are as in (3.2.13), (3.2.14) and (3.2.15), respectively.

Using equation (3.1.2) for $\hat{L}_p$ and applying an approximation suggested by Cochran (1977, p.31)

$$(3.2.17) \qquad \hat{L}_p - L_p = \frac{-\hat{\theta}_1 + \log(p/(1-p))}{\hat{\theta}_2} - L_p$$

$$= \frac{-\hat{\theta}_1 + \log(p/(1-p)) - L_p\hat{\theta}_2}{\hat{\theta}_2}$$

$$\cong \frac{-\hat{\theta}_1 + \log(p/(1-p)) - L_p\hat{\theta}_2}{\theta_2}$$

for a large sample of size $(2m+n_1+n_2)$. Furthermore,

$$(3.2.18) \qquad (\hat{L}_p - L_p)^2 \cong \frac{1}{(\theta_2)^2}\left[-\hat{\theta}_1 + \log(p/(1-p)) - L_p\hat{\theta}_2\right]^2$$

$$= \frac{1}{(\theta_2)^2}\Big[-(\hat{\theta}_1)^2 + [\log(p/(1-p))]^2 + L_p^2(\theta_2)^2 - 2\hat{\theta}_1\log(p/(1-p)) + 2L_p\hat{\theta}_1\hat{\theta}_2 - 2L_p\hat{\theta}_2\log(p/(1-p))\Big].$$

Since asymptotically $E(\hat{L}_p) = L_p$, the variance of $\hat{L}_p$ is

$$(3.2.19)\quad Var(\hat{L}_p) \cong E\left[\hat{L}_p - L_p\right]^2$$

$$= \frac{1}{(\theta_2)^2}\Big[E(\hat{\theta}_1)^2 + [\log(p/(1-p))]^2 + L_p^2\, E(\hat{\theta}_2)^2 - 2\log(p/(1-p))E(\hat{\theta}_1) + 2L_pE(\hat{\theta}_1\hat{\theta}_2) - 2L_p\log(p/(1-p))E(\hat{\theta}_2)\Big]$$

$$= \frac{1}{(\theta_2)^2}\Big[Var(\hat{\theta}_1)+[E(\hat{\theta}_1)]^2+[\log(p/(1-p))]^2 + L_p^2\, Var(\hat{\theta}_2)+L_p^2\,[E(\hat{\theta}_2)]^2 -2\log(p/(1-p))E(\hat{\theta}_1)+2Cov(\hat{\theta}_1,\hat{\theta}_2)L_p +2L_pE(\hat{\theta}_1)E(\hat{\theta}_2)-2L_p\log(p/(1-p))E(\hat{\theta}_2)\Big]$$

$$= \frac{1}{(\theta_2)^2}\Big[\frac{c}{ac-b^2}+(\theta_1)^2+[\log(P/(1-P))]^2 + L_p^2\,\frac{a}{ac-b^2} +L_p^2(\theta_2)^2-2\log(P/(1-P))\theta_1 -2L_p\frac{b}{ac-b^2}+2\theta_1\theta_2L_p-2L_p\log(P/(1-P))\theta_2\Big]$$

$$= \frac{1}{(\theta_2)^2}\Big[\frac{c}{ac-b^2}+(\theta_1)^2+[L_p\theta_2+\theta_1]^2 + L_p\frac{a}{ac-b^2} + L_p^2(\theta_2)^2-2(L_p\theta_2+\theta_1)\theta_1-2L_p\frac{b}{ac-b^2} + 2L_p\theta_2\theta_1-2L_p(L_p\theta_2+\theta_1)\theta_2\Big]$$

$$= \frac{1}{(\theta_2)^2}\left[\frac{c}{ac-b^2} + L_p^2\frac{a}{ac-b^2} - 2L_p\frac{b}{ac-b^2} + [\theta_1+L_p\theta_2]^2 - [\theta_1+L_p\theta_2]^2\right]$$

$$= \frac{aL_p^2 + c - 2bL_p}{(\theta_2)^2(ac-b^2)}$$

If m is large then x_{ij} and $x_i^* \longrightarrow L_{p_i}$ and $P_{ij} \rightarrow \pi_i = \exp(\theta_1+\theta_2 L_{p_i})/(1+\exp(\theta_1+\theta_2 L_{p_i}))$. Then for $n_1=n_2=n$,

(3.2.20) $$a \longrightarrow \sum_{i=1}^{2} (m+n)\pi_i(1-\pi_i)$$

(3.2.21) $$b \longrightarrow \sum_{i=1}^{2} (m+n)L_{p_i}\pi_i(1-\pi_i)$$

(3.2.22) $$c \longrightarrow \sum_{i=1}^{2} (m+n)L_{p_i}\pi_i(1-\pi_i)$$

Now for both m and n large the asymptotic variance given in (3.2.16) becomes

(3.2.23) $$\mathrm{Var}\begin{bmatrix}\hat\theta_1\\ \hat\theta_2\end{bmatrix} = \frac{1}{(m+n)(a'c'-b'^2)}\begin{bmatrix} c' & -b' \\ -b' & a'\end{bmatrix}.$$

This makes the value of (3.2.19)

$$(3.2.24) \qquad Var(\hat{L}_p) = \frac{a' L_p^2 - 2L_p b' + c'}{(m+n)(\theta\ )^2 (a'c' - b'^2)}$$

where $a' = \sum_{i=1}^{2} \pi_i(1-\pi_i)$, $b' = \sum_{i=1}^{2} L_{p_i} \pi_i(1-\pi_i)$ and $c' = \sum_{i=1}^{2} L_{p_i}^2 \pi_i(1-\pi_i)$.

The previous results can be summarized in the following theorems.

THEOREM 1. *Consider the two parameter logit model defined by (3.1.1). Suppose* x_i^* *is the* m^{th} *sequential estimate of* L_{p_i}, $i = 1, 2$ *with* n_i *observations of* y_{ij}, $j = m+1, m+2, \ldots, m+n_i$ *at* x_i^*. *Let* $\hat{\theta}_1$ *and* $\hat{\theta}_2$ *satisfy the conditions of (3.2.6) and (3.2.7). For* n_i *large and any* m, *the asymptotic variance of* $\begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix}$ *is is given by (3.2.16).*

COROLLARY *Assume the conditions of theorem 1 hold. Let* $n_1 = n_2 = n$ *and both* m *and* n *be large. The asymptotic variance of* $\begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix}$ *is given by (3.2.23).*

THEOREM 2. *From (3.2.2) the minimal sufficient statistics for*

$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \text{ is } (y_{11}, y_{12}, \ldots, y_{1m}, y_{21}, y_{22}, \ldots, y_{2m}, \sum_{j=m+1}^{m+n_1} y_{1j}, \sum_{j=m+1}^{m+n_2} y_{2j}).$$

THEOREM 3. *Assume the conditions of theorem 1 hold. Then for* *large n_i (i=1,2) and any m, the asymptotic variance of the TSE of L_p is given by (3.2.19).*

<u>COROLLARY</u> *Let the conditions of theorems 1 and 2 be satisfied and $n_1 = n_2 = n$. If both m and n are large then the asymptotic variance of the TSE of L_p is given by (3.2.24).*

Optimal Selection of p_1 and p_2

In this section we will search for the optimum values of p_1 and p_2 that minimize (3.2.19) under certain conditions.

Suppose we are interested in estimating L_p where $p \epsilon \left[\nu, 1-\nu\right]$, $0 < \nu < \frac{1}{2}$ and we use $p_1 = p^*$ and $p_2 = 1 - p^*$. Assume that f(.) is uniform density function over the interval $(\nu, 1-\nu)$. This means we are equally interested in each quantile L_p, $P \in [\nu, 1-\nu]$. We want to choose p^* such that

$$\int_{\nu}^{1-\nu} \mathrm{Var}(\hat{L}_p) f(p)\, dp \tag{3.3.1}$$

is minimum where

$$f(p) = \frac{1}{(1-\nu)-\nu} = \begin{cases} 1 & \text{if } p \in [\nu, 1-\nu] \\ 0 & \text{otherwise} \end{cases} \tag{3.3.2}$$

and $Var(\hat{L}_p)$ is as in (3.2.19). Without loss of generality consider the case $\theta_1=0$ and $\theta_2 = 1$. Note that if $p_1 = p^*$, $p_2 = 1-p^*$, then $L_{p_1} = L_{p^*}$, $L_{p_2} = L_{1-p^*} = -L_{p^*}$. Hence $a' = 2p^*(1-p^*)$ $b' = 0$ and $c' = 2p^*(1-p^*)L^2_{p^*}$. Then (3.2.19) becomes

$$(3.3.3) \quad Var(\hat{L}_p) = \frac{2p^*(1-p^*)[(\log(p^*/(1-p^*))^2-(\log(p/(1-p))^2]}{(m+n)[2p^*(1-p^*)]^2[\log(p^*/(1-p^*)]^2}$$

$$= [2(m+n)p^*(1-p^*)]^{-1}\left[1+\frac{[\log(p/(1-p))]^2}{[\log(p^*/(1-p^*))]^2}\right].$$

Therefore, it is sufficient to find p^* that minimizes

$$(3.3.4) \qquad \int_{\nu}^{1-\nu} [2p^*(1-p^*)]^{-1}\left[1+\frac{[\log(p/(1-p))]^2}{[\log(p^*/(1-p^*))]^2}\right]dp \ .$$

Faries (1990) has shown that (3.3.1) is minimized if $p^*=.2$. Hence the optimal choices of p_1 and p_2 that minimize the variance are .2 and .8, respectively. This is illustrated in Figure 2.

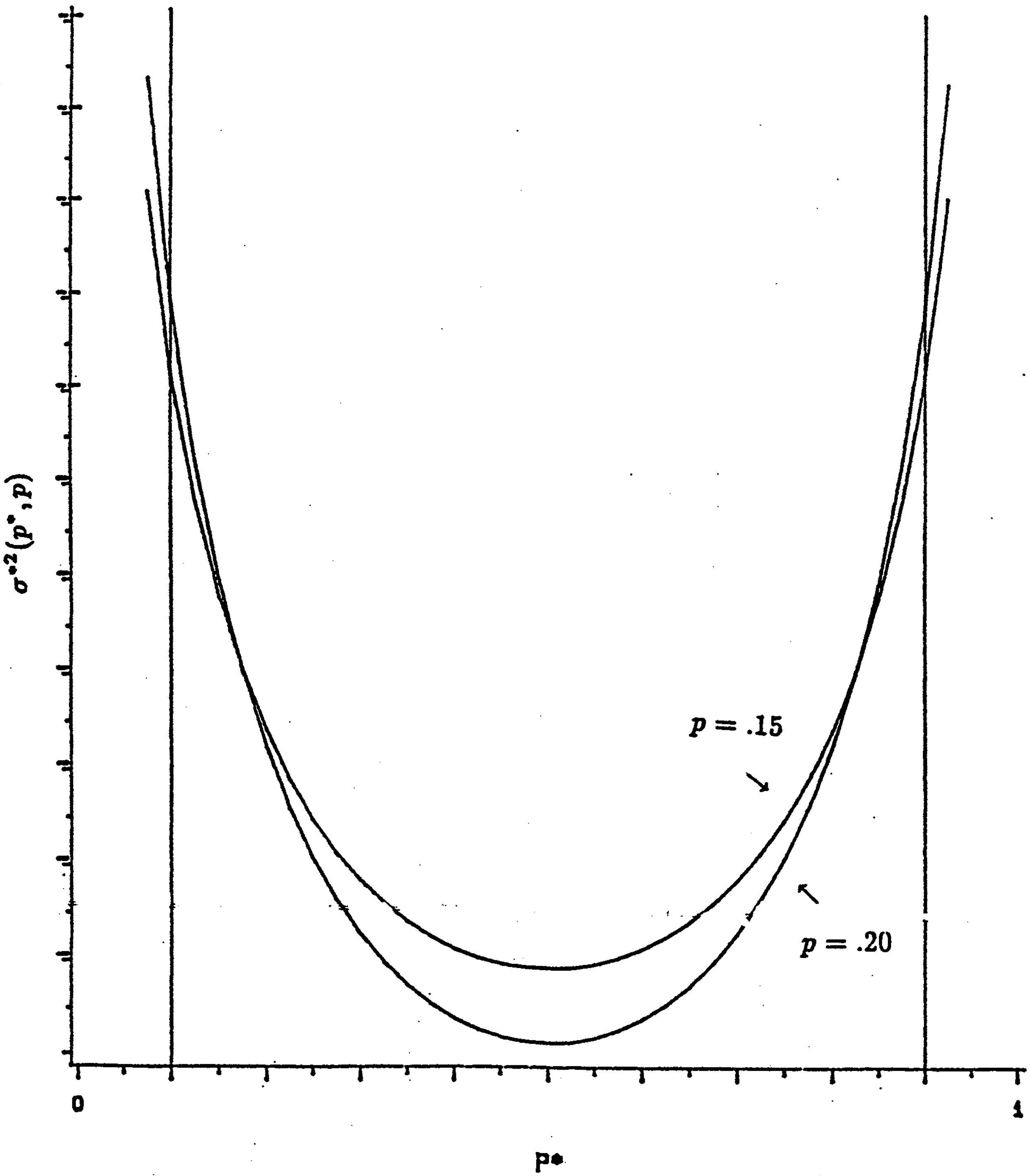


FIGURE 2 AVERAGE MINIMUM p

Asymptotic Distribution of the TSE of L_p

A sequence of solutions $\{\hat{\theta}_M\} = \{(\hat{\theta}_1,\hat{\theta}_2)'_M\}$ of the likelihood equations have an asymptotic normal distribution if the following regularity conditions are satisfied (Serfling 1980). Note that in this section asymptotic refers to large n_i and any m.

Let Θ be an open interval in R and $\ell(\theta) = \log(f(y;\theta))$, where $f(y;\theta)$ is the pmf of the random variable y.

(R1) For each $\theta \in \Theta$, the derivatives $\frac{\partial \ell(\theta)}{\partial\theta}$, $\frac{\partial^2 \ell(\theta)}{\partial\theta^2}$, $\frac{\partial^3 \ell(\theta)}{\partial\theta^3}$ exist, all y.

(R2) For each $\theta_0 \in \Theta$, there exist functions q(y), h(y) and Z(y) such that for θ in a neighborhood $N(\theta_M)$ the relations $\left|\frac{\partial f(y;\theta)}{\partial\theta}\right| \le q(y)$, $\left|\frac{\partial^2 f(y;\theta)}{\partial\theta^2}\right| \le h(y)$, $\left|\frac{\partial^3 \ell(\theta)}{\partial\theta^3}\right| \le Z(y)$ hold, all y, and $\sum q(y) < \infty$, $\sum h(y) < \infty$, $E_\theta\{Z(y)\} < \infty$ for $\theta \in N(\theta_0)$;

(R3) For each $\theta \in \Theta$, $0 < -E_\theta\left\{\frac{\partial^2 \ell(\theta)}{\partial\theta^2}\right\} < \infty$.

From (3.2.4), (3.2.5) and (3.2.12) we can see that conditions (R1) and (R3) hold. To check condition (R2) note that $f(y;\theta_1,\theta_2) = P^y(1-P)^{1-y}$ where $P = Pr(Y(x)=1)$ given by (1.14). $\left|\frac{\partial}{\partial\theta_1}\{f(y;\theta_1,\theta_2)\}\right| = \left|P^y(1-P)^{1-y}(y-P)\right| \le y.$

Also $\left|\frac{\partial^2}{\partial\theta_1^2}\{f(y;\theta_1,\theta_2)\}\right| = \left|P^y(1-P)^{1-y}\{y\ (y-P)+2P^2-P(y+1)\}\right|$

$\le y^2+2$ and $\left|-\frac{\partial^3 \ell(\theta_1,\theta_2)}{\partial\theta_1^3}\right| = \sum_{i=1}^{2}\sum_{j=1}^{m}\{P_{ij}(1-P_{ij}) \quad -2P_{ij}[1-P_{ij}]^2\}$

$$+\sum_{i=1}^{2}\{P_i(1-P_i)-2P_i[1-P_i]^2\}.$$

Setting $q(y) = y$, $h(y) = y^2+2$ and

$$Z(y) = \sum_{i=1}^{2}\sum_{j=1}^{m}\{P_{ij}(1-P_{ij})-2[P_{ij}]^2[1-P_{ij}]\ \}+$$
$$\sum_{i=1}^{2}\{P_i(1-P_i)-2[P_i]^2[1-P_i]\ \},$$

we note that $\Sigma\, q(y) = \Sigma\, y_i < \infty$, $\Sigma\, h(y) = \Sigma\,(y_i^2+2) < \infty$ where the summation is taken over the sample size of interest. Also note that $E\{Z(y)\} < \infty$. Similarly, these properties hold for θ_2, so, the three regularity conditions are satisfied. Thus, the solution $\begin{bmatrix}\hat{\theta}_{1M}\\ \hat{\theta}_{2M}\end{bmatrix}$ of the maximum likelihood equations (3.2.6) and (3.2.7) is asymptotically normal with asymptotic mean $\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix}$ and variance given by

(3.2.16). Set $\text{Var}\begin{bmatrix} \hat{\theta}_{1M} \\ \hat{\theta}_{2M} \end{bmatrix} = \sum$.

Hence

(3.4.1) $$\begin{bmatrix} \hat{\theta}_{1M} \\ \hat{\theta}_{2M} \end{bmatrix} \sim AN\left(\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}, \sum \right).$$

To find the asymptotic distribution of $\hat{L}_p$ given in (3.1.2), we note that $\hat{L}_p$ is a real valued function of $\hat{\theta}_1$ and $\hat{\theta}_2$ from $R^2 \longrightarrow R$ which may be expressed as

(3.4.2) $$\hat{L}_p(\hat{\theta}_1, \hat{\theta}_2) = \frac{-\hat{\theta}_1 + \log(p/(1-p))}{\hat{\theta}_2}.$$

$$\frac{\partial \hat{L}_p(\hat{\theta}_1, \hat{\theta}_2)}{\partial \hat{\theta}_1} = -\frac{1}{\hat{\theta}_2} \quad \text{and} \quad \frac{\partial \hat{L}_p(\hat{\theta}_1, \hat{\theta}_2)}{\partial \hat{\theta}_2} = -\frac{[-\theta_1 + \log(p/(1-p))]}{(\hat{\theta}_2)^2}$$

$$\text{Then } D = \left(\frac{\partial \hat{L}_p(\hat{\theta}_1, \hat{\theta}_2)}{\partial \hat{\theta}_1}, \frac{\partial \hat{L}_p(\hat{\theta}_1, \hat{\theta}_2)}{\partial \hat{\theta}_2} \right) \Bigg|_{(\hat{\theta}_1, \hat{\theta}_2) = (\theta_1, \theta_2)}$$

$$= \left(\frac{-1}{\theta_2}, -\frac{(-\theta_1 + \log(p/(1-p)))}{(\theta_2)^2} \right).$$

From large sample theory (Serfling, 1980, p.123), we know that $\hat{L}_p(\hat{\theta}_1, \hat{\theta}_2)$ is asymptotically normally distributed with mean $L_p(\theta_1, \theta_2)$ and variance $D \sum D'$. Hence we compute $D \sum D'$ below.

$$D \sum D' = \left[\frac{-1}{\theta_2}, -\frac{(-\theta_1+\log(p/(1-p)))}{(\theta_2)^2}\right] \frac{1}{(ac-b^2)} \begin{bmatrix} c & -b \\ -b & a \end{bmatrix} D'$$

$$= \left[\frac{-c}{\theta_2} + \frac{b}{(\theta_2)^2}(-\theta_1+\log(p/(1-p))),\right.$$

$$\left.\frac{b}{\theta_2} - \frac{a}{(\theta_2)^2}(-\theta_1+\log(p/(1-p)))\right] \left[\begin{array}{c} -1/\theta_1 + \\ -(-\theta_1 \\ \log(p/(1-p)))/(\theta_2)^2 \end{array}\right] \frac{1}{(ac-b^2)}$$

$$= \frac{1}{(ac-b^2)} \left(\frac{c}{(\theta_2)^2} - \frac{2b}{(\theta_2)^2}\left\{\frac{-\theta_1+\log(p/(1-p))}{\theta_2}\right\} + \frac{a}{(\theta_2)^2}\left[\frac{-\theta_1+\log(p/(1-p))}{\theta_2}\right]^2\right)$$

$$= \frac{1}{(ac-b^2)} \left(\frac{c}{(\theta_2)^2} - \frac{2b}{(\theta_2)^2} L_p + \frac{a}{(\theta_2)^2} L_p^2\right)$$

$$= \frac{aL_p^2 - 2bL_p + c}{(\theta_2)^2(ac - b^2)}$$

since $(-\theta_1+\log(p/(1-p)))/\theta_2 = L_p$. It is interesting to note that the last quantity is equal to (3.2.19), as it should be. Hence, the following result has been established.

THEOREM 4. *In the presence of the conditions of Theorem 1*

$$\hat{L}_p \sim AN\left(L_p, \frac{aL_p^2 - 2bL_p + c}{(\theta_2)^2(ac - b^2)}\right).$$

CHAPTER IV

SIMULATION STUDY

The Setup

In the first two chapters asymptotic properties of the sequential and two-stage procedures have been presented. In this chapter, a simulation study is performed to study the properties of the two-stage procedure and compare it with the sequential procedure for small binary logit data samples. The SAS code (SAS Institute, Inc.) for carrying out the simulation operation is presented in Appendix B, where the sections of the program that perform the two stage and sequential methods are given beginning at pages 103 and 107, respectively. The simulation process creates 1000 samples each, for total samples of sizes 20, 40, 80 and 160 observations, for each of the two procedures, paying attention to the sizes of m and n in each sample. For each of these samples, estimates of L_{p_1} and L_{p_2} are obtained using both methods of estimation where the pair (p_1,p_2) is (.2,.8), (.35,.65), (.75,.9), (.9,.95). Consequently, the roots L_p for p = .10, .25, .50, .75 and .90 are estimated. Without loss of generality the values of Y_{ij} are generated using the logit model with $\theta_1 = 0$ and $\theta_2 = 1$. For each sample, the existence of estimates of L_p depend on the existence of the maximum likelihood estimates of θ_1 and θ_2. Silvapulle (1981) has established that the maximum likelihood estimators of (θ_1, θ_2) exist, and are unique, if and only if,

$$(4.1) \qquad (x^+_{min}, x^+_{max}) \cap (x^-_{min}, x^-_{max})$$

is nonempty or

$$(4.2) \qquad x^+_{min} < x^-_{min} = x^-_{max} < x^+_{max}$$

or

$$(4.3) \qquad x^-_{min} < x^+_{min} = x^+_{max} < x^-_{max},$$

where $x^+_{max} = \max\{ x_i : y_i = 1 \}$, $x^-_{max} = \max\{ x_i : y_i = 0 \}$, $x^+_{min} = \min\{ x_i : y_i = 1 \}$, and $x^-_{min} = \min\{ x_i : y_i = 0\}$.

The part of the SAS program that checks Silvapulle's condition is given in Appendix B on pages 103 and 104. For each method of estimation, the sample mean, sample variance, and bias of $\hat{L}_p$ is calculated as

$$(4.4) \qquad \widehat{Mean}(\hat{L}_p) = \sum_{i=1}^{I} (\hat{L}_p)_i / I$$

$$(4.5) \qquad \widehat{Var}(\hat{L}_p) = \frac{1}{(I-1)} \left\{ \sum_{i=1}^{I} (\hat{L}^2_p)_i - I \, [Mean(\hat{L}_p)]^2 \right\}$$

(4.6) $\hat{Bias}(\hat{L}_p) = Mean(\hat{L}_p) - L_p$

where i = 1, 2, 3, ...,I, I≤ 1000 and I is the number of simulated samples for which $\hat{L}_p$ exists.

Results of the Simulation Study

Variance Estimates of $\hat{L}_p$

Tables 2 and 3 present the estimated variances of the estimates of $L_{.10}$, $L_{.25}$, $L_{.50}$, $L_{.75}$, and $L_{.90}$ for both two stage and sequential procedures for various sample sizes and starting values.

If the total sample size is 20 (m = 5, n = 5), the estimates of the variance for the TSE of $L_{.50}$ are .369, .268, 3.752, and 2.428 provided that the pairs of initial values ($L_{.2}$, $L_{.8}$), ($L_{.35}$, $L_{.65}$), ($L_{.75}$, $L_{.9}$) and ($L_{.9}$, $L_{.95}$) are used, respectively. The corresponding estimates of the variances of $\hat{L}_{.50}$ for the sequential procedure are .253, .211, .464, and .489, respectively. For the two stage procedure, the estimated variances of $\hat{L}_{.50}$ are smaller and closer to each other if we use the pair of initial values ($L_{.2}$, $L_{.8}$) and ($L_{.35}$, $L_{.65}$), than if we use the pair ($L_{.75}$, $L_{.9}$) and ($L_{.9}$, $L_{.95}$). Similarly, for the total sample size of 20, the variance estimates for the sequential method are smallest for starting values ($L_{.2}$,$L_{.8}$) and ($L_{.35}$,$L_{.65}$), with the smallest variance estimate being obtained for $(L_{.2}, L_{.8})$. If the total sample size is 40 ($m = 5$, $n = 15$) the variance estimates for the TSE of $L_{.50}$ are .211, .199, .993, and 3.502 where the pair of initial values $(L_{.2}, L_{.8})$, $(L_{.35}, L_{.65})$, $(L_{.75}, L_{.9})$ and $(L_{.9}, L_{.95})$ are used, respectively. The associated estimates of variances of $\hat{L}_{.50}$ for the sequential method are .121, .129, .290 and .345, respectively. The variance estimates of $\hat{L}_{.50}$ have been reduced for both procedures as a result of an increase in total sample size. For both methods, the smallest and the largest variance estimates are obtained if the pair of initial values $(L_{.2}, L_{.8})$ and $(L_{.9}, L_{.95})$ are used, respectively. The variance estimates of $\hat{L}_{.50}$ are smaller for the sequential method than for the two stage method. Similar properties hold for the variance estimate of $\hat{L}_{.50}$ for various total sample sizes such as 40, and 80. That is, both the two stage and sequential methods provide small variance estimates of $\hat{L}_{.50}$ for symmetrical initial values $(L_{.2}, L_{.8})$ and $(L_{.35}, L_{.65})$. Larger variance estimates of $\hat{L}_{.50}$ are obtained if the asymmetric initial values $(L_{.75}, L_{.9})$ and $(L_{.9}, L_{.95})$ are used. The variance estimate decreases as the total sample size increases. The most interesting result of the variance estimate of $\hat{L}_{.50}$ is obtained when the total sample size is 160 ($m = n = 40$). Using the four pairs of initial values $(L_{.2}, L_{.8})$, $(L_{.35}, L_{.65})$, $(L_{.75}, L_{.9})$ and $(L_{.9}, L_{.95})$, the variance estimates of $\hat{L}_{.50}$ for the two stage and sequential methods are, respectively, .023, .053, .131, .164 and .023, .053, .130, and .163. This shows that if we take 160 observations with the same initial values, then the variance estimates for both methods are nearly equal.

Now consider estimating $L_{.10}$ with each of the methods using each of the pairs of initial values ($L_{.2}$, $L_{.8}$), ($L_{.35}$, $L_{.65}$), ($L_{.75}$, $L_{.9}$) and ($L_{.9}$, $L_{.95}$).

If the total sample size is 20, the variance estimates of $\hat{L}_{.10}$ for the two stage method are 13.567, 20.096, 20.929 and 17.068 using the initial values in the previous order. Correspondingly, the variance estimates of $\hat{L}_{.10}$ for the sequential method are 1.698, 1.528, 2.201 and 1.843. The estimates of the variance for the sequential method are smaller than those of the two stage method. However, as we increase the total sample size, the variance estimate of $\hat{L}_{.10}$ reduces faster for the two stage than for the sequential method. When the total sample size reaches 160, we find that the variance estimates of $\hat{L}_{.10}$ for the two stage method are .406, .490, .599, and .491, corresponding to the initial values ($L_{.2}$, $L_{.8}$), ($L_{.35}$, $L_{.65}$), ($L_{.75}$, $L_{.9}$) and ($L_{.9}$, $L_{.95}$). Taking the initial values in the same order, the corresponding variance estimates of $\hat{L}_{.10}$ for the sequential method are .404, .489, .591, and .486. Hence we note that the estimated variances are nearly equal for both methods.

With regard to estimating $L_{.25}$ using a total sample size of 20, the variance estimates of the two stage method are 3.830, 5.204, 10.431, and 7.931. For these estimates, the starting values $(L_{.2}, L_{.8})$, $(L_{.35}, L_{.65})$, $(L_{.75}, L_{.9})$ and $(L_{.9}, L_{.95})$ are used, respectively. For the sequential method with the same number of observations and initial values used in the same order, we obtain the variance estimates of $\hat{L}_{.25}$ to be .622, .538, 1.067 and .969. Using $(L_{.2}, L_{.8})$ and $(L_{.35}, L_{.65})$ as initial values will produce small variance estimates of $\hat{L}_{.25}$ for both methods. For any starting value, the variance estimates of $\hat{L}_{.25}$ decrease as we increase the total number of observations. When the total number of observations reaches 160, the variance estimates from both methods are nearly equal. At this stage the variance estimates of $\hat{L}_{.25}$ for the two stage method are .145, .180, .291, and .224 whereas the corresponding variance estimates for the sequential method are .144, .178, .288, and .273.

Now consider estimating $L_{.75}$ and $L_{.9}$ with a total of 20 observations. In this case, small variance estimates are obtained if we use the two pairs of initial values $(L_{.75}, L_{.9})$ and $(L_{.9}, L_{.95})$. This is due to the fact that the initial values are chosen near the roots to be estimated. For total observations up to 80, the variance estimates for the sequential method are smaller than that of the two stage method. For both methods, the variance estimates of the estimator of either root decrease as we increase the total sample size. If the total sample size is 160, the variance estimates of $\hat{L}_{.75}$ for the two stage method are .152, .148,

.105, and .154, whereas for the sequential method the variance estimates of $\hat{L}_{.75}$ are .150, .146, .101, and .151, using the initial values ($L_{.2}$, $L_{.8}$), ($L_{.35}$, $L_{.65}$), ($L_{.75}$, $L_{.9}$) and ($L_{.9}$, $L_{.95}$), respectively. For the same total sample size, the variance estimates of $\hat{L}_{.90}$ for the two stage method are .425, .450, .226 and .247, but for the sequential method these variance become .423, .444, .222, and .243. Thus we see that for both roots, $L_{.75}$ and $L_{.90}$, the two stage and the sequential methods give nearly equal estimated variances.

Bias Estimates of $\hat{L}_p$

Table 4 provides the bias estimates for the estimators of $L_{.10}$, $L_{.25}$, $L_{.50}$, $L_{.75}$, and $L_{.90}$ for both the two stage and sequential procedures. The biases for the sequential method are shown in brackets.

Consider estimating the bias of $\hat{L}_{.50}$ using a total of 20 ($m = n = 5$) observations. For the two stage method the bias estimators are .000, -.006, .291, and .086 whereas for the sequential method the corresponding bias estimates are -.010, .008, .328 and .453, using ($L_{.2}$, $L_{.8}$), ($L_{.35}$, $L_{.65}$), ($L_{.75}$, $L_{.9}$) and ($L_{.9}$, $L_{.95}$), respectively. Thus we note that for this total sample size the TSE of $L_{.50}$ is less biased than the sequential estimator. If we take a sample of 80 or more observations, the two stage estimator of $L_{.50}$ is essentially unbiased. The bias estimate for the sequential estimator of $L_{.50}$ is generally greater than that of the TSE. Thus we note that the TSE of $L_{.50}$ is less biased than the sequential estimator. Using $(L_{.35}, L_{.65})$ as initial values to estimate $L_{.50}$ gives the bias estimates of -.001 and .005 for the two stage and sequential methods, respectively, if 160 observations are taken. If we use $(L_{.9}, L_{.95})$ as initial values, the smallest bias estimate of $\hat{L}_{.50}$ is -.023 for 160 observations. The corresponding sequential estimator of $\hat{L}_{.50}$ gives bias estimate of .147. So again, the TSE of $L_{.50}$ is less biased.

Now consider estimating the biases of $\hat{L}_{.10}$ and $\hat{L}_{.90}$. Suppose 20 observations are taken with the initial values $(L_{.2}, L_{.8})$, the bias estimates of the TSE of $L_{.10}$ and $L_{.90}$ are .598 and -.599, respectively. The corresponding bias estimates for the sequential method are .551 and -.572. As we increase the the total sample size, we find that the bias estimates steadily decrease in absolute value, reaching minimum values of .098 and -.100 for $\hat{L}_{.10}$ and $\hat{L}_{.90}$ for the two stage procedure, respectively, when 160 observations are taken. With 20 observations the estimators of $L_{.10}$ and $L_{.90}$ are more biased for the two stage than the sequential estimators. However, for observations of more than 40, we find that the TSEs of $L_{.10}$ and $L_{.90}$ are less biased. If we use $(L_{.35}, L_{.65})$ as initial values, we note that the bias estimates for the TSE of $L_{.10}$ and $L_{.90}$ are always less than that of the sequential estimator. The bias estimates for both methods generally decrease as we increase the total sample size. We obtain the smallest bias for 160 observations regardless of method used. Using ($L_{.75}$, $L_{.9}$) as initial values with 20 observations, we obtain the bias estimates of the TSE of $L_{.10}$ and $L_{.90}$ as .962 and -.380, respectively. The corresponding sequential estimates are .661 and -.012. Thus the biases for the two stage method are greater in absolute value than that of the sequential method. However, if we take 80 observations with m = 20 and n = 20 or more and ($L_{.75}$,$L_{.9}$) as starting values, we find that the bias estimate for the TSE of $L_{.10}$ is smaller. With ($L_{.9}$, $L_{.95}$) as the starting values with 20 observations, the bias estimates of the TSEs of $L_{.10}$ and $L_{.90}$ are .361 and -.189, respectively. The corresponding bias estimates for the sequential method are .556 and .350. As we increase the number of observations, the bias estimate for the TSE of $L_{.90}$ steadily decreases in absolute value reaching its minimum for 160 observations. On the other hand, the bias estimate for the TSE of $L_{.10}$ generally decreases. The biases for the sequential method are always greater than the biases for the two stage method when ($L_{.9}$, $L_{.95}$) are used as initial values.

If the initial values ($L_{.2}$, $L_{.8}$) are used, and we increase the sample size from 20 to 160, the bias estimates of both $\hat{L}_{.25}$ and $\hat{L}_{.75}$ decrease for the two stage method. In general the bias estimates for the sequential method follow a similar pattern as two stage method. For 80 or more observations, the bias estimates for the TSEs of $L_{.25}$ and $L_{.75}$ are less in absolute value than those of the sequential estimator. If $(L_{.35}, L_{.65})$ are used as initial values, the TSEs of $L_{.25}$ and $L_{.75}$ are less biased than those of the sequential estimator for all observations considered. With $(L_{.75}, L_{.9})$ as initial values, the biases for the TSE of $L_{.25}$ are less than the biases for the sequential estimator of $L_{.25}$ for 40 or more observations. However, the TSE of $L_{.75}$ is less biased than the sequential estimator of $L_{.75}$ for all observations considered. For starting values $(L_{.9}, L_{.95})$, the bias estimates of the TSEs of $L_{.25}$ and $L_{.75}$ are always less than the bias estimates of the sequential estimator.

Simulated Frequency Distribution of $\hat{L}_p$

As can be seen in Appendix B, each root estimator is standardized according to

$$(\hat{L}_p - L_p)/ \text{std}(\hat{L}_p)$$

so that its relative frequency distribution can be compared with that of the standard normal random variable. The frequency distributions of the estimated roots $L_{.10}$, $L_{.25}$, $L_{.50}$, $L_{.75}$, and $L_{.90}$ is divided into 26 cells: Cell 1 from $-\infty$ to -3, cell 2 from -3 to -2.75, cell 3 from -2.75 to -2.50 cell 4 from -2.5 to -2.25, ..., cell 25 from 2.75 to 3 and cell 26 from 3 to ∞. For each of the two methods, the standardized frequency distributions are given in Figures 3 through 42 for 20 and 160 observations using the initial values ($L_{.2}$, $L_{.8}$) and ($L_{.9}$, $L_{.95}$).

Figures 3 and 4 give the frequency distributions of $\hat{L}_{.10}$ for the two stage and sequential methods for 20 observations and ($L_{.2}$, $L_{.8}$) as initial values. The distributions of both estimators are skewed to the right. However, the probability of each estimator taking a value to the right of 2.25 is small. Now consider the distributions of both estimators of $L_{.10}$ when 160 observations are taken. These are shown in Figures 13 and 14 for the two stage and sequential methods, respectively. The graph of the TSE of $L_{.10}$ is more symmetric with respect to 0 than the graph of the sequential estimator. Figures 23 and 24 present the graphical representation of $\hat{L}_{.10}$ using 20 observations with ($L_{.9}$, $L_{.95}$) as initial values. The tail probabilities for the two stage method are approximately equal, whereas this property is not true for the sequential method. The probability of the sequential estimator of $L_{.10}$ taking a value less than -3 is 0. For 160 observations, the two stage and sequential estimators of $L_{.10}$ are graphically depicted in Figures 33 and 34. These graphs look more alike than the graphs we obtain using 20 observations with ($L_{.9}$, $L_{.95}$) as initial values.

Figures 5 and 6 give the graphical pictures of $\hat{L}_{.25}$ for both methods, with 20 observations and using ($L_{.2}$, $L_{.8}$)

as initial values. As we see from the figures, these estimators are skewed to the right and have similar tail probabilities. In Figures 15 and 16, we have the graphs of the two stage and sequential estimators of $L_{.25}$, respectively, for 160 observations. For both estimators the graphs are skewed to the right. However, the degree of skewness is smaller for 160 observations than for 20. Figures 25 and 26 present the frequency distribution of the two stage and sequential estimators of $L_{.25}$, respectively, for 20 observations, with $(L_{.9}, L_{.95})$ as the initial values. The TSE of $L_{.25}$ is distributed over $(-\infty, \infty)$, but the sequential estimator of $L_{.25}$ takes no value below -3. Figures 35 and 36 give the frequency distributions for the same situation but with 160 observations. It can be seen from Figure 35, that the TSE of $L_{.25}$ is symmetric with respect to 0, but that the sequential estimator is not.

For 20 observations with $(L_{.2}, L_{.8})$ as initial values, the frequency distribution of the two stage and sequential estimators of $L_{.50}$ are shown in Figures 7 and 8, respectively. The distributions of both estimators are symmetric with respect to 0. Also, for both distributions, the probabilities to the left of -3 and to the right of 3 are small. However, the distribution of the TSE of $L_{.50}$ is more dispersed than that of the sequential estimator. For 160 observations, the frequency distributions of $L_{.50}$ for both methods are given in Figures 17 and 18. In this case both distributions satisfy the properties of the standard normal variable. For 20 observations with $(L_{.9}, L_{.95})$ as initial values, the frequency distributions of the two stage and sequential estimators of $L_{.50}$ are displayed in Figures 27 and 28, respectively. The TSE of $L_{.50}$ is distributed over $(-\infty, \infty)$ with the probability of the right tail slightly more than that of the left tail. The sequential estimator of $L_{.50}$ is not distributed below -2.25 at all and its right tail probability is larger than that of the TSE. For 160 observations, the distributions are given in Figures 37 and 38 . The TSE of $L_{.50}$, shown in Figure 37, is symmetric with respect to 0 and has approximately equal tail probabilities. However, the sequential estimator of $L_{.50}$ has larger right tail probability than left tail. The increased sample size has resulted in the TSE of $L_{.50}$ to take a bell shape, whereas for the sequential estimator, the effect is not observable.

With initial values $(L_{.2}, L_{.8})$, the estimated frequency distributions of $\hat{L}_{.75}$ are given in Figures 9 and 10 for 20 observations and in Figures 19 and 20 for 160 observations. These distributions show properties identical to $\hat{L}_{.25}$, except that the distribution of $\hat{L}_{.75}$ is skewed to the left. Similarly, the frequency distributions of $\hat{L}_{.90}$, given in Figures 11 and 12 for 20 observations and in Figures 21 and 22 for 160 observations, are identical to those $\hat{L}_{.10}$, except that the distribution of $\hat{L}_{.90}$ is skewed to the left.

CHAPTER V

CONCLUSION

The two stage procedure represents a new approach to estimating multiple roots of the logit model. It can be used to estimate multiple roots of an expectation of other similar binary models such as the probit and log log models, as given in Appendix A.

Both the two stage and sequential methods provide small variance estimates of $L_{.50}$ for symmetric initial values $(L_{.2}, L_{.8})$ and $(L_{.35}, L_{.65})$. Larger variance estimates are obtained if the asymmetric initial values $(L_{.75}, L_{.9})$ and $(L_{.9}, L_{.95})$ are used. For a given sample size and any initial value the maximum variance estimate for both methods is obtained by estimating $L_{.10}$, not because $L_{.10}$ has some bad properties but because $L_{.10}$ was the root always farthest away from the starting values. Using $(L_{.2}, L_{.8})$ and $(L_{.35}, L_{.65})$ as initial values for both methods will give smaller variance estimates of $L_{.25}$ than $(L_{.75}, L_{.9})$ and $(L_{.9}, L_{.95})$. Small variance estimates of $L_{.75}$ and $L_{.9}$ are obtained if we use $(L_{.75}, L_{.9})$ and $(L_{.9}, L_{.95})$ as initial values. This is because the initial values are near the roots to be estimated. In general, for any starting value, the variance estimates of the roots from the two stage decrease as we increase the total number of observations. When the total number of observations reaches 160, the variance estimates from both methods are nearly equal.

With any starting value and using 80 or more observations, the two stage estimators of $L_{.10}$, $L_{.25}$, $L_{.50}$, $L_{.75}$, and $L_{.90}$ are less biased than the sequential estimators. For 80 observations with $m = n = 20$ and using $(L_{.2}, L_{.8})$ as the initial values, we find that the two stage estimator of $L_{.50}$ is unbiased whereas the corresponding sequential estimator is not unbiased. If $(L_{.35}, L_{.65})$ are used as the initial values, the two stage estimators of $L_{.25}$ and $L_{.75}$ are less biased than those of the sequential estimator for all observations considered. With $(L_{.9}, L_{.95})$ as the initial values, the bias estimates of the two stage estimators of $L_{.10}$ and $L_{.90}$ are smaller than those of the sequential estimators for all the sample sizes considered in the simulation study.

With 20 observations and $(L_{.2}, L_{.8})$ as the initial values, the relative frequency distributions of both estimators of $L_{.50}$ are symmetric with respect to 0. Also, for both distributions, the probabilities to the left of -3 and to the right of +3 are small. For each method, with 160 observations, we note that:

$$E[\hat{L}^*_{.50}] = 0,$$

$$Pr\{-1 < \hat{L}^*_{.50} < 1\} \approx .68,$$

$$Pr\{-2 < \hat{L}^*_{.50} < 2\} \approx .95,$$

$$Pr\{-3 < \hat{L}^*_{.50} < 3\} \approx .99,$$

where $\hat{L}^*_{.50}$ is standardized estimator of $\hat{L}_{.50}$. It follows that for observations of 80 or more, both estimators of $L_{.50}$ approach the normal distribution. Using $(L_{.9}, L_{.95})$ provides a more skewed distribution for the sequential estimator than the two stage. The relative frequency distributions of $\hat{L}_{.10}$, $\hat{L}_{.25}$, $\hat{L}_{.75}$, and $\hat{L}_{.90}$ give skewed distributions. The degree of skewness is improved more for the two stage estimator than for the sequential estimator when the number of observations is increased from 20 to 160.

BIBLIOGRAPHY

Albert, A. E and Gardner, L. A. (1967). Stochastic Approximation and Nonlinear Regression. Research Monograph No. 42 MIT Press.

Anbar, D. (1977). "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.

Burkill, J. R. and Burkill, H. (1970). A Second Course in Math.Analysis, Cambridge University Press, New Delhi.

Cocrhan, W. G. (1977). Sampling Techniques, 3rd ed., John Wiley and Sons, Inc., New york.

Chung, K, L. (1977). "On Stochastic Approximation Method," Annals *of Mathematical Statistics*, 25, 463-483.

Dobson, A. J. (1990). 2nd ed., An Introduction to Generalized Linear Models, Chapman and Hall, London.

Dvoretzky, A. (1956)." On Stochastic Approximation," Proc. Third Berkeley Symposium ," *Math. Stat. Prob.,1*, 39-56.

Faries, D. (1990). "A New Sequential Approximation Method," Ph.D. Dissertation, Department of Statistics, Oklahoma State University.

Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). Introduction to the Theory of Statistics, McGraw-Hill, Tokyo.

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

Moser, B. K. and Payton, M. E. (1992). "Hypothesis Testing for the Quantiles of the Logit Model," *Submitted to a Journal for Publication.*

Rao, C. R. (1973). Linear Statistical Inference and Its Applications, John Wiley & Sons, New York.

Robbins, H. and Monro, S. (1951). "A Stochastic Approximation," *Annals of Mathematical Statistics*, 29, 400-407.

Robbins, H. and Siegmund, D. (1971). "A Convergent Theorem for Non-negative Almost Supermartingales and Some Applications," *Opti. Methods in Stat.* 233-257.

Sacks, J. (1958). "Asymptotic Distribution of Schismatic Approximation Procedures," *Annals of Mathematical Statistics*, 29, 373-405.

Sakrison, D. J. (1965). "Efficient Recursive Estimation: Application to Estimate the Parameters of a Covariance Function," *Int. J. Eng. Sci.*, 3, 461-483.

SAS Institute Inc., (1985), SAS User's Guide : Basics, Version 5 Edition, Cary N.C..

Serfling, R. J. (1980). Approximation Theorems of Mathematical Statistics, John Wiley & Sons , New York.

Silvapulle, M. J. (1981). "On the Existence of Maximum Likelihood Estimators for the Binomial Response Model," *Journal of the Royal Statistical Society, Series B*, 43, 310-313.

Ventor, J. H. (1967) "An Extension of the Robbins-Monro Procedure," *Annals of Mathematical Statistics*, 38, 181 - 190.

Wetherill, G.B.(1963). "Sequential Estimation of Quantal Response Curves," *Journal of Royal Society, Series B*, 25, 1-48.

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

TABLE 2

SIMULATION RESULTS

ESTIMATED VARIANCES FOR TWO STAGE METHOD

INITIAL VALUES

TSS	ROOTS	(L.2,L.8)	(L.35,L.65)	(L.75,L.9)	(L.9,L.95)
	L10	13.567	20.096	20.929	17.068
	L25	3.830	5.204	10.431	7.931
m=5	L50	0.369	0.268	3.752	2.428
n=5	L75	3.182	5.288	0.890	0.559
	L90	12.269	20.263	1.846	2.323
	L10	1.750	4.315	5.858	12.543
	L25	0.634	1.249	2.850	7.266
m=5	L50	0.211	0.199	0.993	3.502
n=15	L75	0.481	1.166	0.288	1.251
	L90	1.443	4.149	0.735	0.512
	L10	1.271	2.843	5.290	11.903
	L25	0.447	0.888	2.481	6.331
m=10	L50	0.156	0.185	0.862	2.597
n=10	L75	0.397	0.734	0.252	0.699
	L90	1.170	2.536	0.832	0.638
	L10	0.880	1.908	2.557	11.226
	L25	0.276	0.539	1.258	6.193
m=10	L50	0.073	0.080	0.460	2.697
n=30	L75	0.271	0.530	0.162	0.738
	L90	0.870	1.889	0.365	0.316
	L10	0.654	0.960	1.054	6.034
	L25	0.220	0.283	0.480	3.152
m=20	L50	0.071	0.066	0.168	1.243
n=20	L75	0.208	0.309	0.119	0.305
	L90	0.630	1.013	0.331	0.339
	L10	0.406	0.490	0.599	0.491
	L25	0.145	0.180	0.291	0.274
m=40	L50	0.023	0.053	0.131	0.164
n=40	L75	0.152	0.148	0.105	0.154
	L90	0.425	0.450	0.226	0.247

TABLE 3

SIMULATION RESULTS

ESTIMATED VARIANCES FOR SEQUENTIAL METHOD

INITIAL VALUES

TSS	ROOTS	(L.2,L.8)	(L.35,L.65)	(L.75,L.9)	(L.9,L.95)
	L10	1.698	1.528	2.201	1.843
	L25	0.622	0.538	1.067	0.969
m=5	L50	0.253	0.211	0.464	0.489
n=5	L75	0.592	0.548	0.393	0.400
	L90	1.639	1.549	0.854	0.705
	L10	0.891	0.968	1.331	1.197
	L25	0.303	0.327	0.657	0.653
m=5	L50	0.121	0.129	0.290	0.345
n=15	L75	0.344	0.375	0.231	0.274
	L90	0.973	1.063	0.478	0.440
	L10	0.949	0.943	1.437	1.279
	L25	0.344	0.337	0.701	0.702
m=10	L50	0.131	0.133	0.295	0.364
n=10	L75	0.308	0.330	0.221	0.266
	L90	0.876	0.927	0.477	0.408
	L10	0.512	0.540	0.636	0.578
	L25	0.185	0.186	0.309	0.319
m=10	L50	0.076	0.068	0.135	0.177
n=30	L75	0.184	0.184	0.113	0.153
	L90	0.511	0.536	0.244	0.247
	L10	0.455	0.514	0.682	0.531
	L25	0.161	0.177	0.332	0.295
m=20	L50	0.065	0.060	0.140	0.170
n=20	L75	0.166	0.164	0.107	0.155
	L90	0.465	0.488	0.233	0.251
	L10	0.404	0.489	0.593	0.486
	L25	0.144	0.178	0.288	0.273
m=40	L50	0.023	0.053	0.130	0.163
n=40	L75	0.150	0.146	0.101	0.151
	L90	0.423	0.444	0.222	0.243

TABLE 4

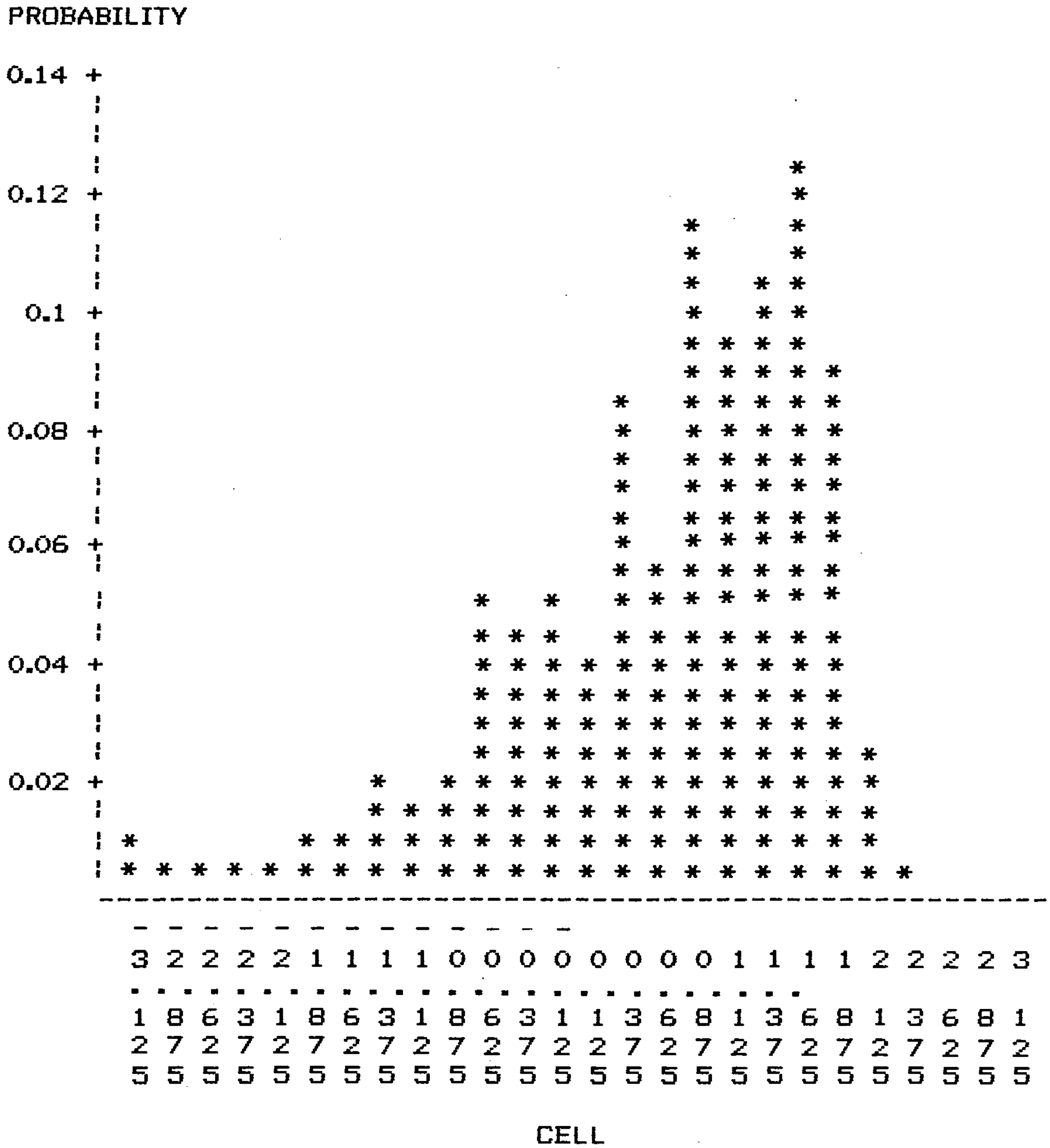
SIMULATION RESULTS

COMPARISON OF ESTIMATED BIASES

INITIAL VALUES

TSS	ROOTS	(L.2,L.8)	(L.35,L.65)	(L.75,L.9)	(L.9,L.95)
	L10	.598 (.551)*	.522 (.657)	.962 (.661)	.361 (.556)
	L25	.299 (.270)	.258 (.332)	.626 (.493)	.224 (.504)
m=5	L50	.000(-.010)	-.006 (.008)	.291 (.324)	.086 (.453)
n=5	L75	-.230(-.291)	-.271 (.317)	-.045 (.156)	-.051 (.401)
	L90	-.599(-.572)	-.535(-.642)	-.380(-.012)	-.189 (.350)
	L10	.433 (.448)	.421 (.477)	.639 (.562)	.260 (.444)
	L25	.219 (.235)	.214 (.241)	.410 (.404)	.155 (.384)
m=5	L50	.004 (.023)	.007 (.006)	.180 (.247)	.050 (.325)
n=15	L75	-.210 (.189)	-.199(-.229)	-.050 (.089)	-.055 (.265)
	L90	-.424(-.401)	-.406(-.464)	-.279(-.068)	-.160 (.205
	L10	.304 (.421)	.464 (.533)	.529 (.527)	.308 (.454)
	L25	.145 (.207)	.231 (.268)	.332 (.389)	.193 (.391)
m=10	L50	-.014(-.007)	-.003 (.003)	.135 (.250)	.078 (.327)
n=10	L75	-.172(-.221)	-.236(-.262)	-.062 (.111)	-.037 (.264)
	L90	-.331(-.435)	-.469(-.528)	-.260(-.028)	-.152 (.201)
	L10	.258 (.338)	.188 (.347)	.188 (.337)	.184 (.270)
	L25	.129 (.170)	.086 (.169)	.110 (.247)	.115 (.233)
m=10	L50	.001 (.003)	-.015(-.008)	.032 (.157)	.046 (.197)
n=30	L75	-.128(-.165)	-.117(-.185)	-.046 (.067)	-.023 (.160)
	L90	-.256(-.333)	-.218(-.363)	-.125(-.024)	-.091 (.123)
	L10	.159 (.315)	.210 (.380)	.085 (.389)	-.107 (.234)
	L25	.079 (.157)	.112 (.187)	.042 (.281)	-.099 (.204)
m=20	L50	.000(-.001)	-.013(-.006)	-.001 (.174)	-.092 (.174)
n=20	L75	-.080(-.159)	-.085(-.199)	-.044 (.066)	-.085 (.143)
	L90	-.160(-.316)	-.184(-.392)	-.087(-.041)	-.078 (.113)
	L10	.098 (.294)	.168 (.317)	.038 (.292)	-.062 (.203)
	L25	.480 (.140)	.066 (.153)	.016 (.069)	-.085 (.179)
m=40	L50	.000 (.000)	-.001(-.005)	.000 (.026)	-.023 (.147)
n=40	L75	-.500(-.153)	-.062(-.172)	-.042 (.024)	-.006 (.128)
	L90	-.100(-.230)	-.155(-.336)	-.061(-.014)	-.067 (.103)

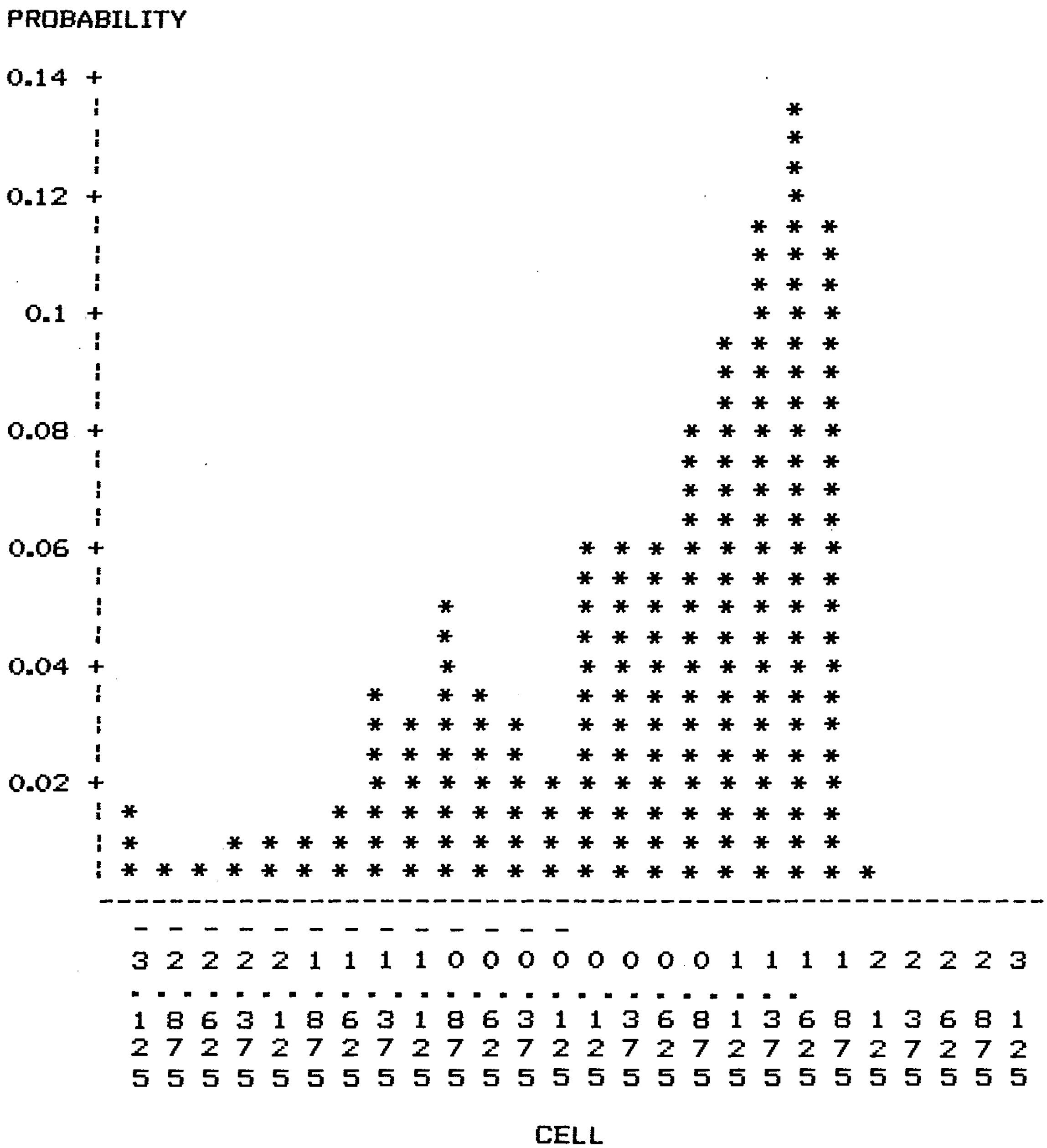
* Biases for the two stage are unbracketed, biases in brackets are for the corresponding sequential estimates.



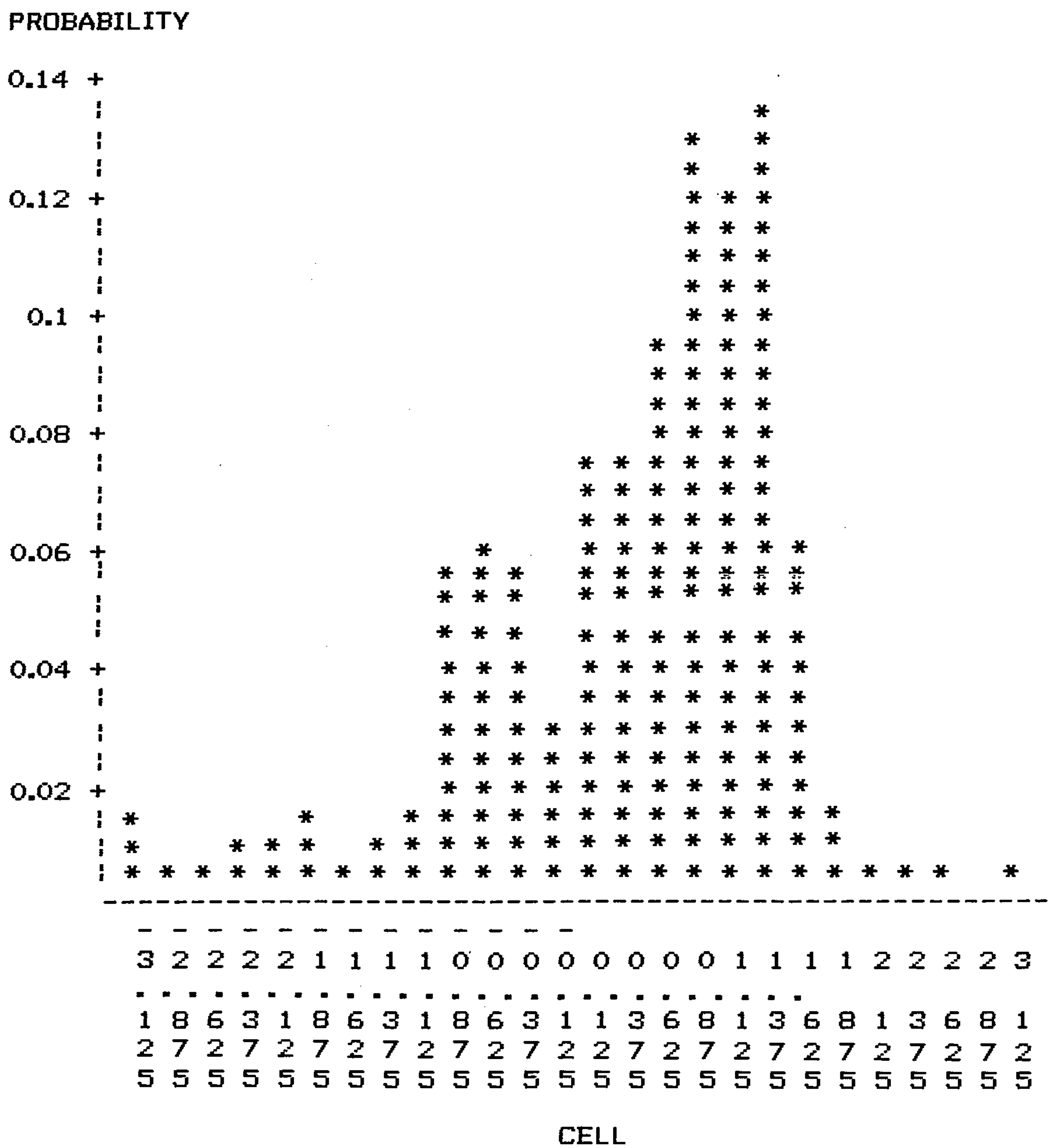
SIMULATION RESULTS FOR TWO STAGE METHOD

(m,n) = (5,5), (X11,X21) = (L.2,L.8)

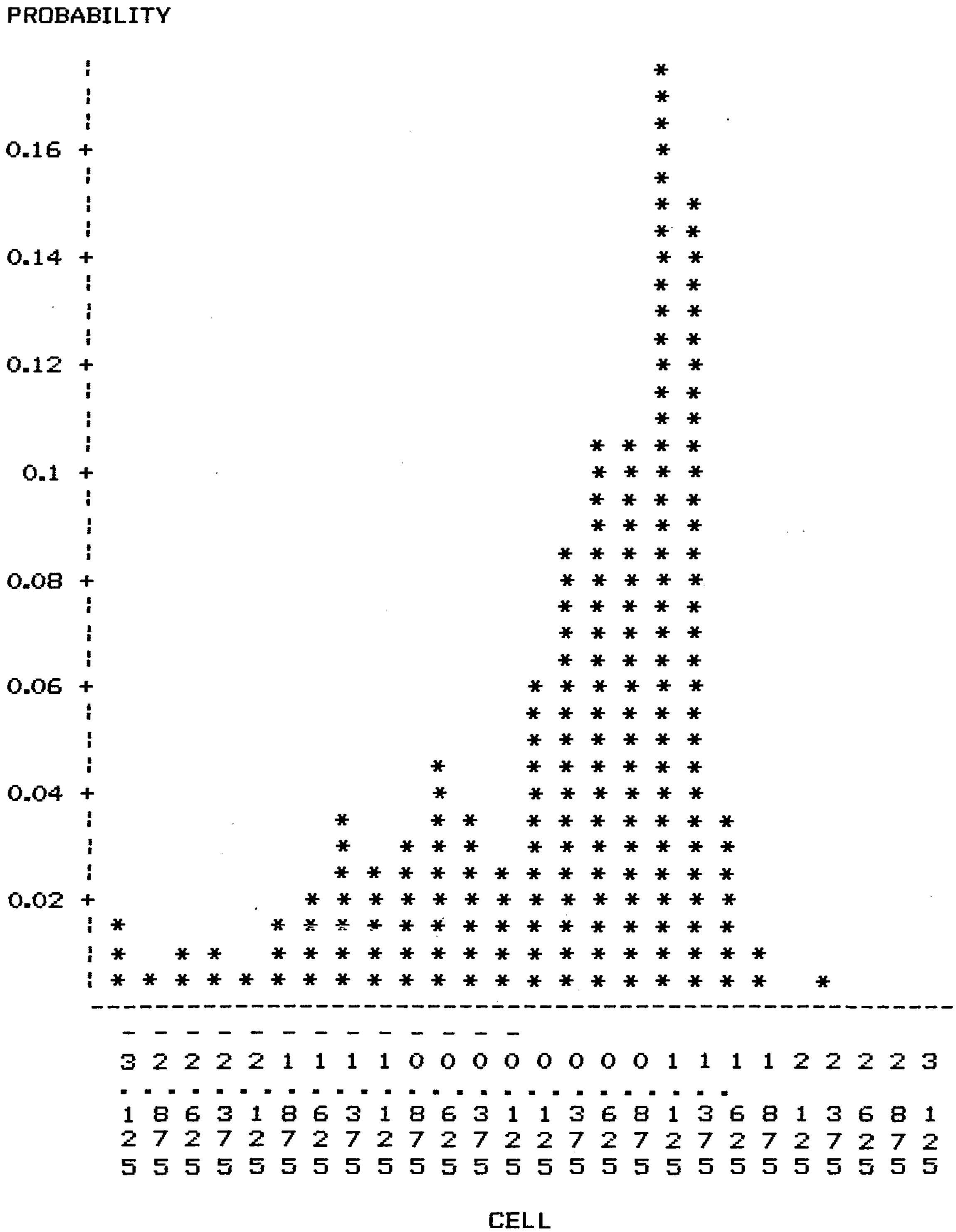
FIGURE 3 FREQUENCY DISTRIBUTION OF L10HAT



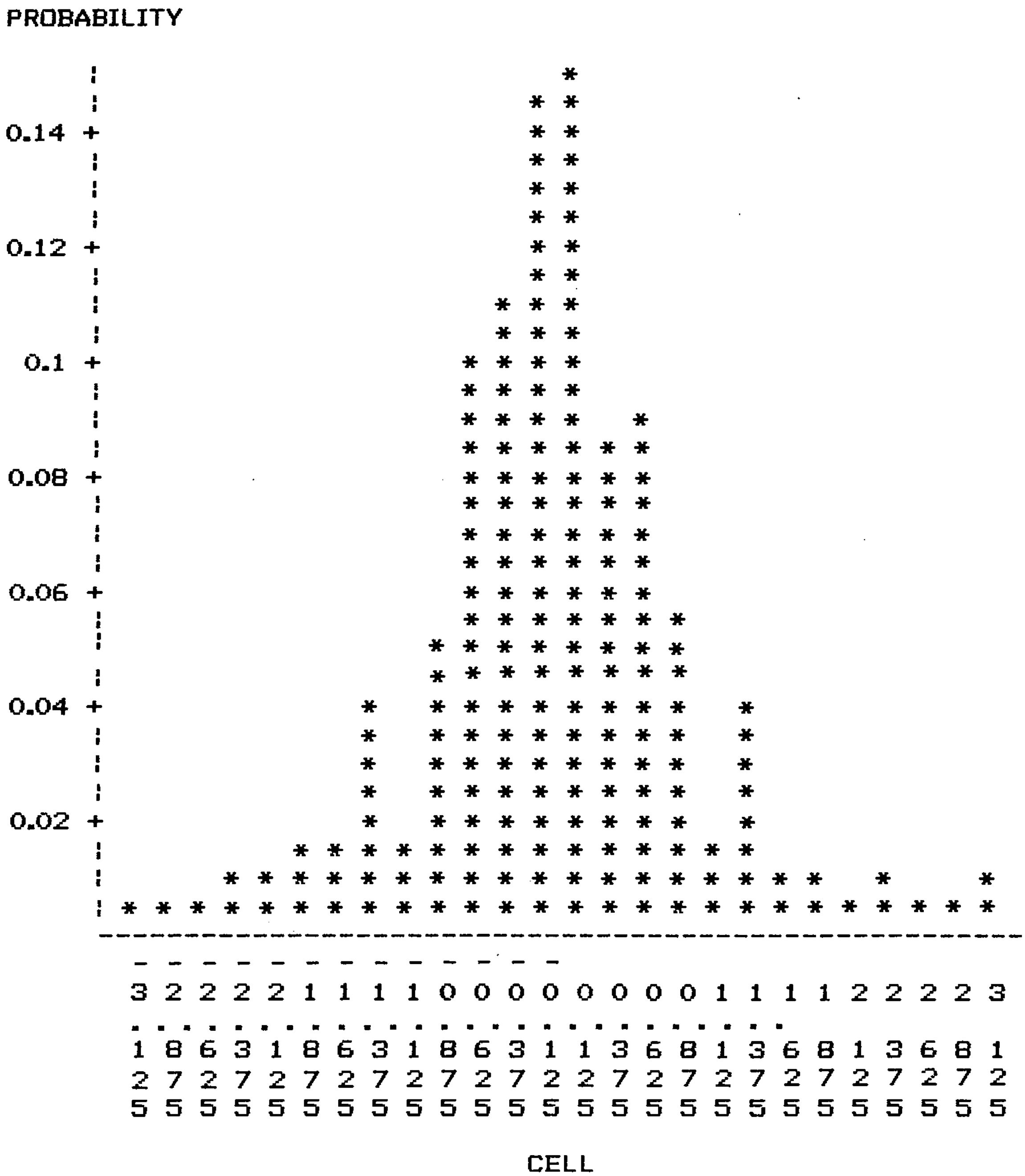
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (5,5), (X11,X21) = (L.2, L.8)
FIGURE 4 FREQUENCY DISTRIBUTION OF L10HAT



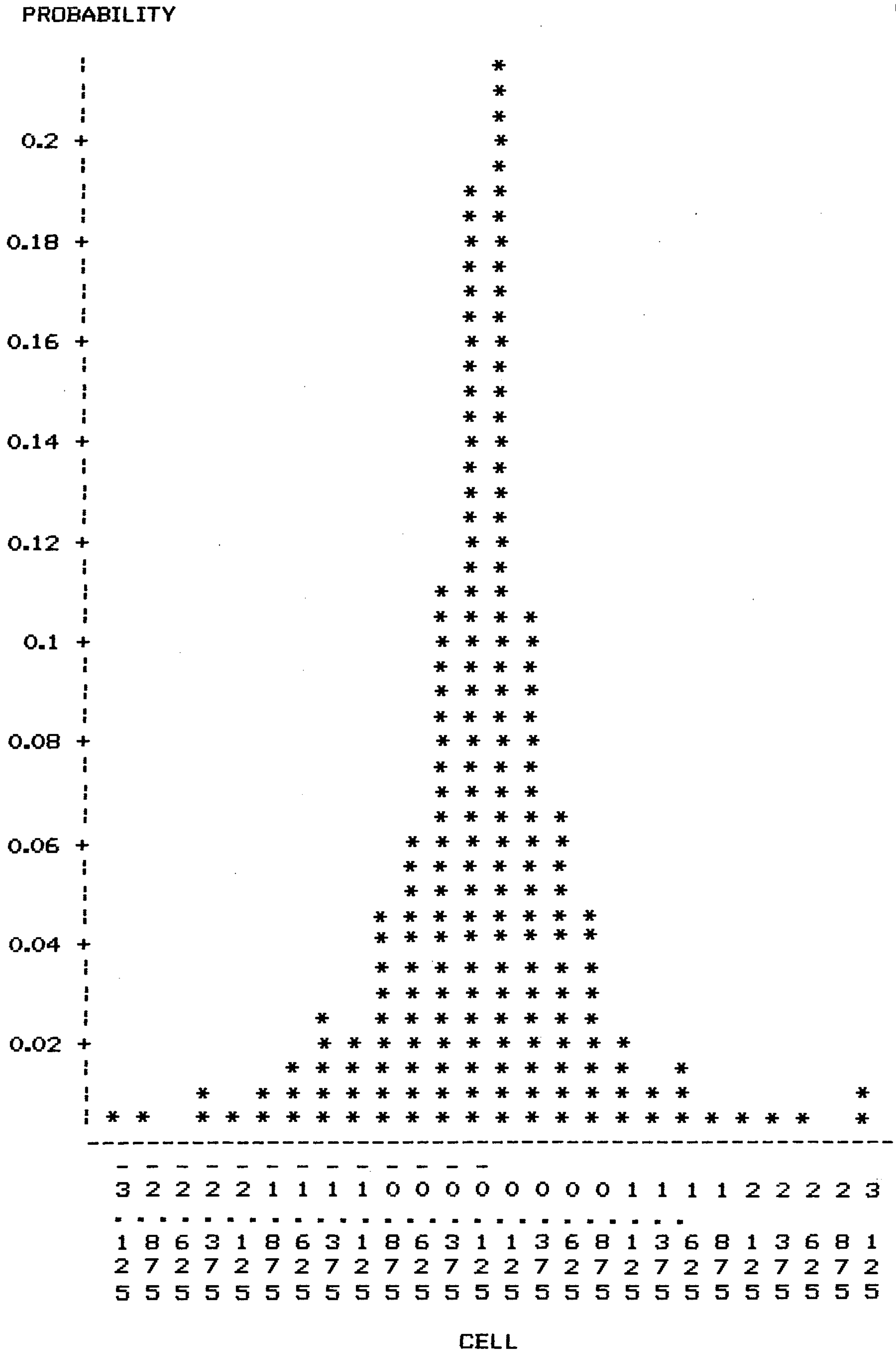
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (5,5), (X11,X21) = (L.2,L.8)
FIGURE 5 FREQUENCY DISTRIBUTION OF L25HAT



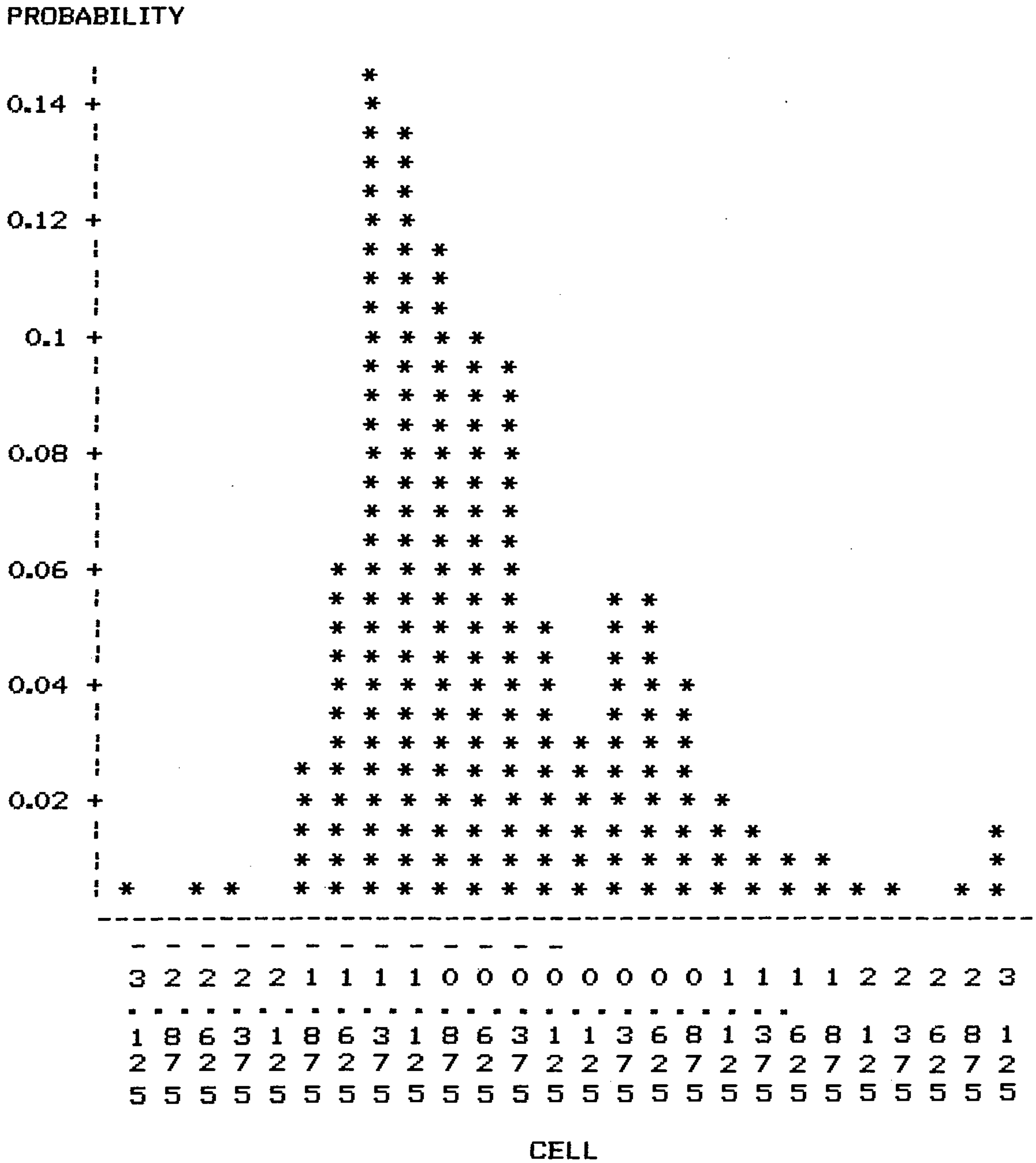
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (5,5), (X11,X21) = (L.2,L.8)
FIGURE 6 FREQUENCY DISTRIBUTION OF L25HAT



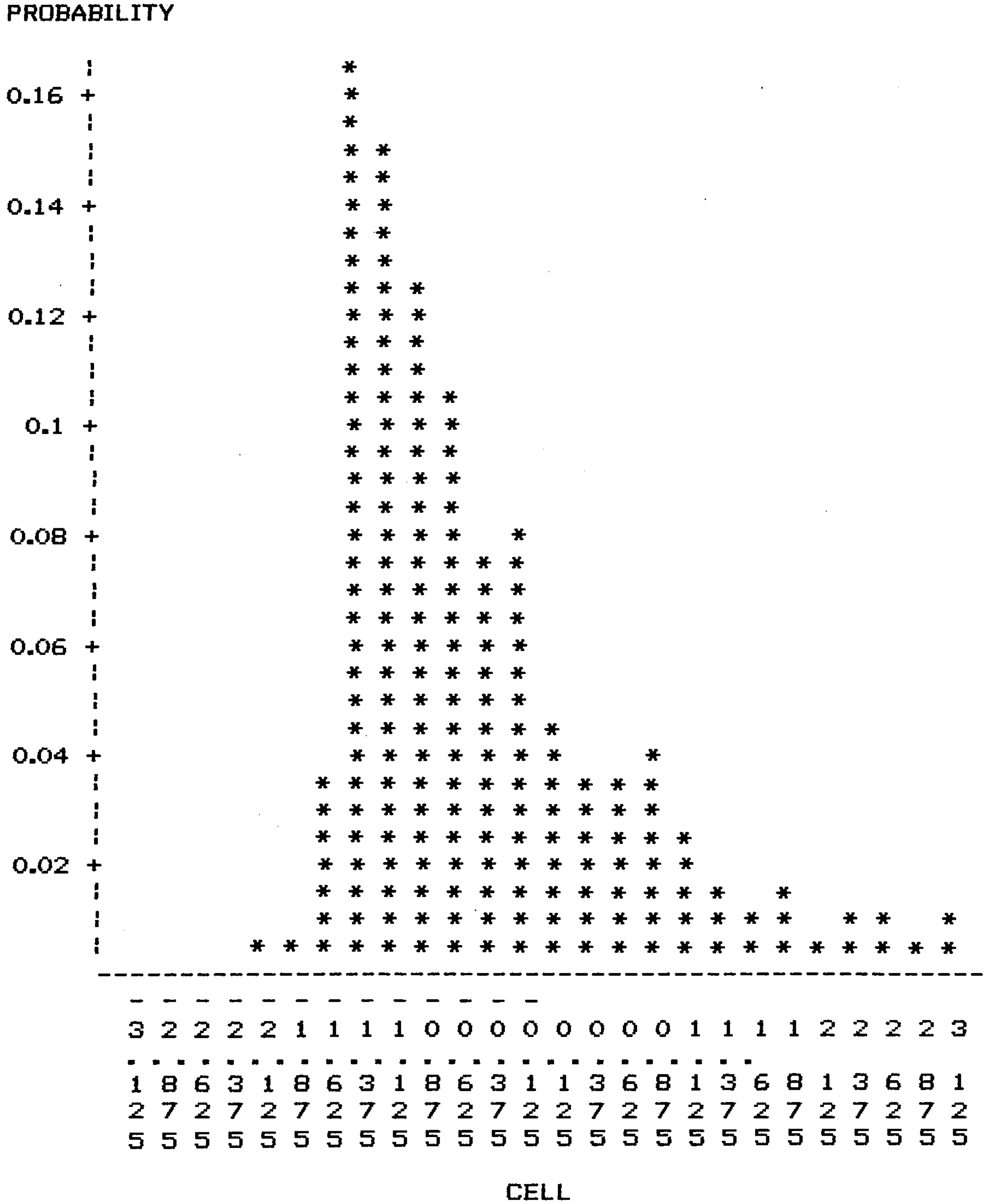
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (5,5), (X11,X21) = (L.2,L.8)
FIGURE 7 FREQUENCY DISTRIBUTION OF L50HAT



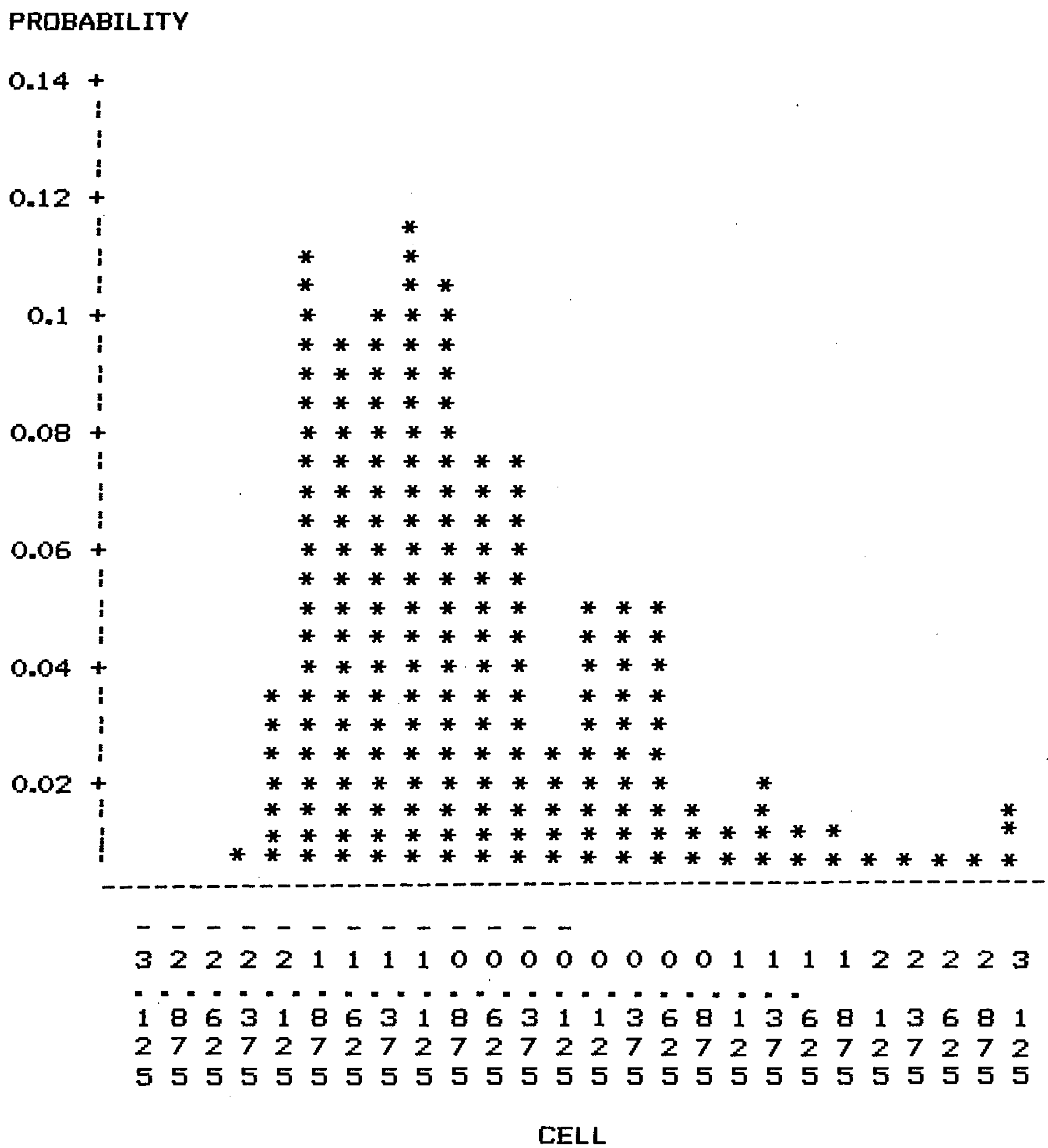
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (5,5), (X11,X21) = (L.2,L.8)
FIGURE 8 FREQUENCY DISTRIBUTION OF L50HAT



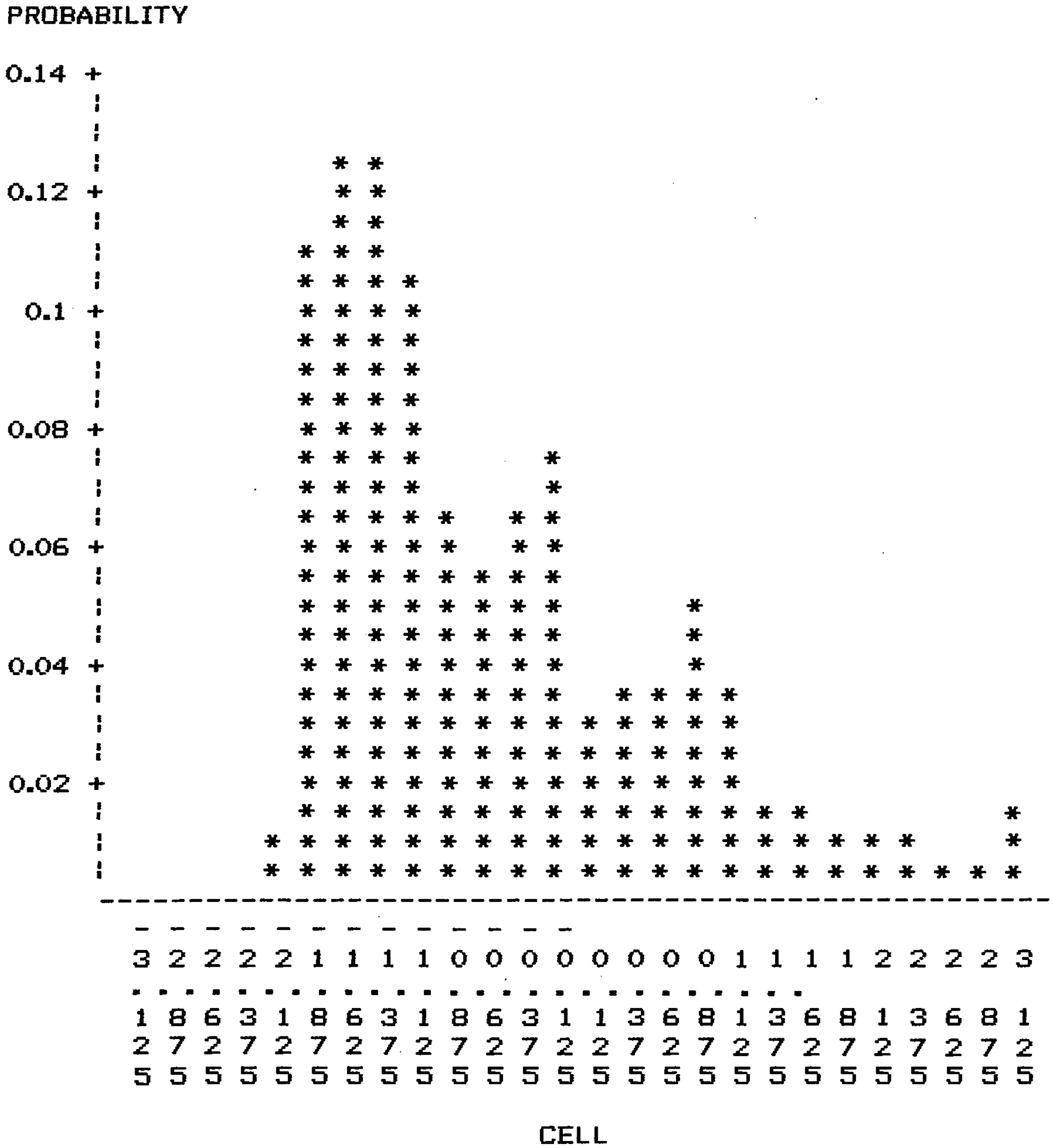
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (5,5), (X11,X21) = (L.2,L.8)
FIGURE 9 FREQUENCY DISTRIBUTION OF L75HAT



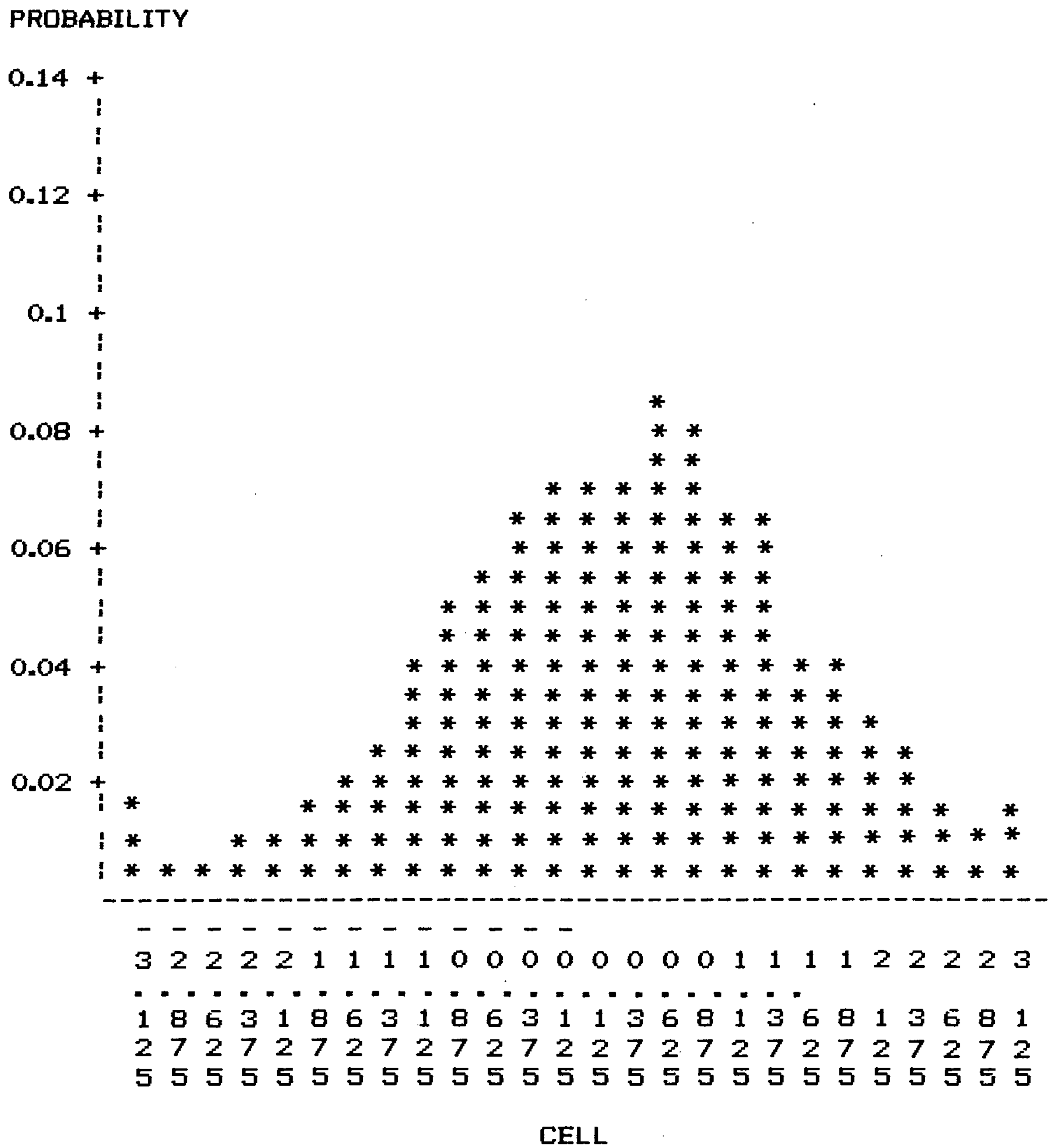
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (5,5), (X11,X21) = (L.2,L.8)
FIGURE 10 FREQUENCY DISTRIBUTION OF L75HAT



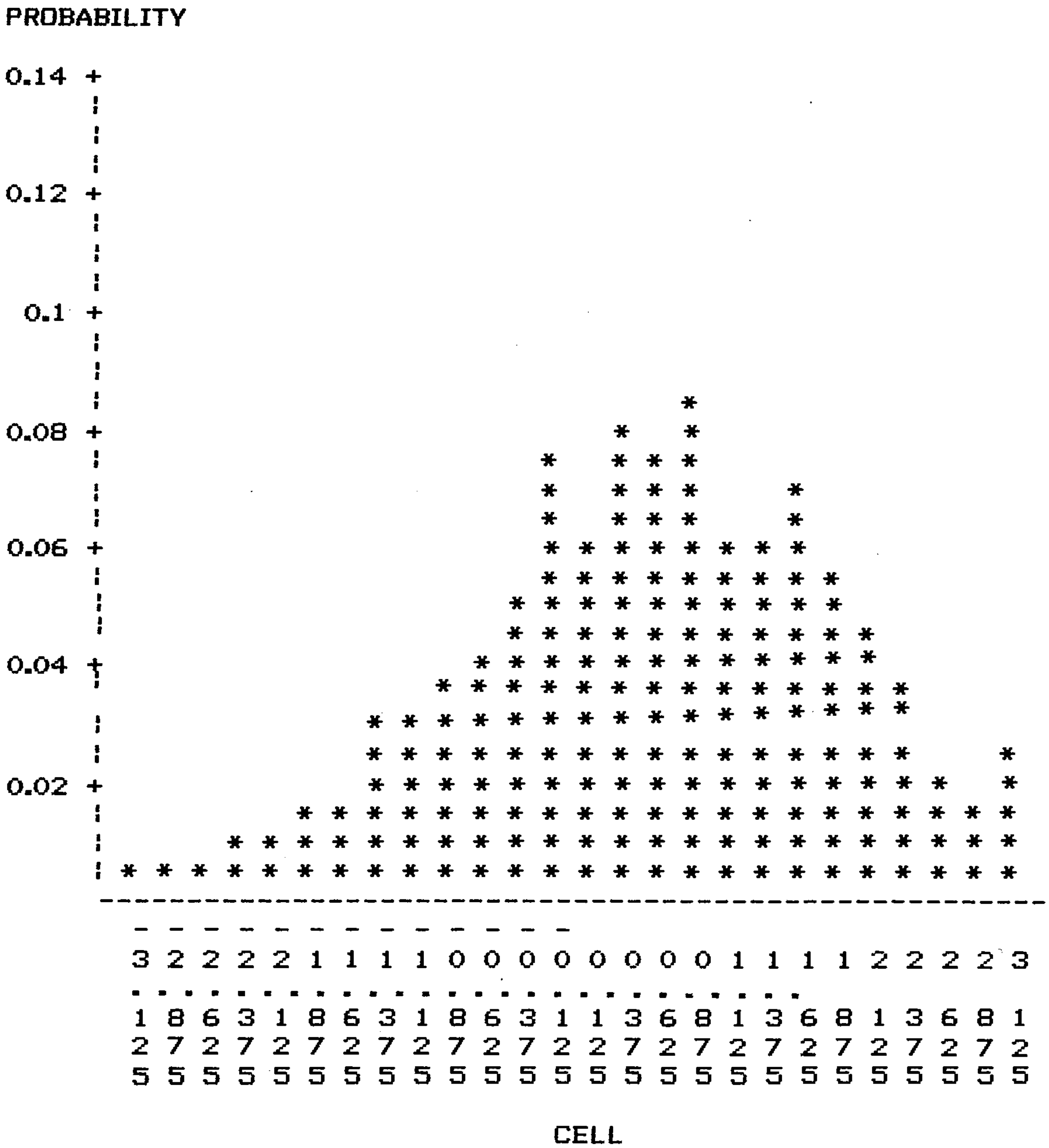
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (5,5), (X11,X21) = (L.2,L.8)
FIGURE 11 FREQUENCY DISTRIBUTION OF L90HAT



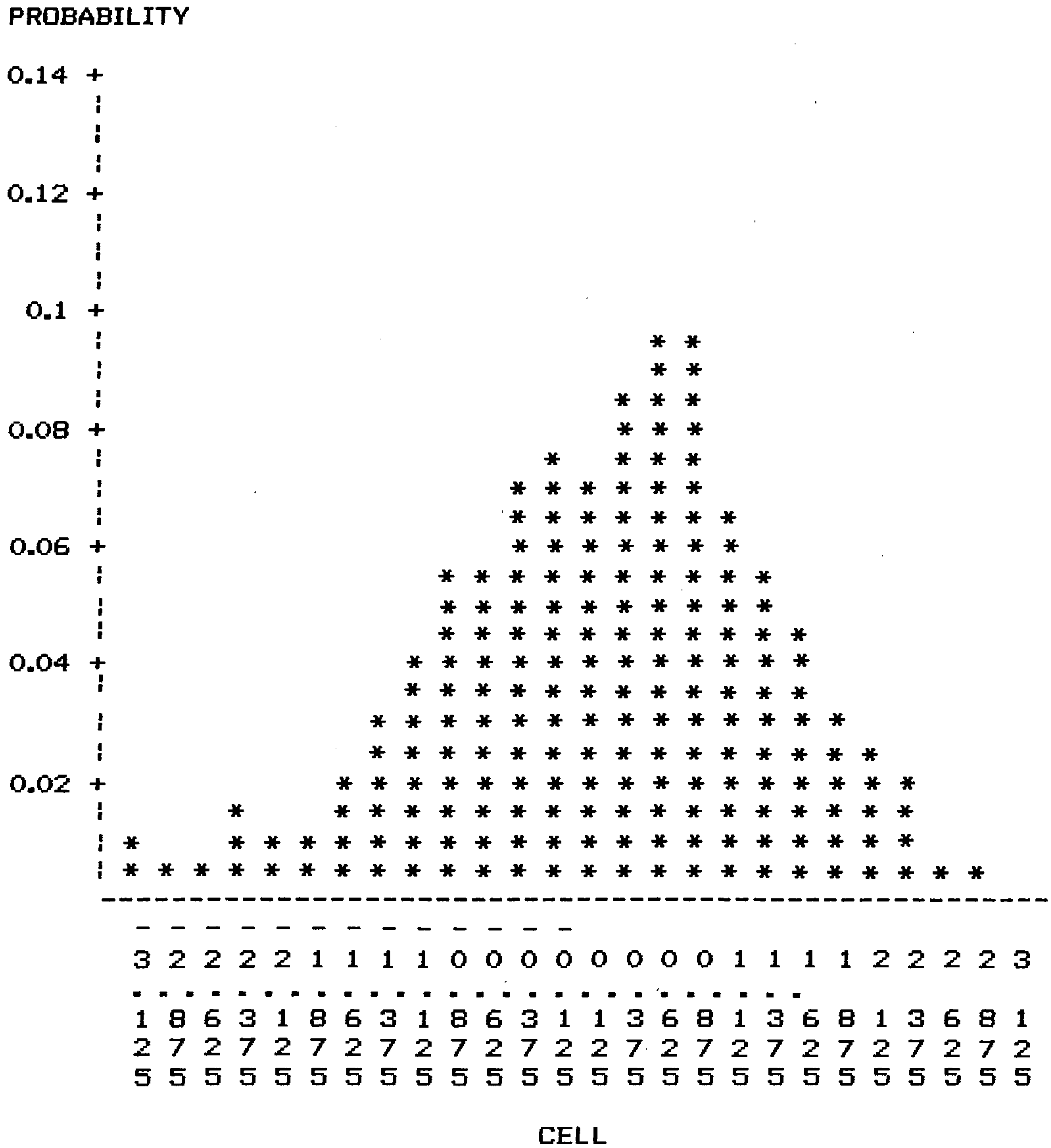
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (5,5), (X11,X21) = (L.2,L.8)
FIGURE 12 FREQUENCY DISTRBUTION OF L90HAT



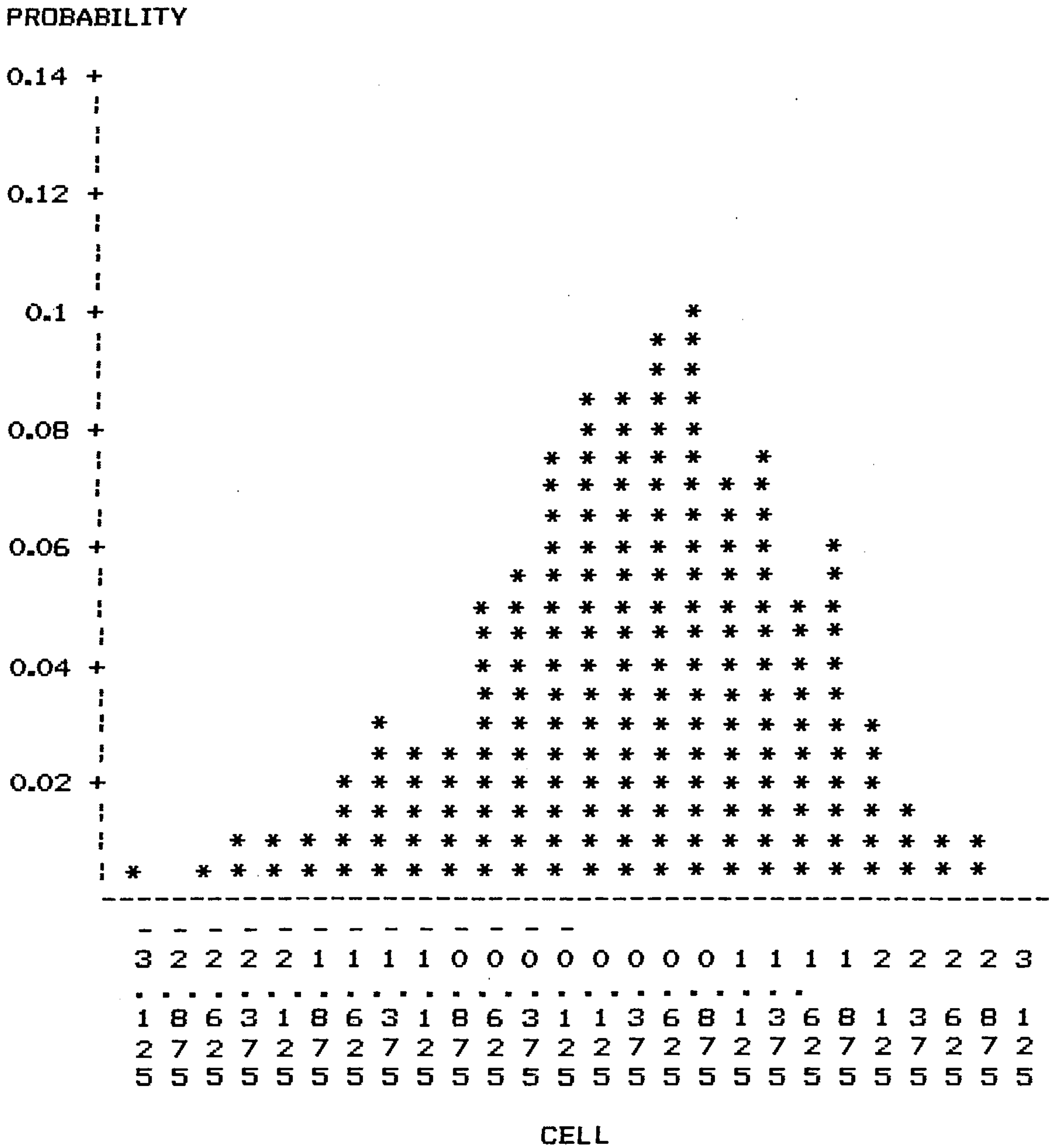
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (40,40), (X11,X21) = (L.2,L.8)
FIGURE 13 FREQUENCY DISTRIBUTION OF L1OHAT



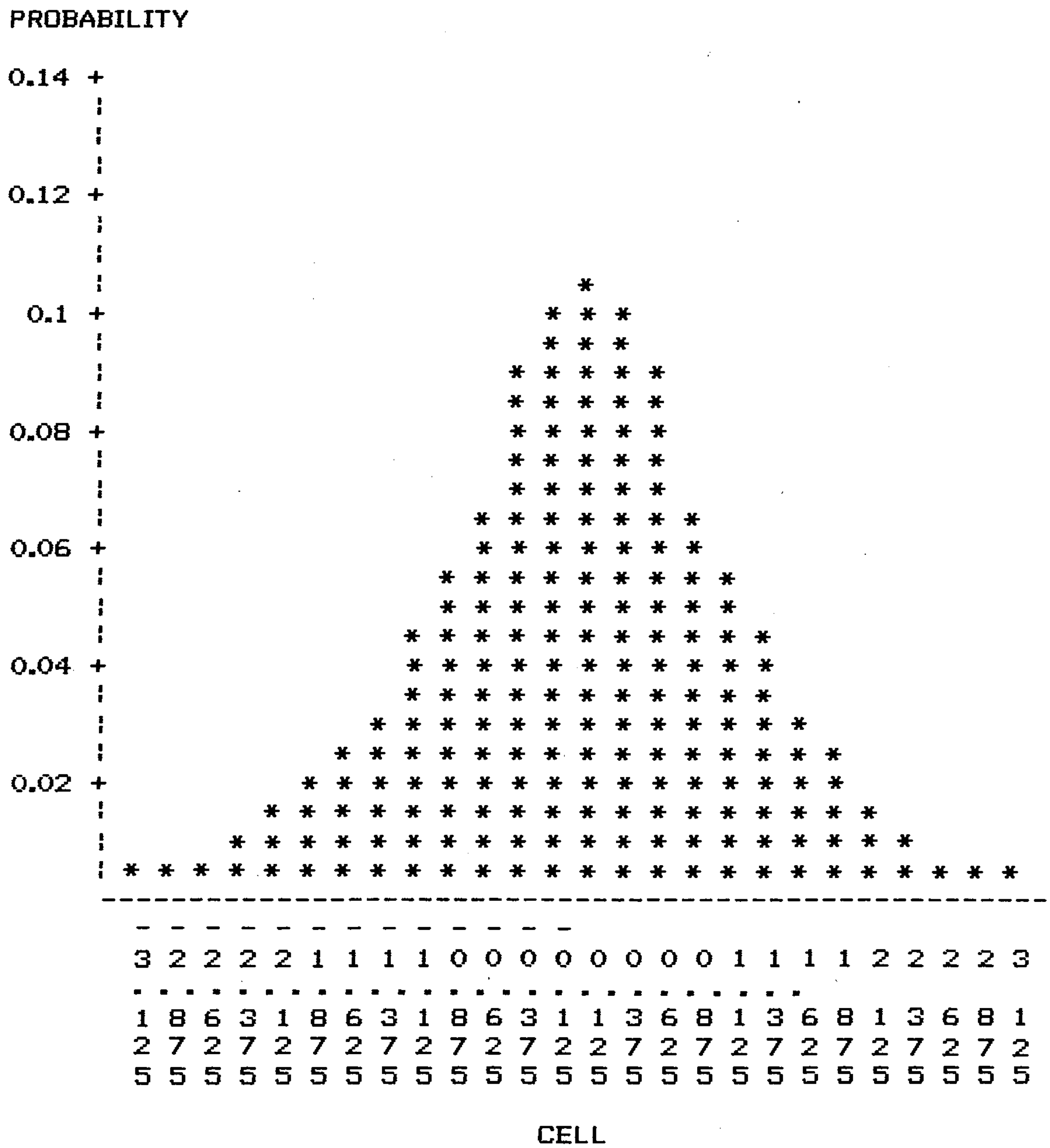
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (40,40), (X11,X21) =(L.2,L.8)
FIGURE 14 FREQUENCY DISTRIBUTION OF L10HAT



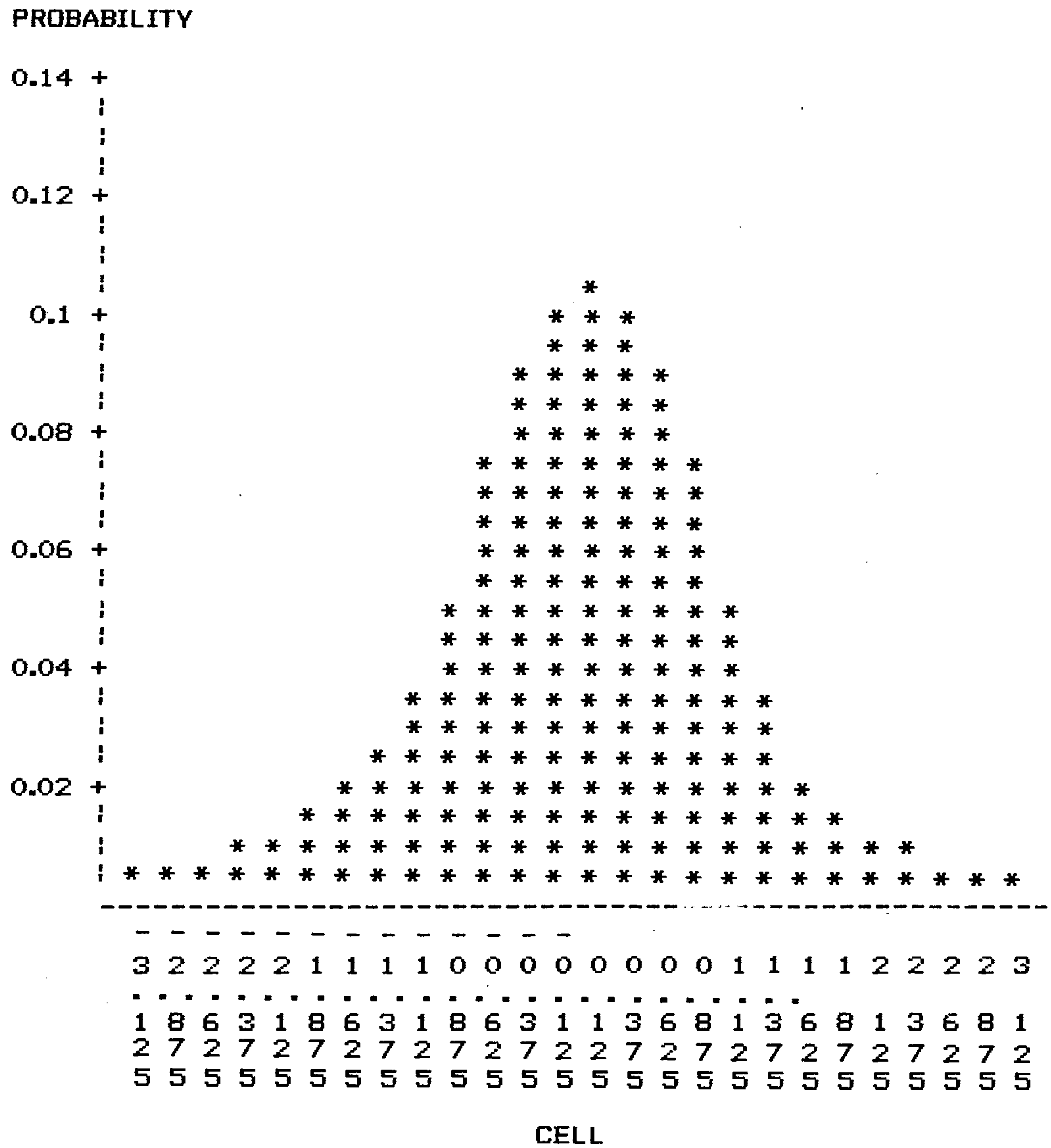
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (40,40), (X11,X21) = (L.2,L.8)
FIGURE 15 FREQUENCY DISTRIBUTION OF L25HAT



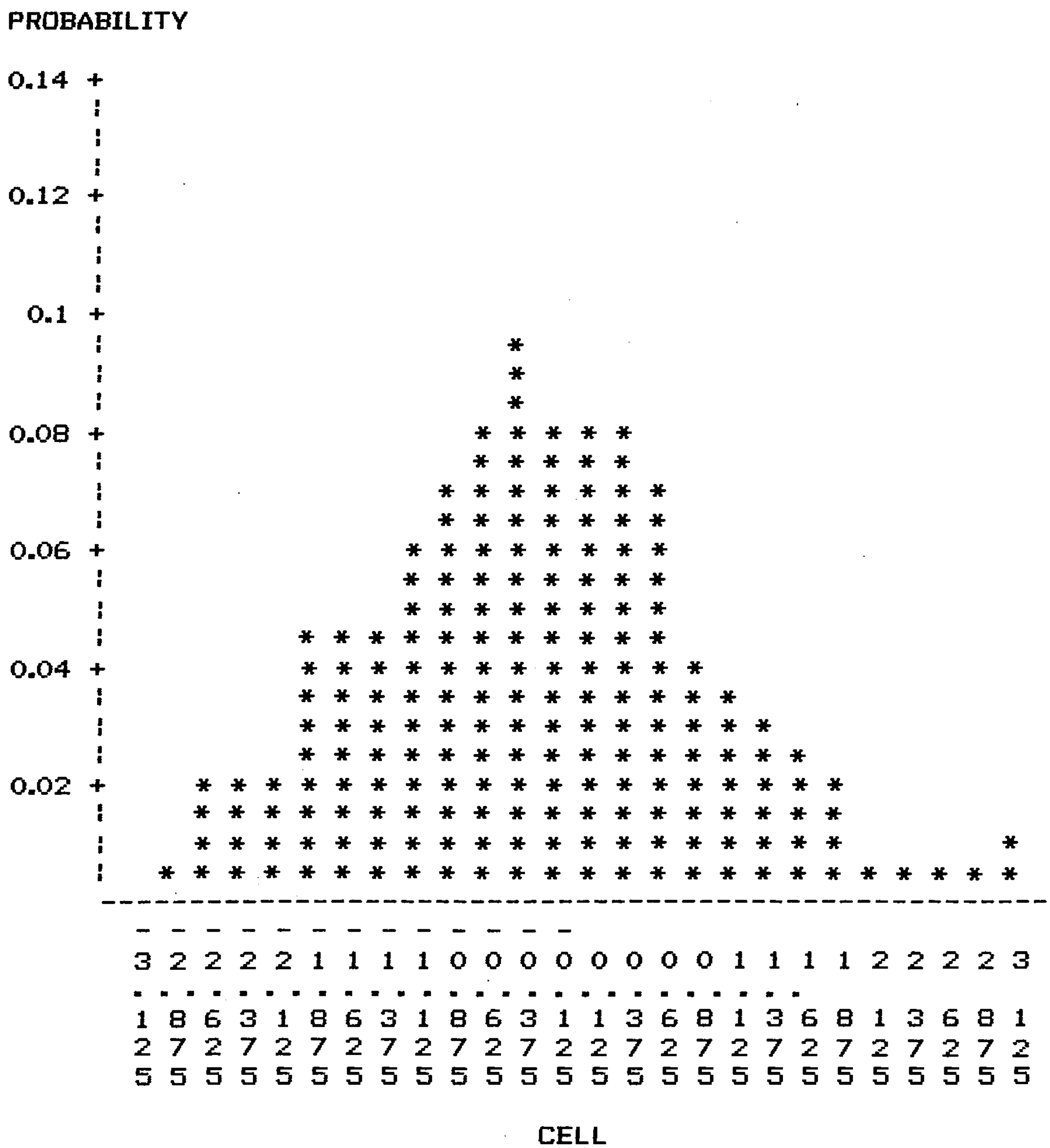
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (40,40), (X11,X21) = (L.2,L.8)
FIGURE 16 FREQUENCY DISTRIBUTION OF L25HAT



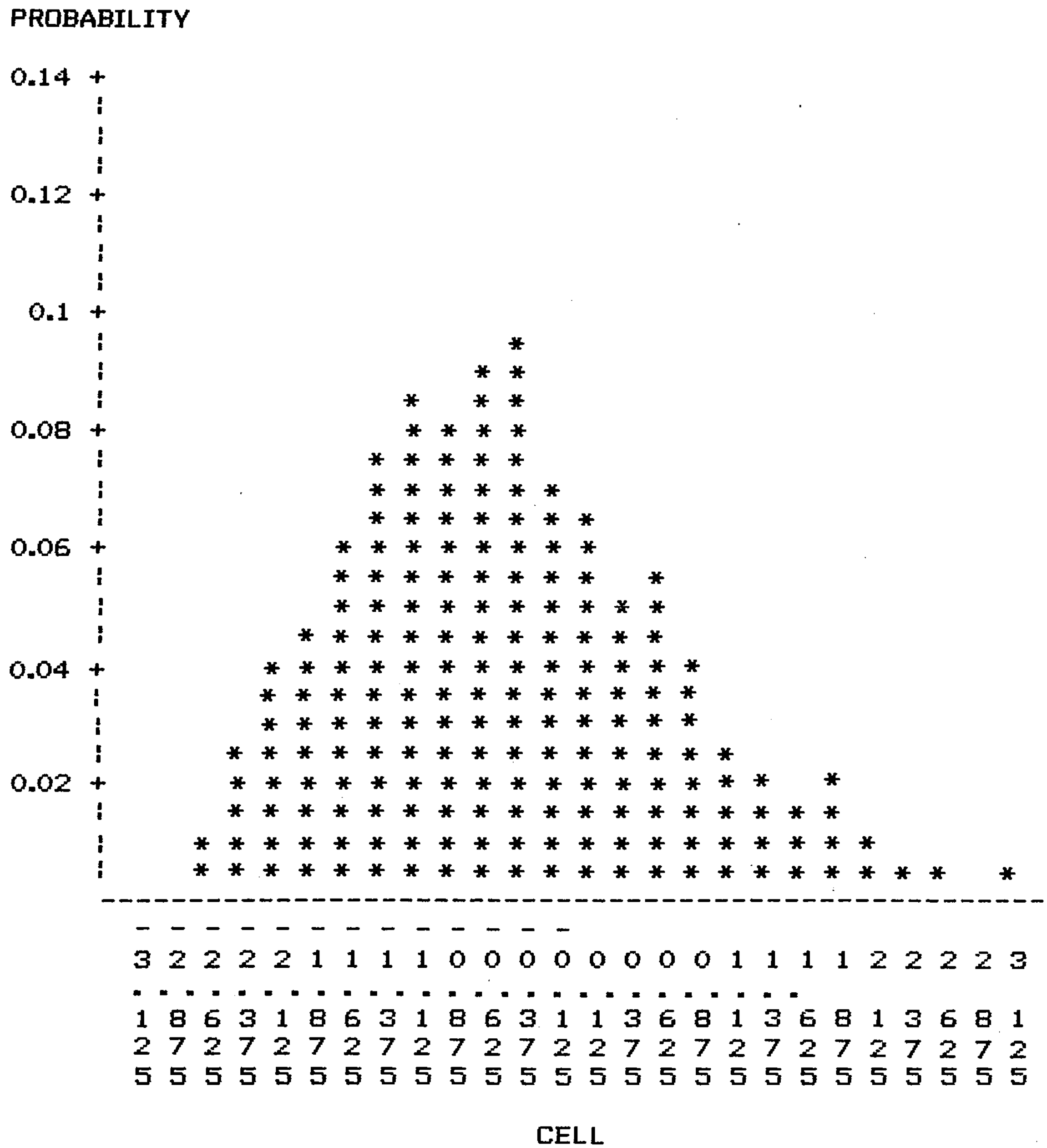
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (40,40), (X11,X21) = (L.2,L.8)
FIGURE 17 FREQUENCY DISTRIBUTION OF L50HAT



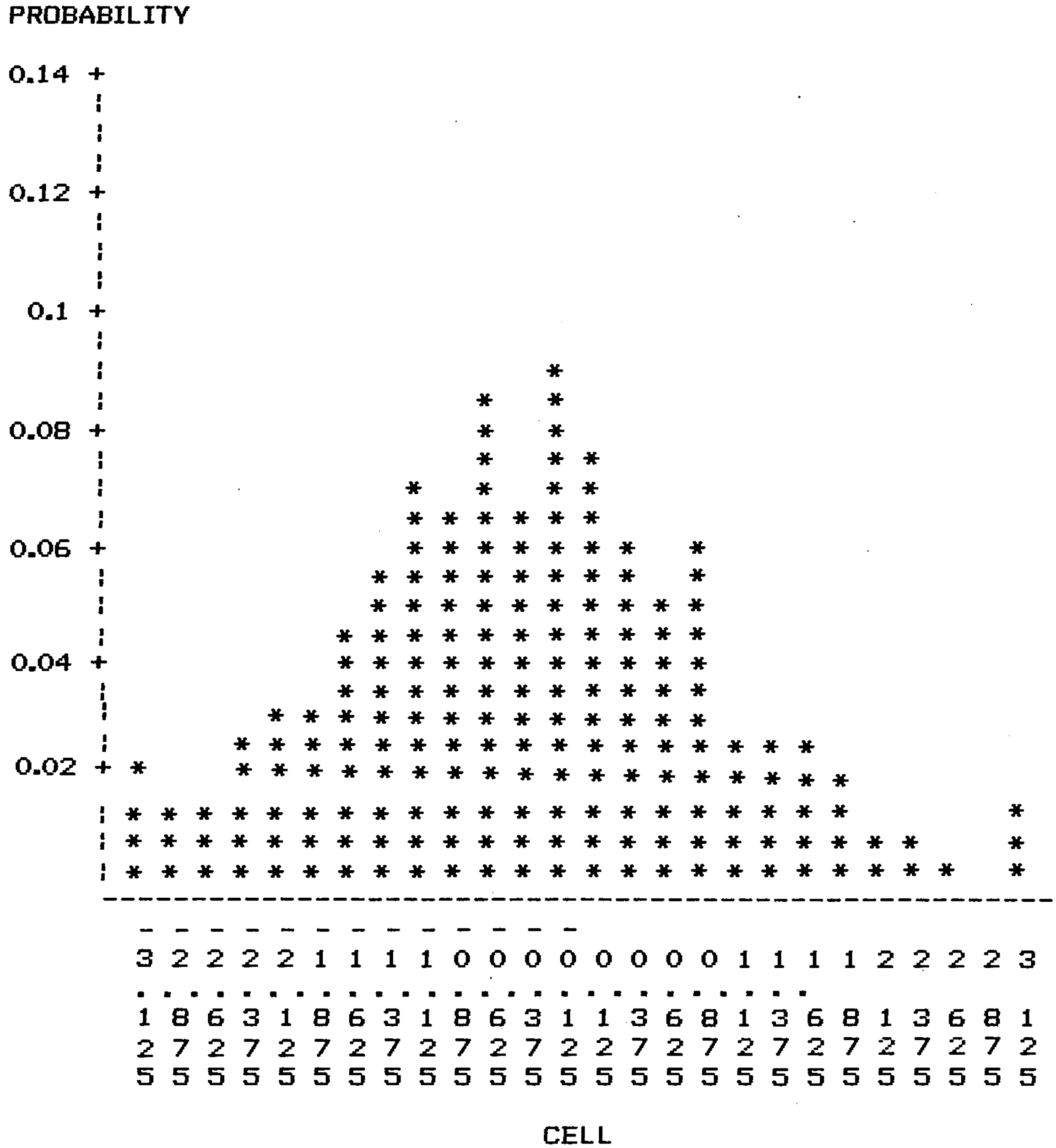
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (40,40), (X11,X21) = (L.2,L.8)
FIGURE 18 FREQUENCY DISTRIBUTION OF L50HAT



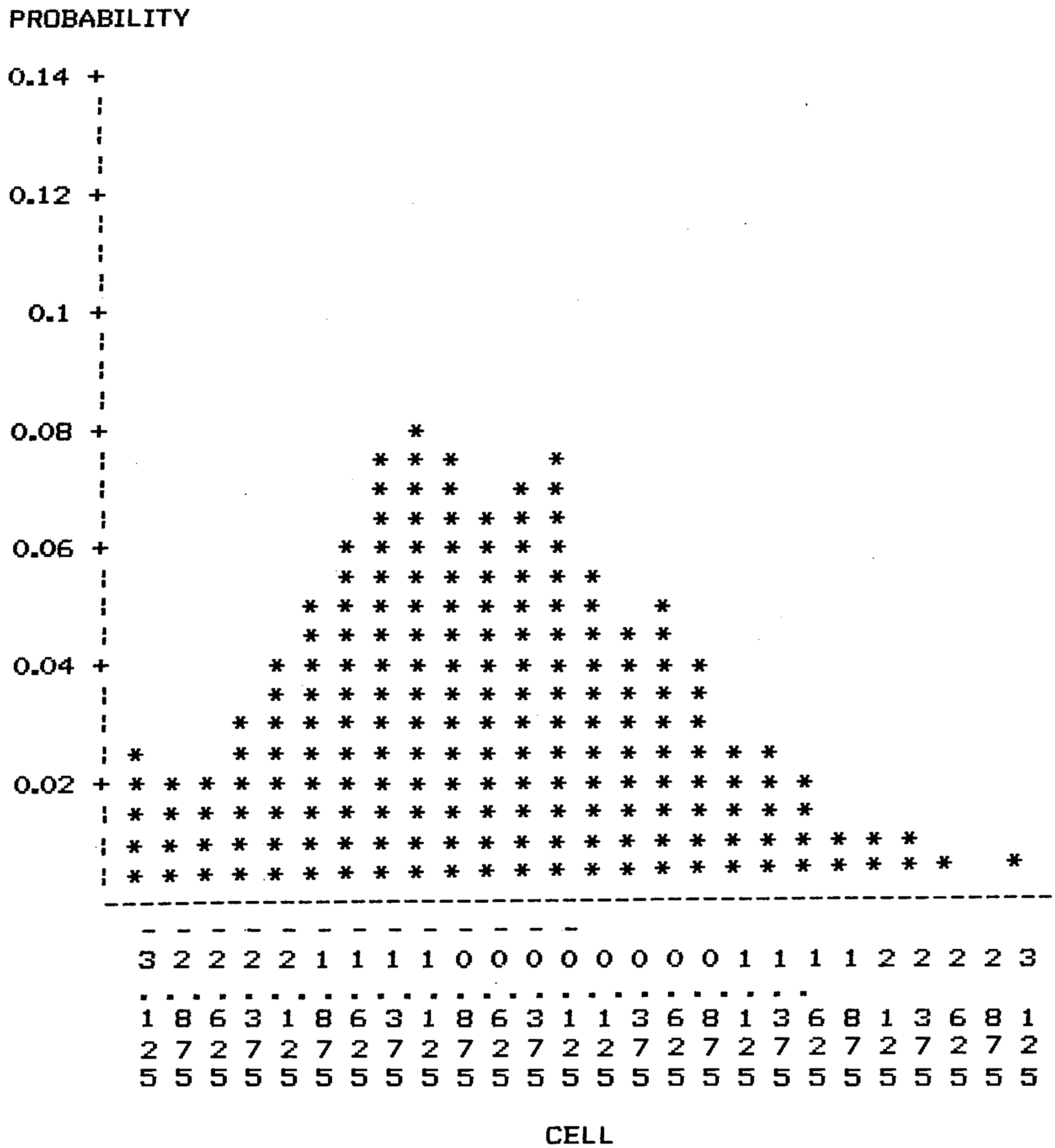
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (40,40), (X11,X21) = (L.2,L.8)
FIGURE 19 FREQUENCY DISTRIBUTION OF L75HAT



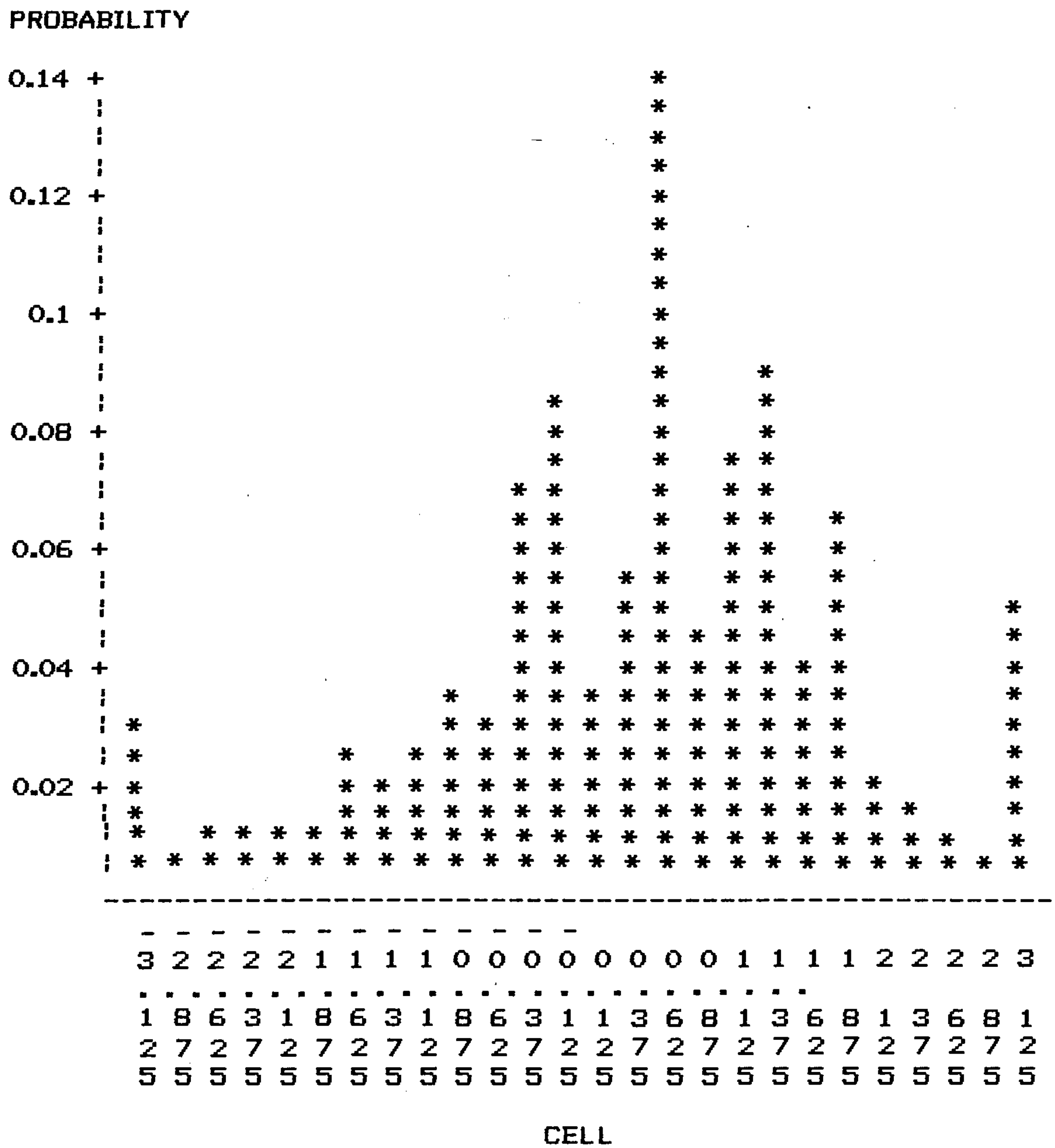
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (40,40), (X11,X21) = (L.2,L.8)
FIGURE 20 FREQUENCY DISTRIBUTION OF L75HAT



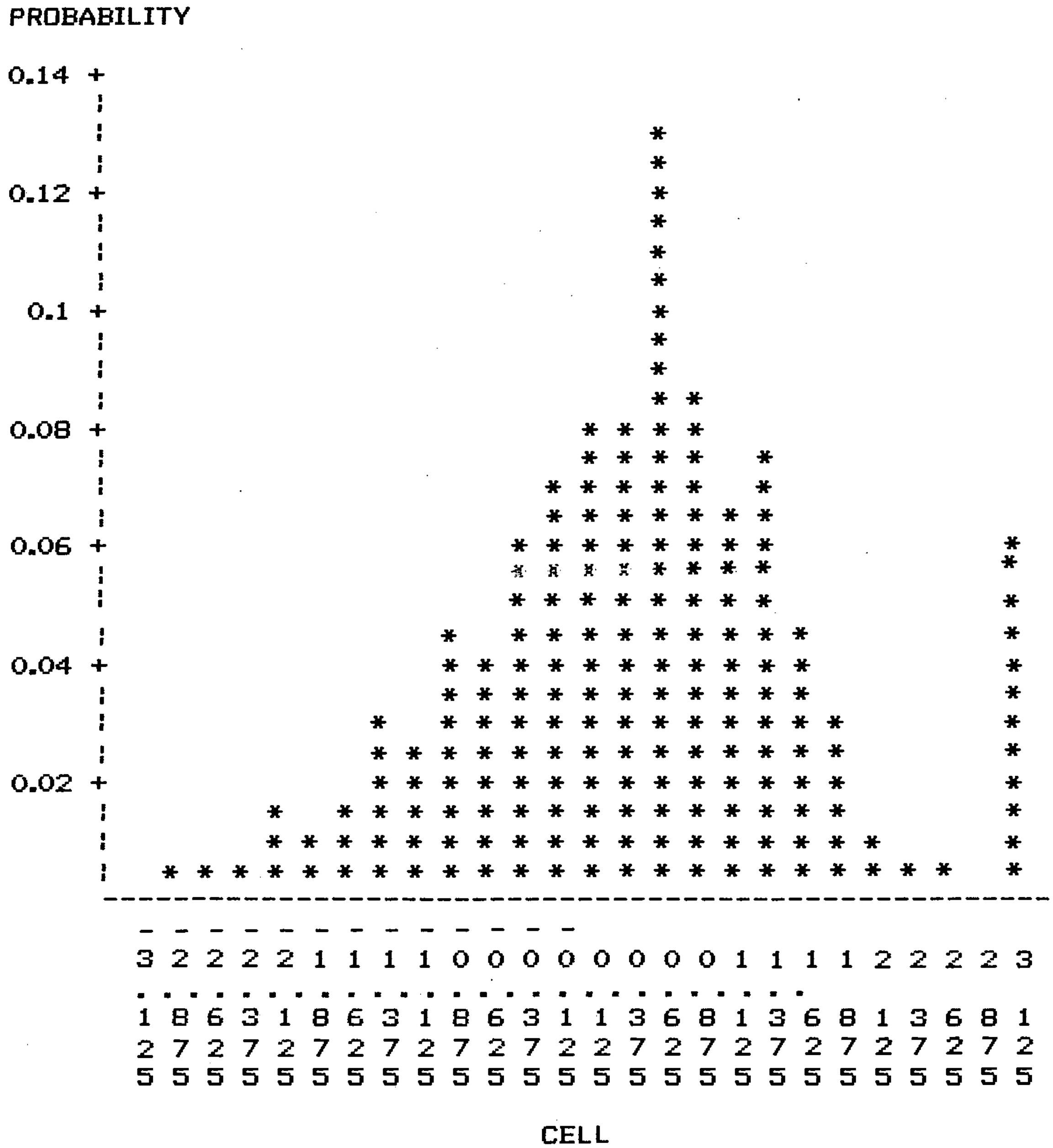
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (40,40), (X11,X21) = (L.2,L.8)
FIGURE 21 FREQUENCY DISTRIBUTION OF L90HAT



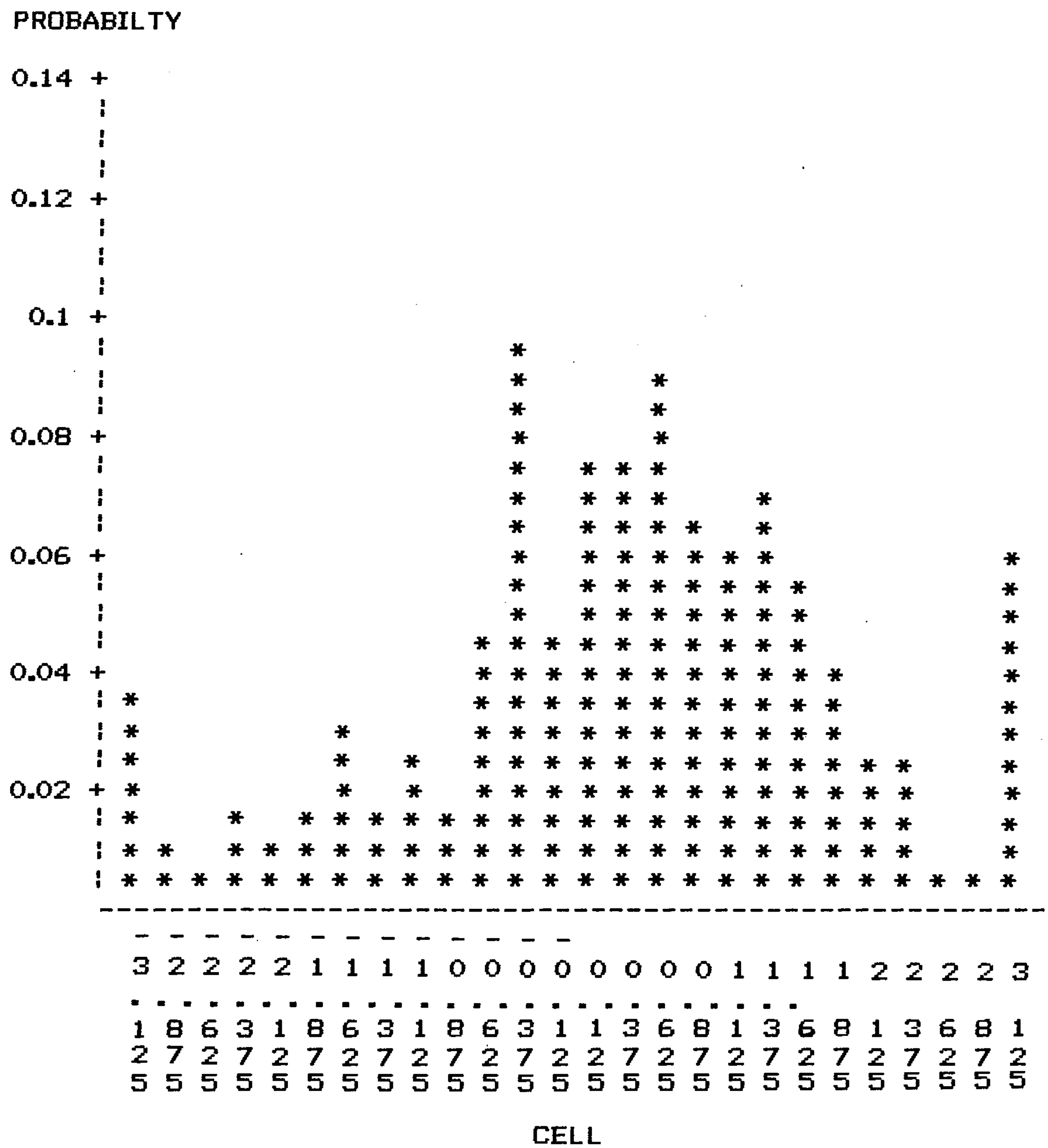
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (40,40), (X11,X21) = (L.2,L.8)
FIGURE 22 FREQUENCY DISTRIBUTION OF L90HAT



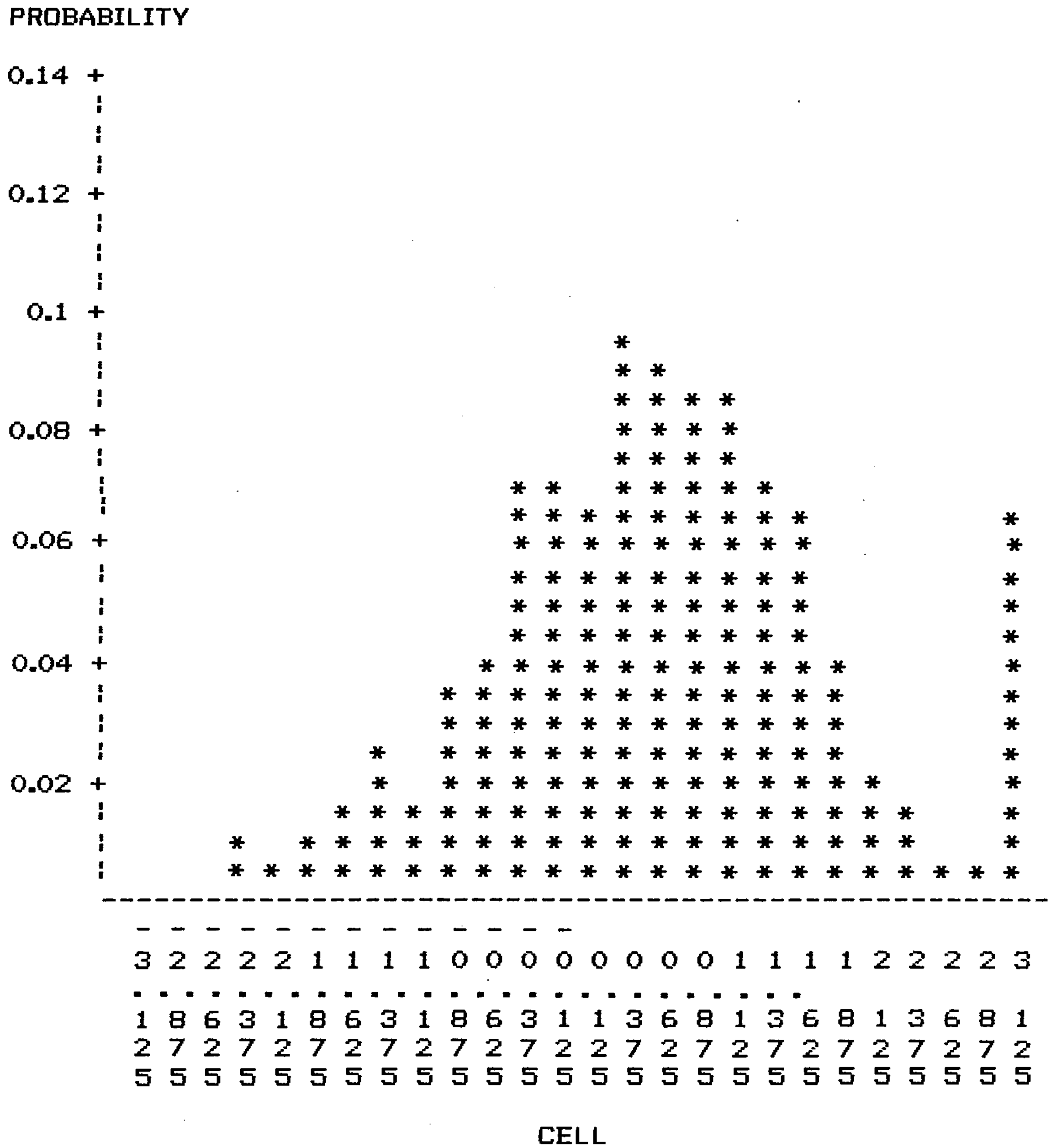
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (5,5), (X11,X21) = (L.9,L.95)
FIGURE 23 FREQUENCY DISTRIBUTION OF L10HAT



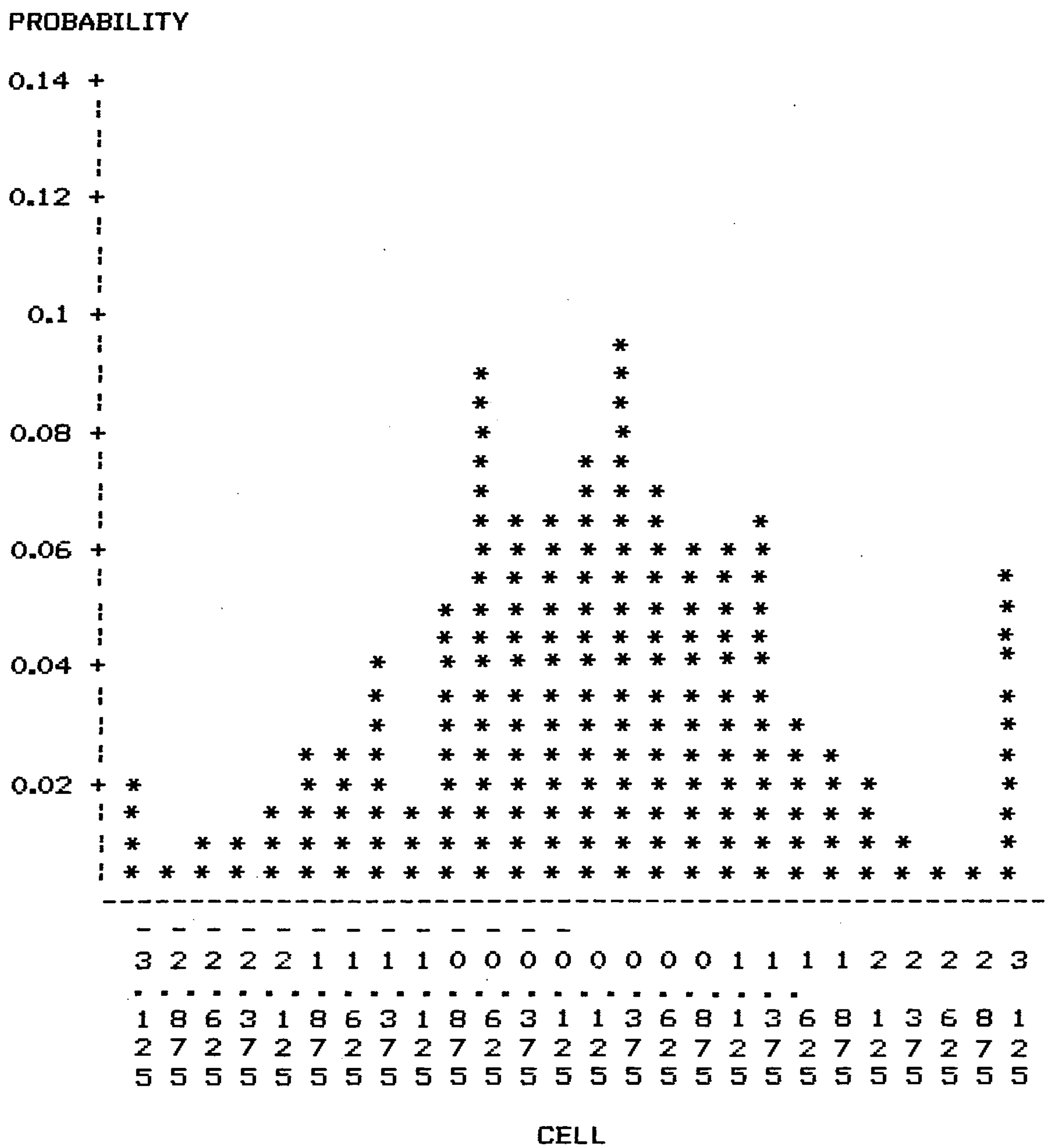
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (5,5), (X11,X21) = (L.9,L.95)
FIGURE 24 FREQUENCY DISTRIBUTION OF L10HAT



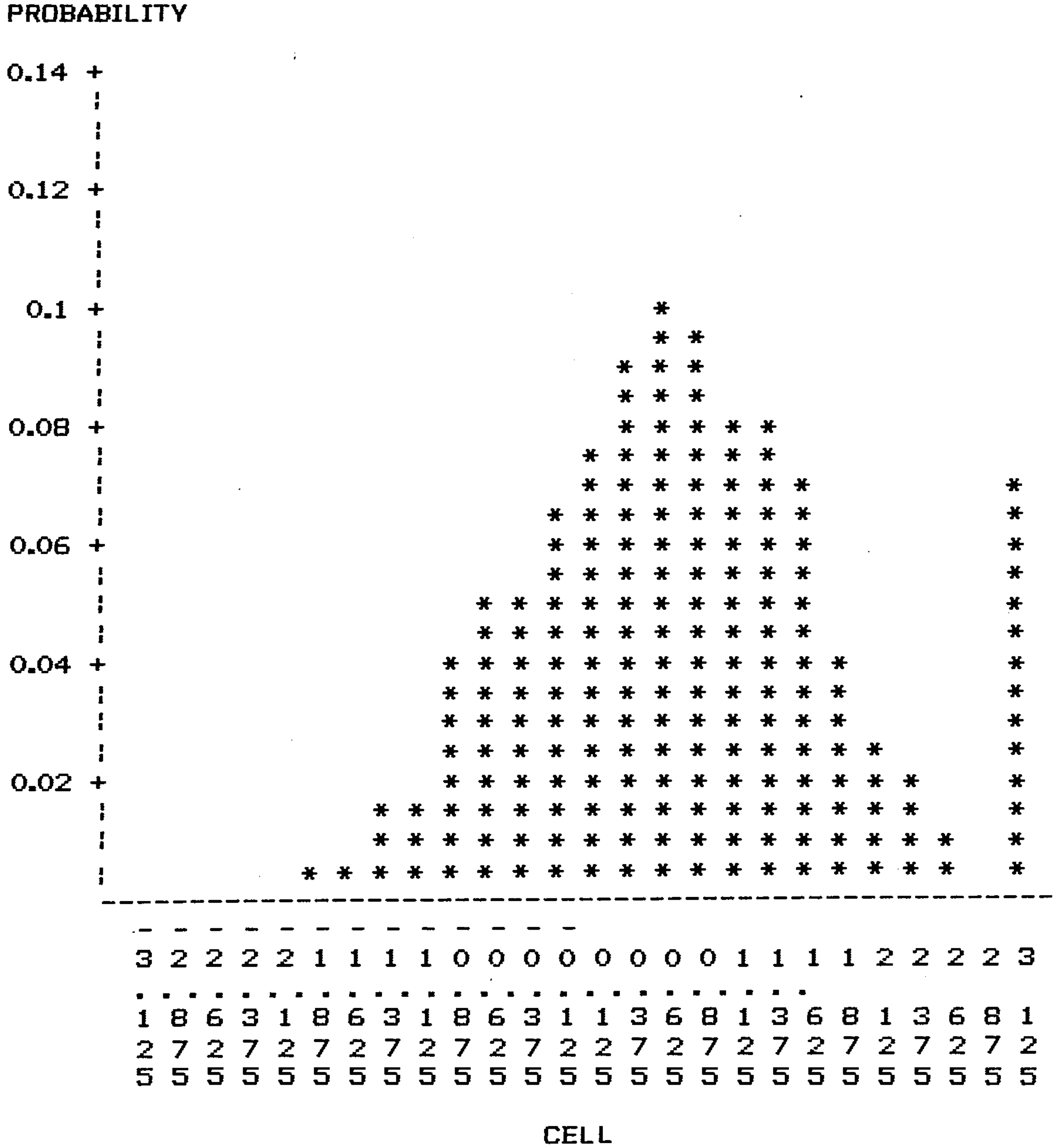
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (5,5), (X11,X21) = (L.9,L.95)
FIGURE 25 FREQUENCY DISTRIBUTION OF L25HAT



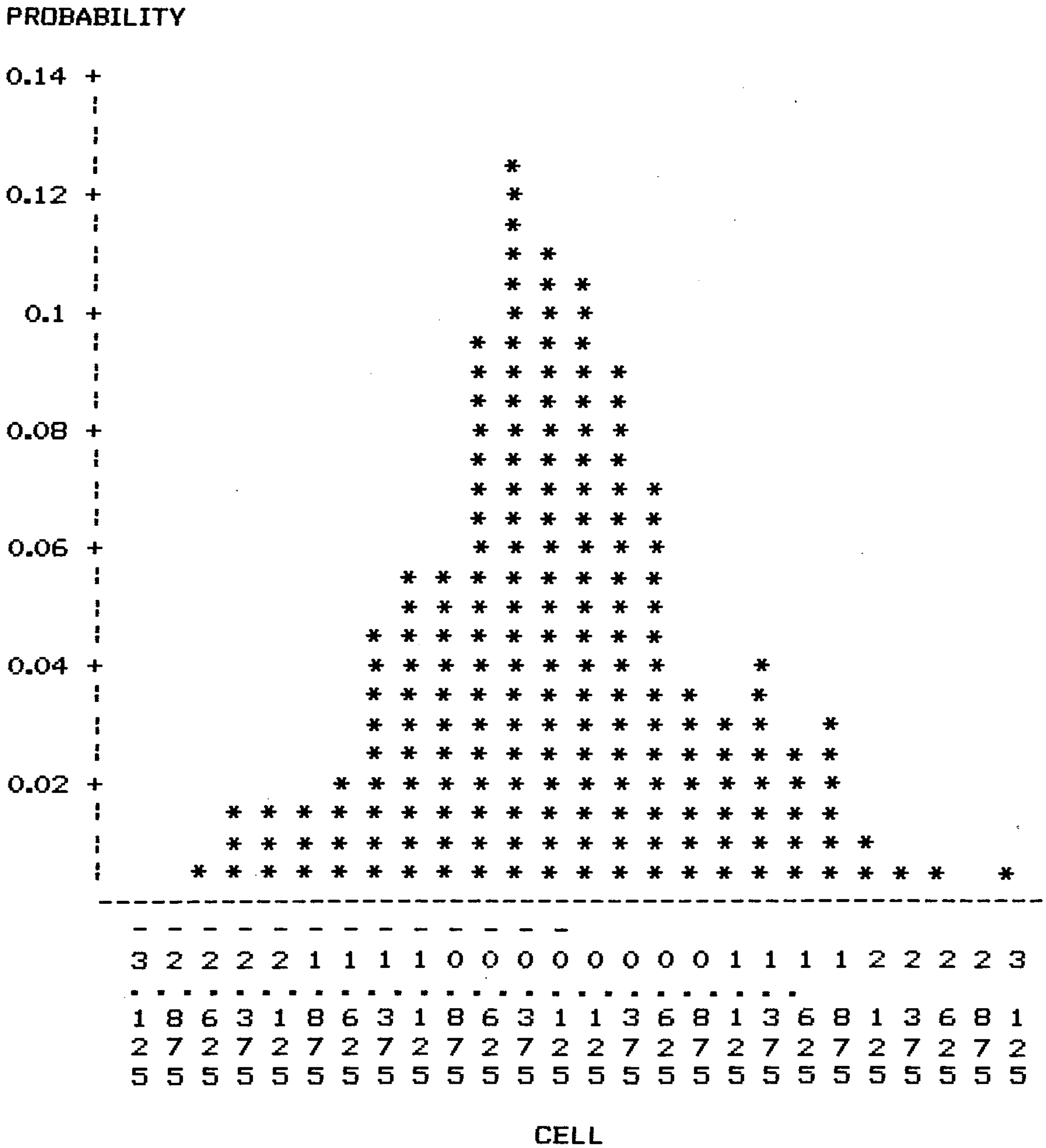
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (5,5), (X11,X21) = (L.9,L.95)
FIGURE 26 FREQUENCY DISTRIBUTION OF L25HAT



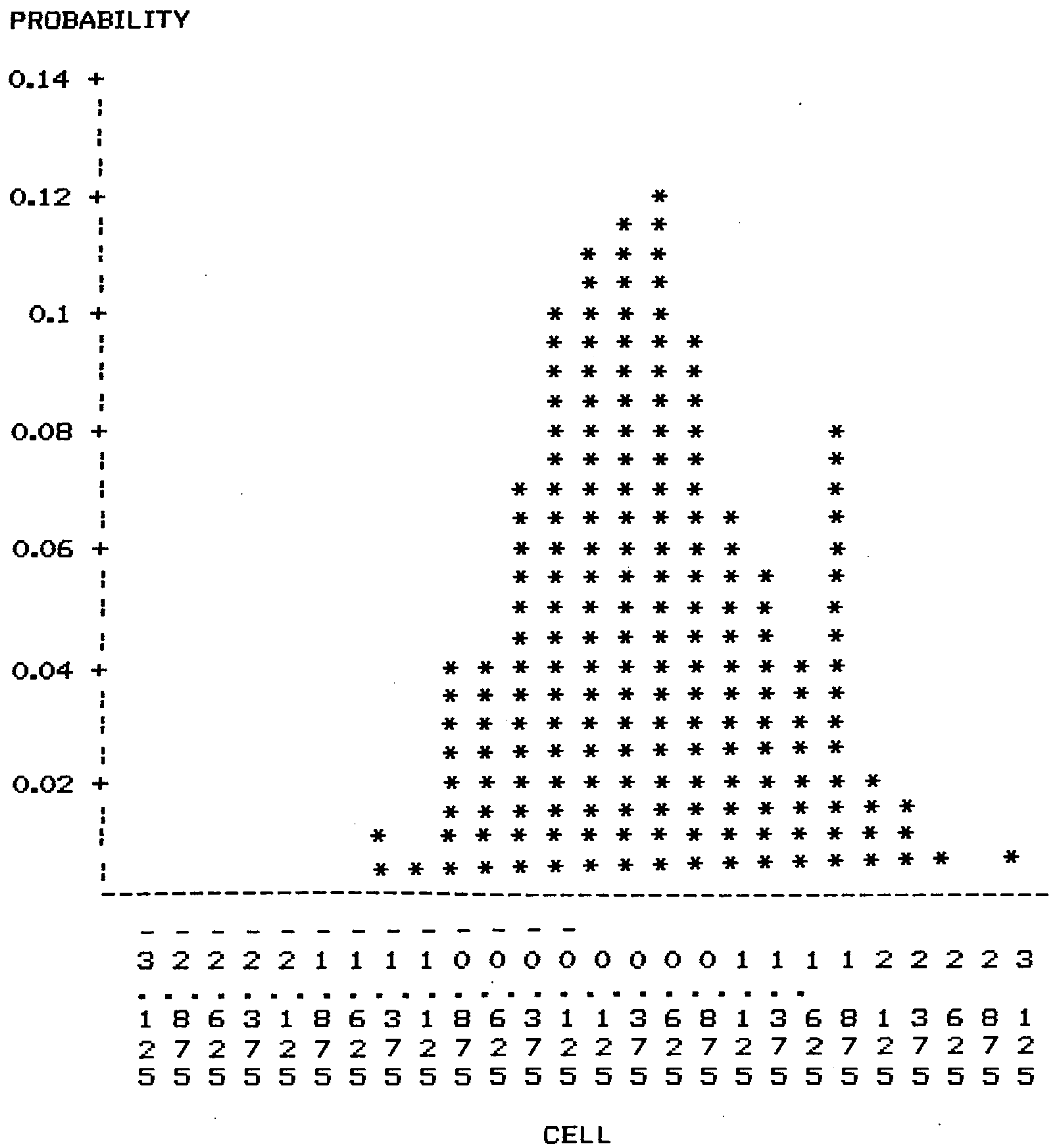
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (5,5), (X11,X21) = (L.9,L.95)
FIGURE 27 FREQUENCY DISTRIBUTION OF L50HAT



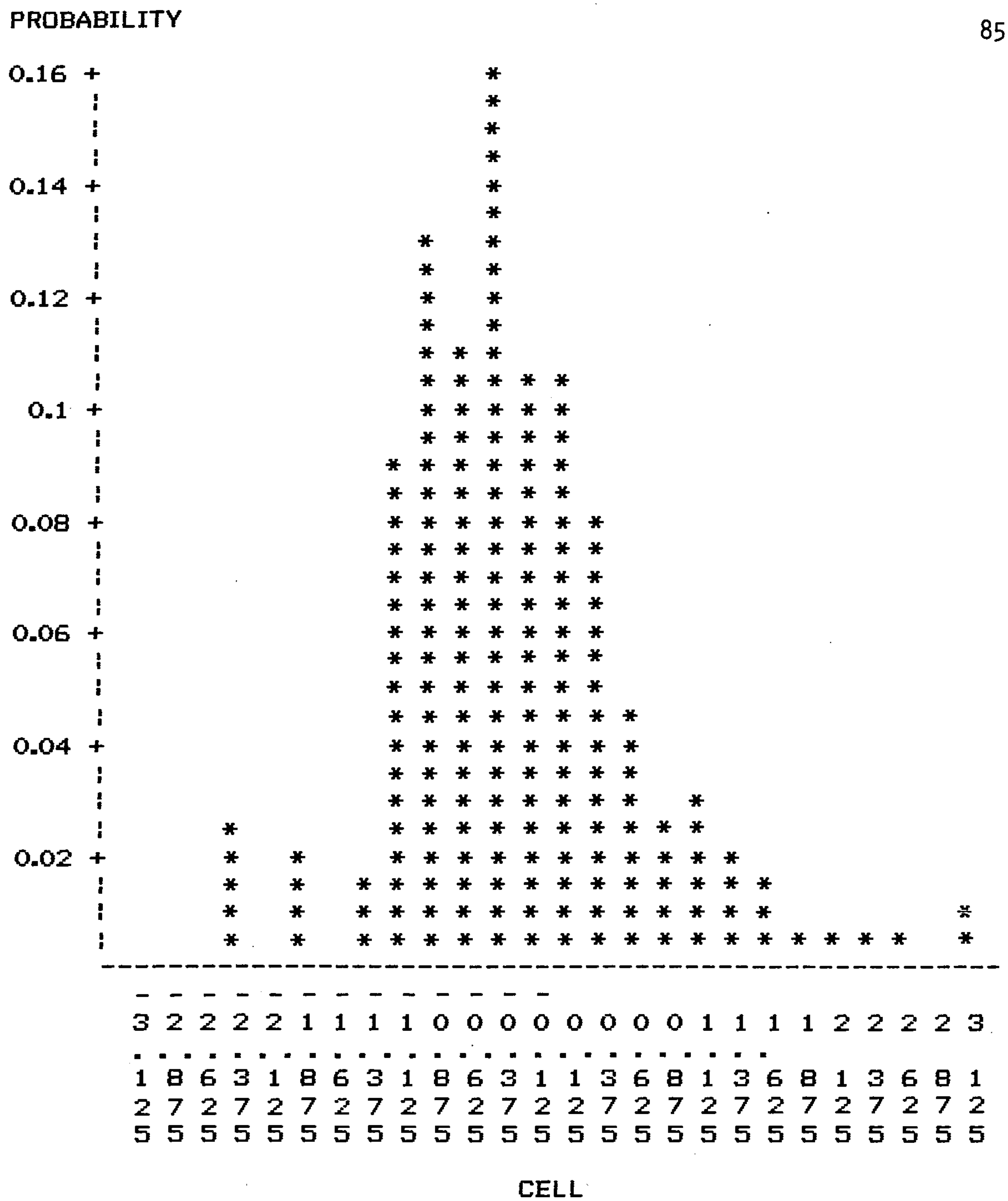
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (5,5), (X11,X21) = (L.9,L.95)
FIGURE 28 FREQUENCY DISTRIBUTION OF L50HAT



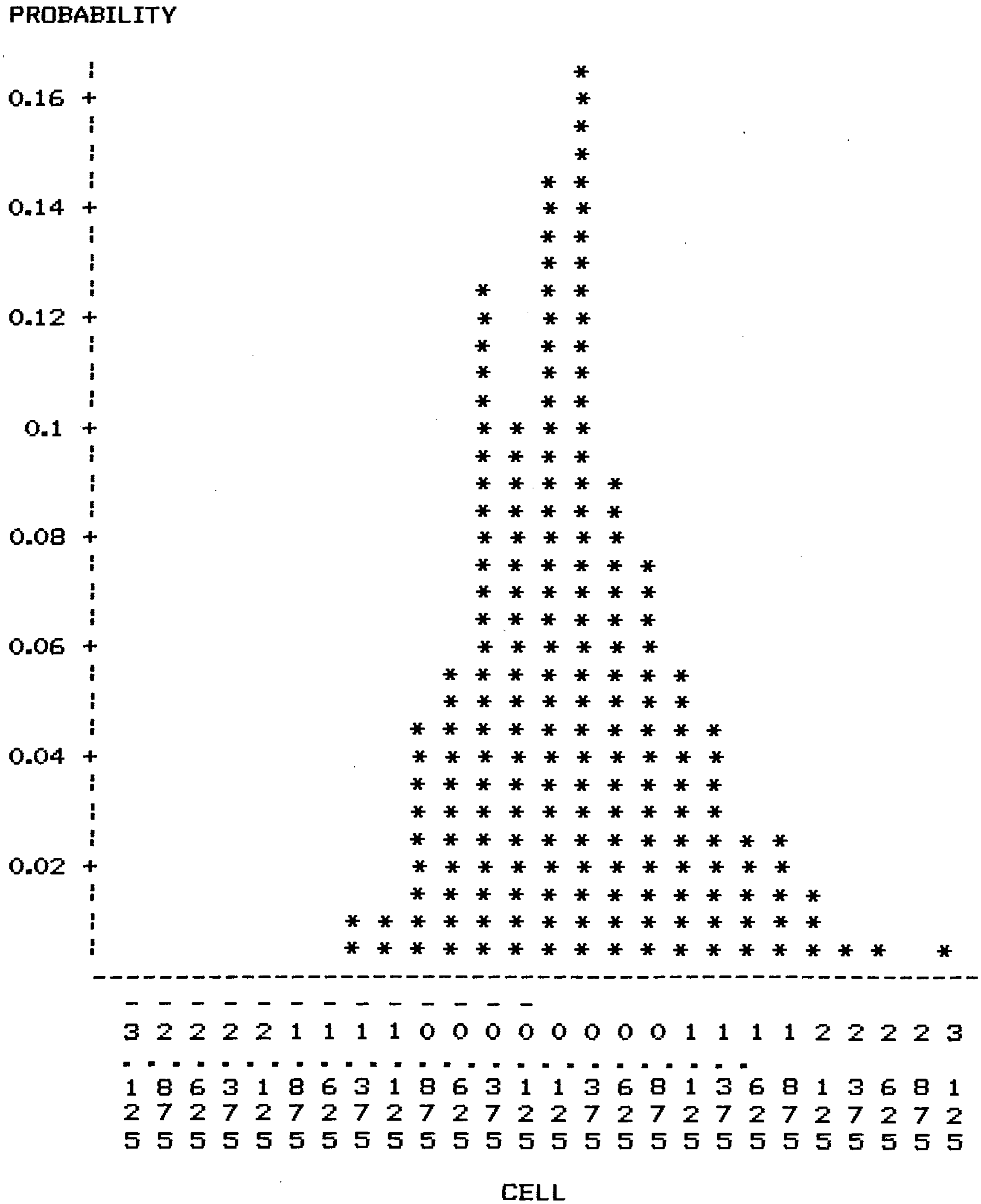
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (5,5), (X11,X21) = (L.9,L.95)
FIGURE 29 FREQUENCY DISTRIBUTION OF L75



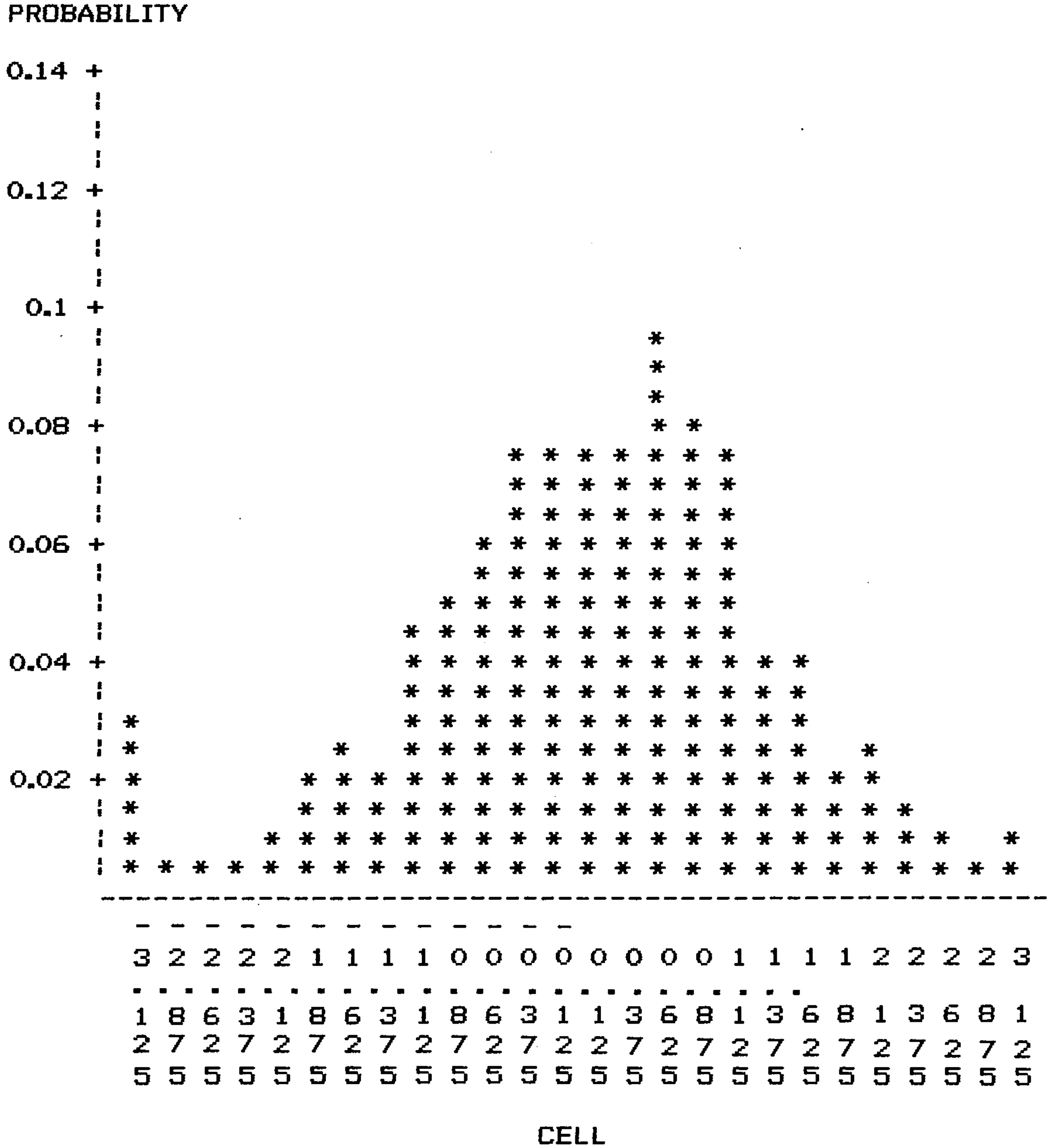
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (5,5), (X11,X21) = (L.9,L.95)
FIGURE 30 FREQUENCY DISTRIBUTION OF L75HAT



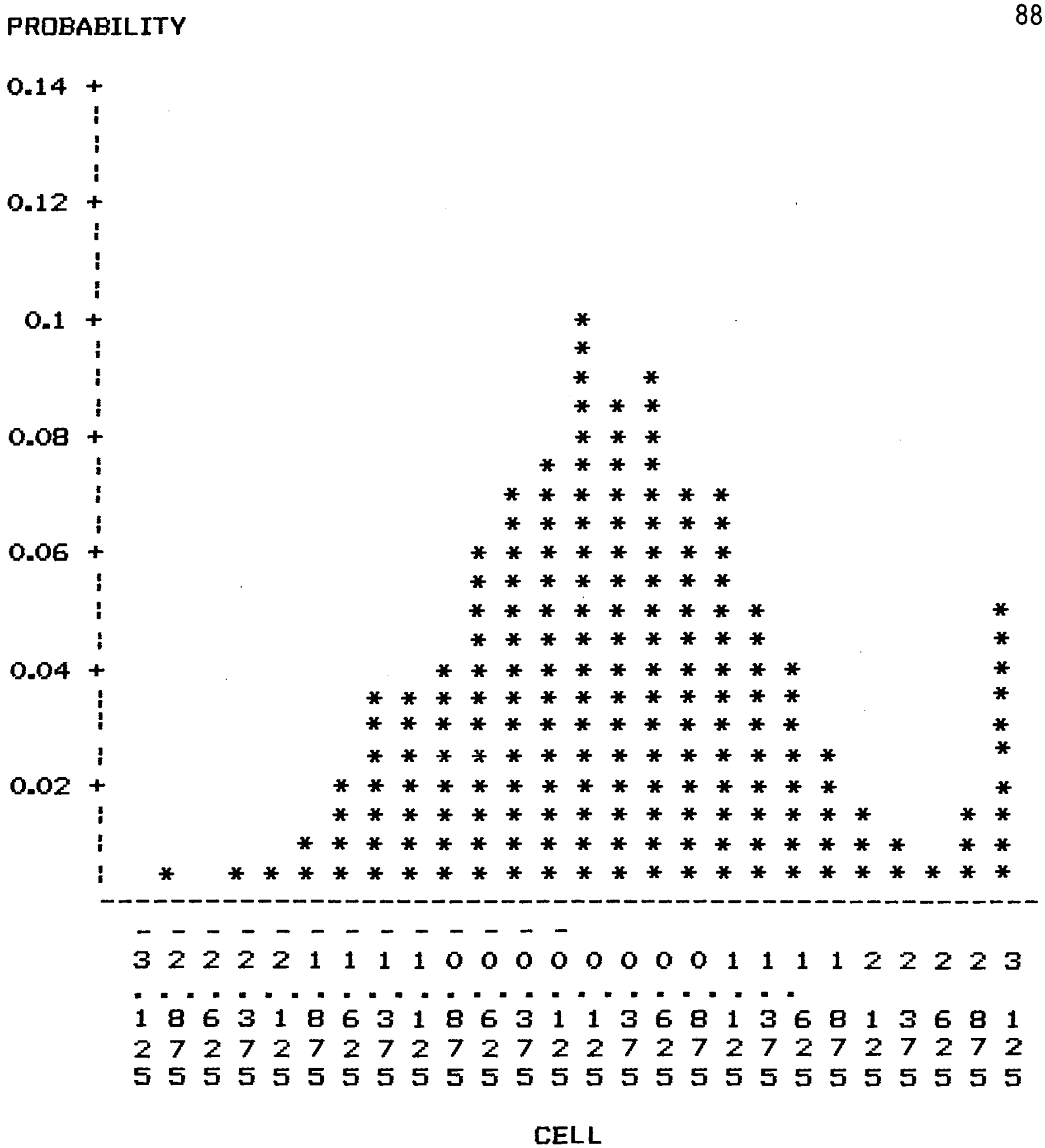
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (5,5), (X11,X21) = (L.9,L.95)
FIGURE 31 FREQUENCY DISTRIBUTION OF L90HAT



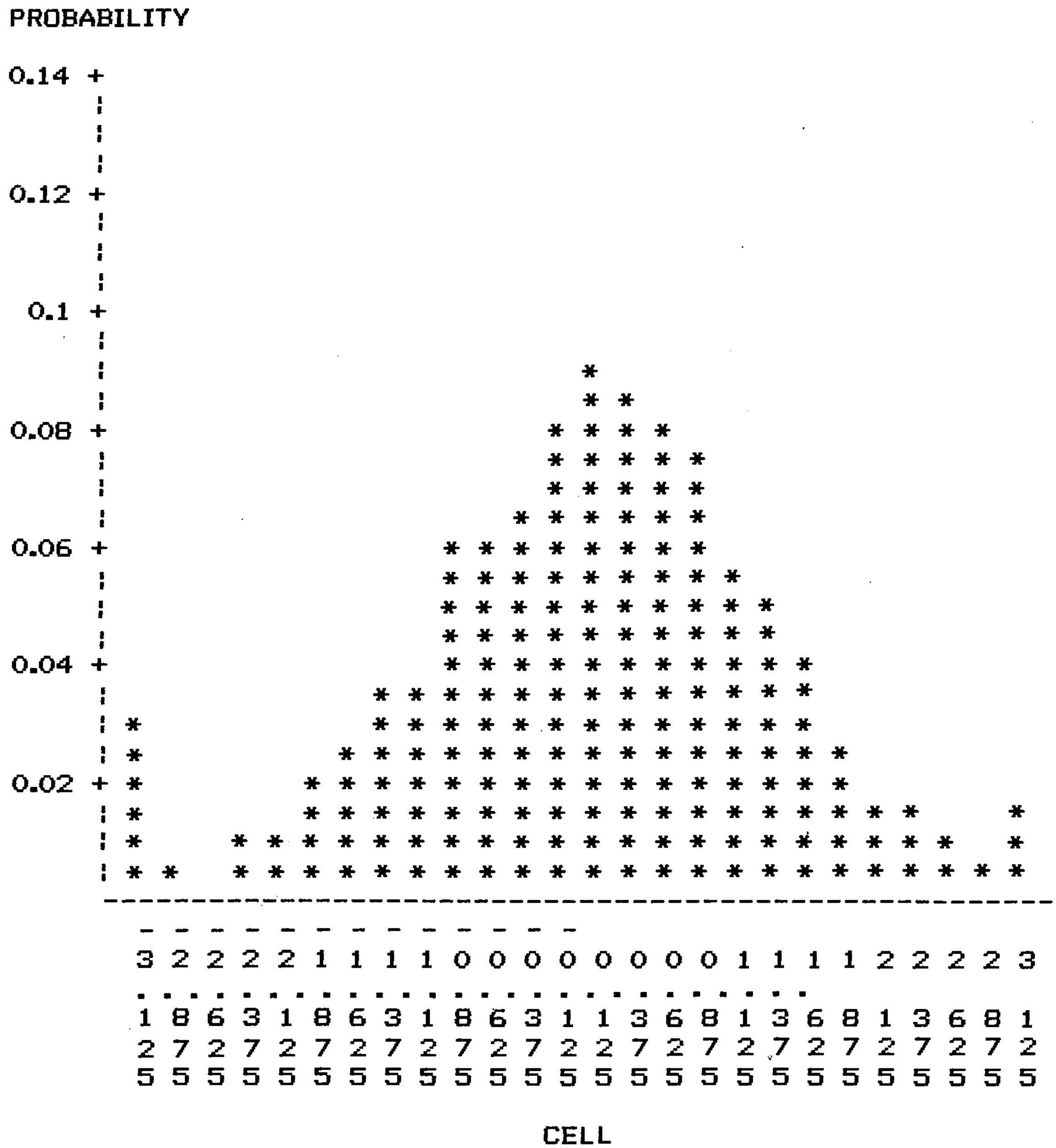
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (5,5), (X11,X21) = (L.9,L.95)
FIGURE 32 FREQUENCY DISTRIBUTION OF L90HAT



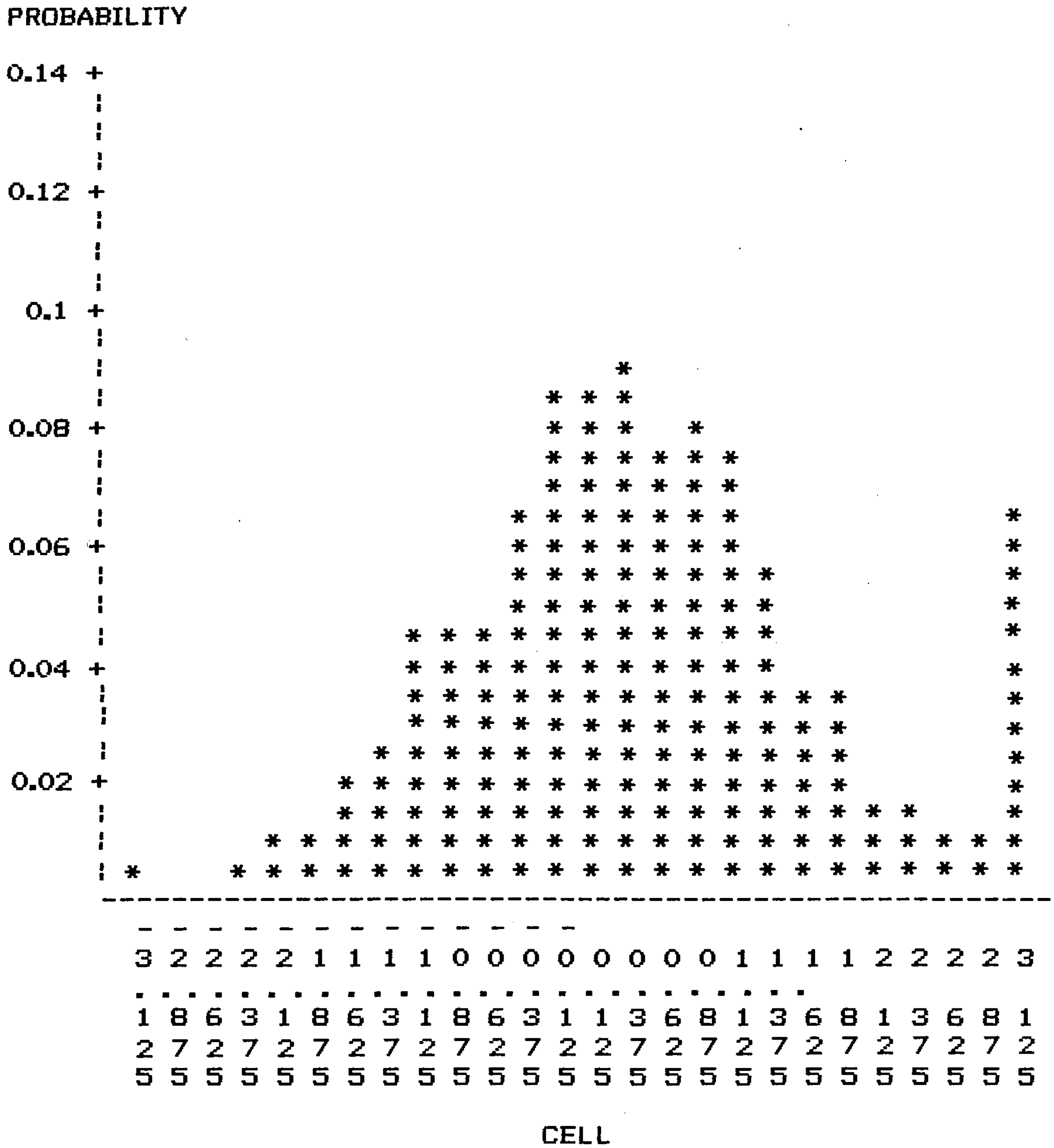
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (40,40), (X11,X21) = (L.9,L.95)
FIGURE 33 FREQUENCY DISTRIBUTION OF L10HAT



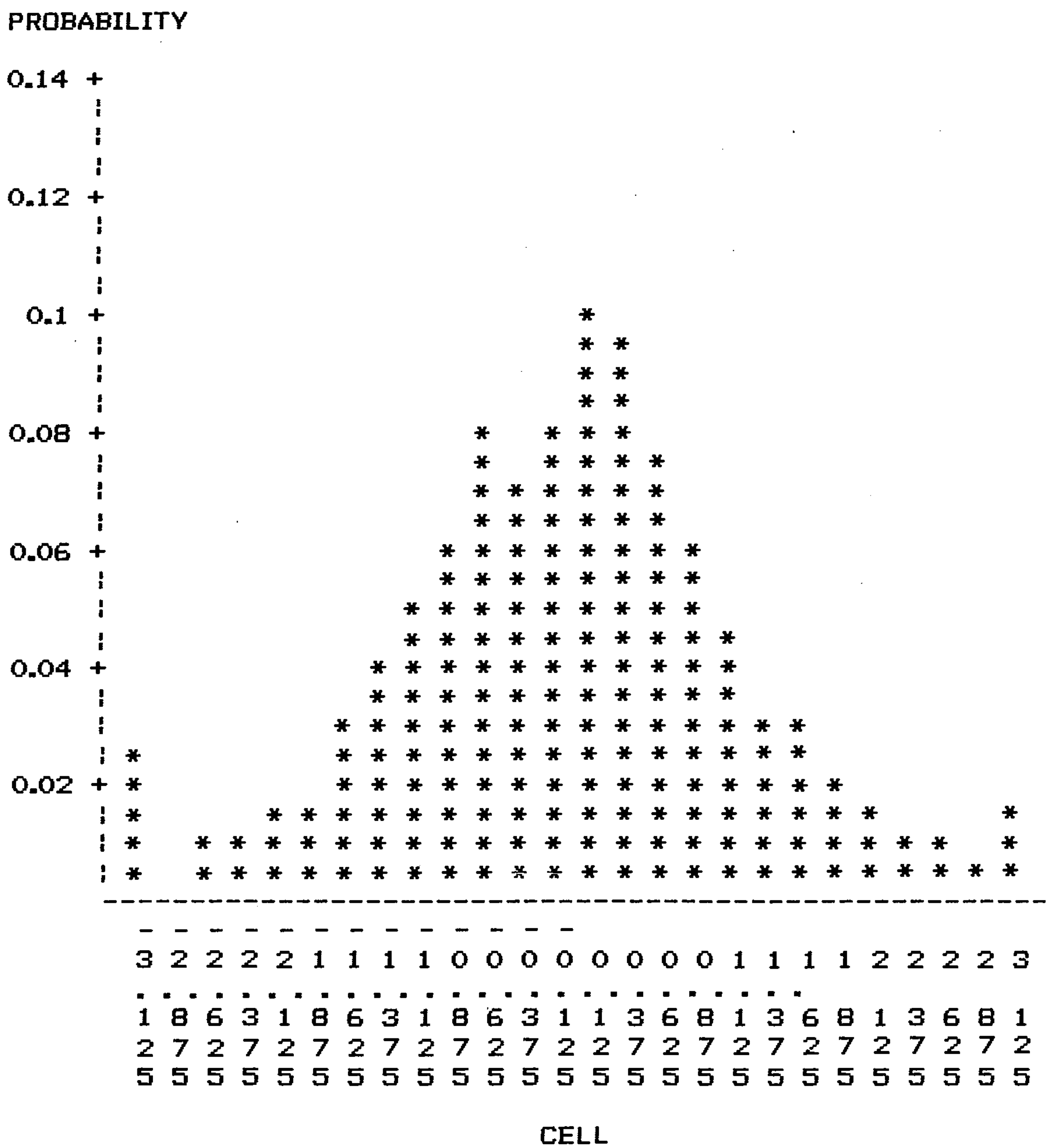
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (40,40), (X11,X21) = (L.9,L.95)
FIGURE 34 FREQUENCY DISTRIBUTION OF L10HAT



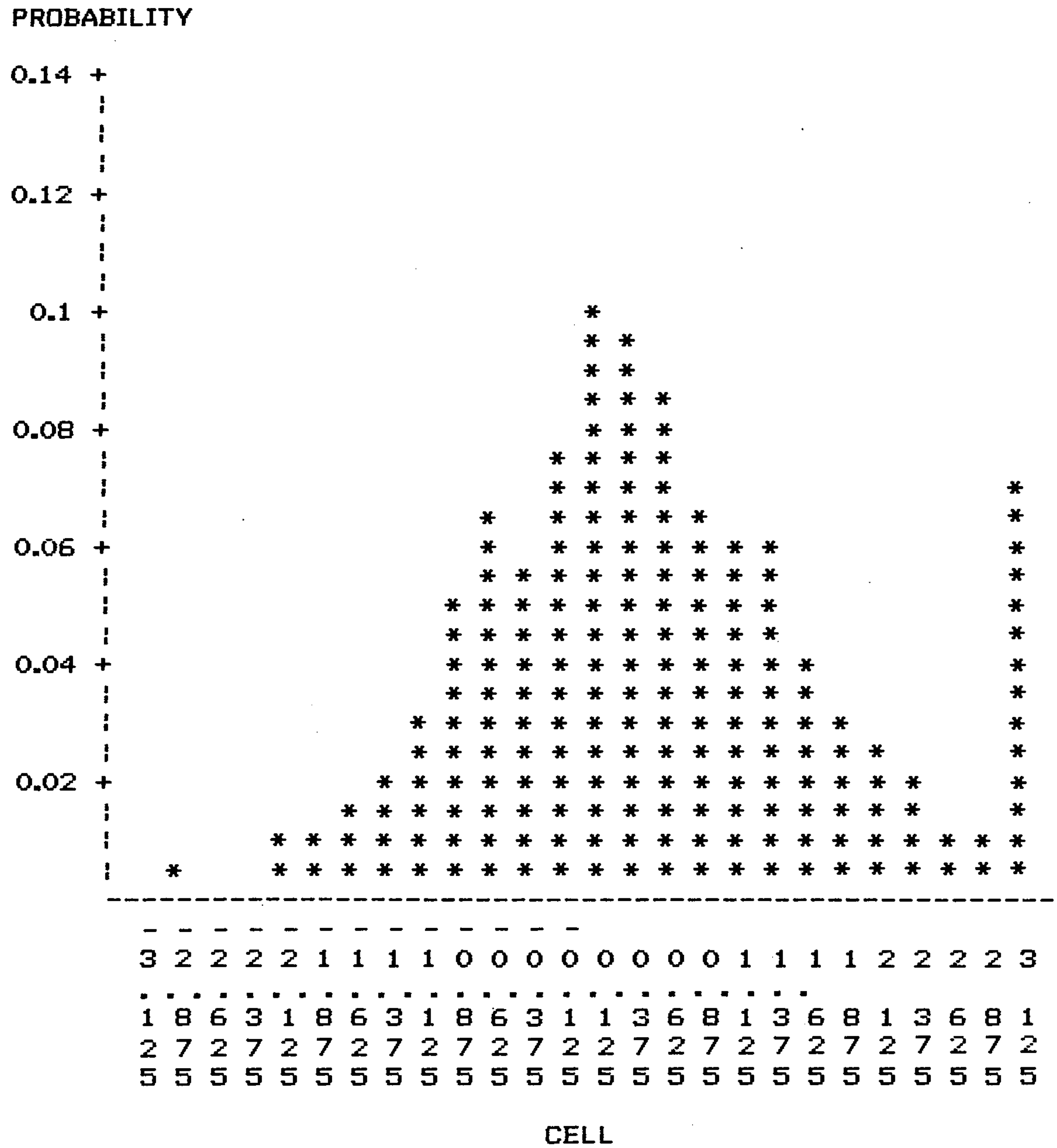
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (40,40), (X11,X21) = (L.9,L.95)
FIGURE 35 FREQUENCY DISTRIBUTION OF L25HAT



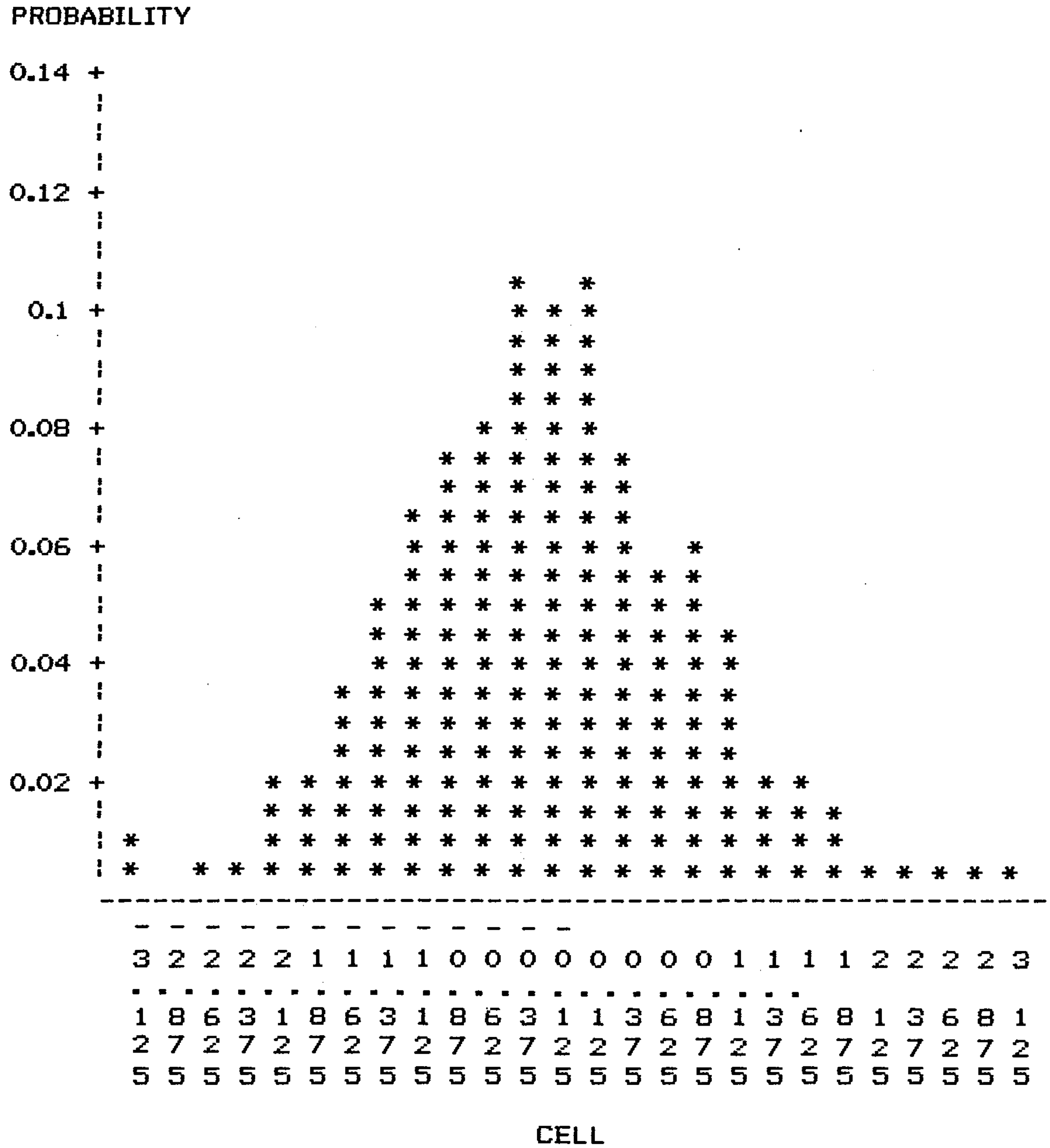
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (40,40), (X11,X21) = (L.9,L.95)
FIGURE 36 FREQUENCY DISTRIBUTION OF L25HAT



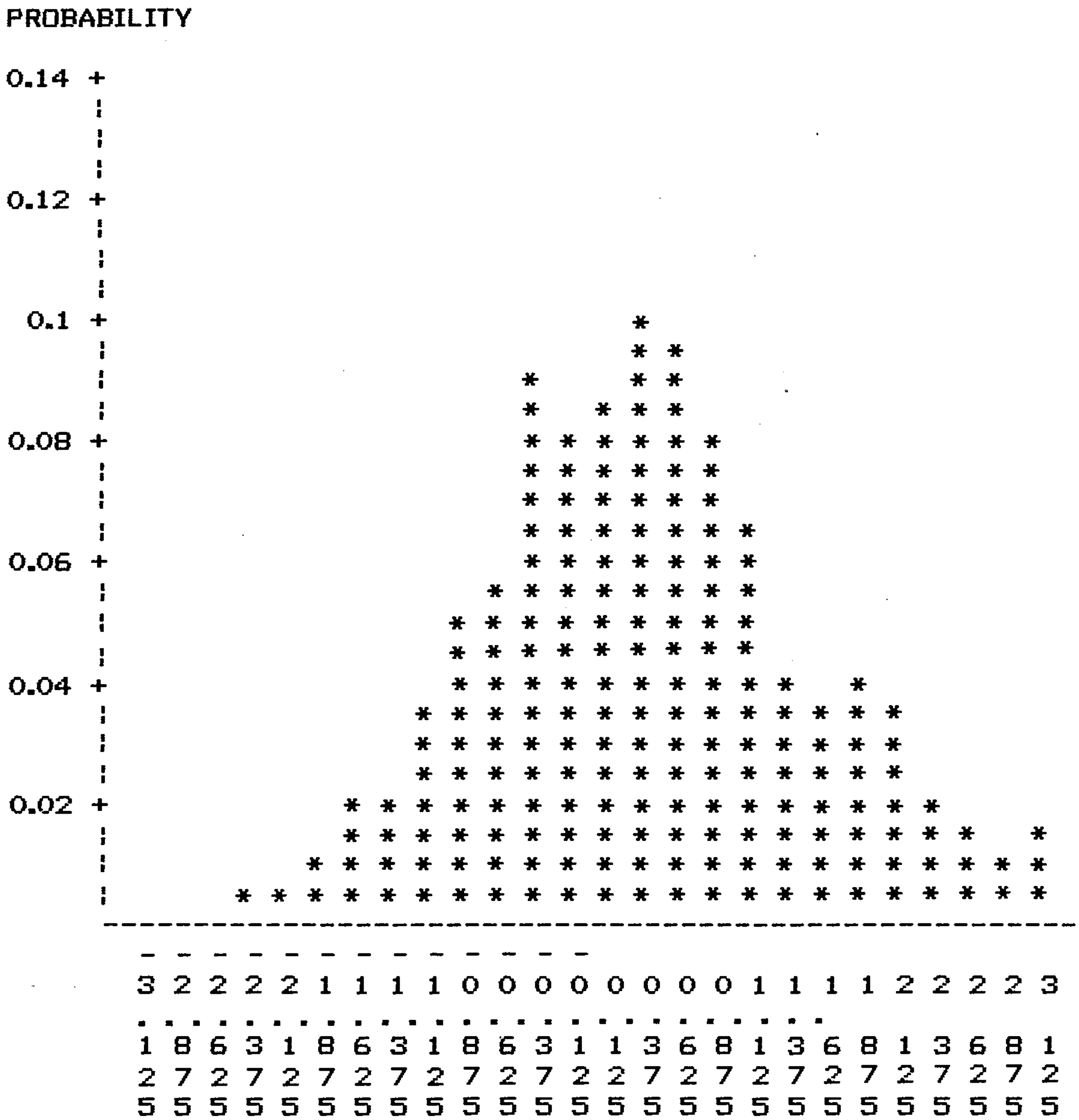
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (40,40), (X11,X21) = (L.9,L.95)
FIGURE 37 FREQUENCY DISTRIBUTION OF L50HAT



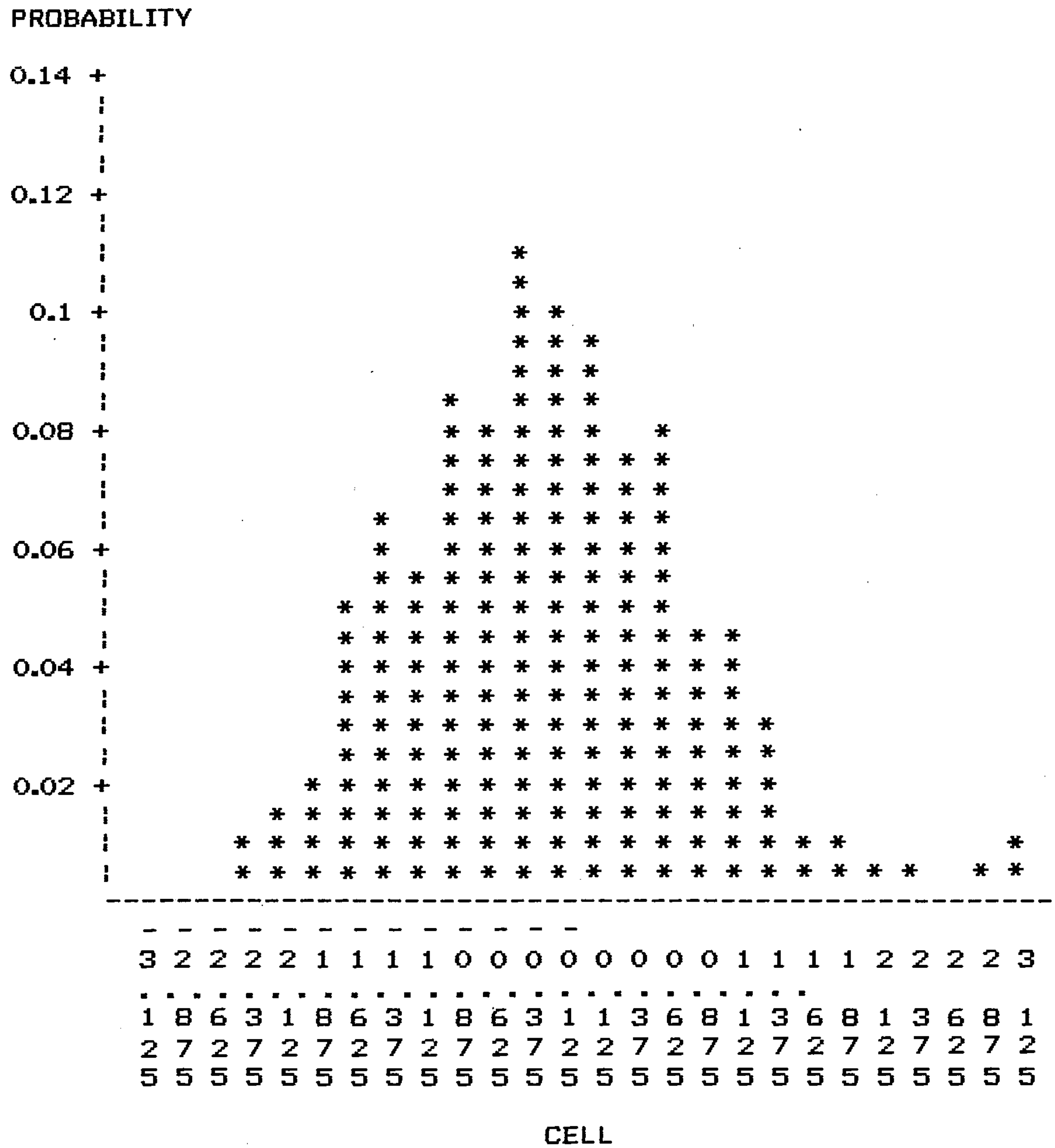
SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (40,40), (X11,X21) = (L.9,L.95)
FIGURE 38 FREQUENCY DISTRIBUTION OF L50HAT



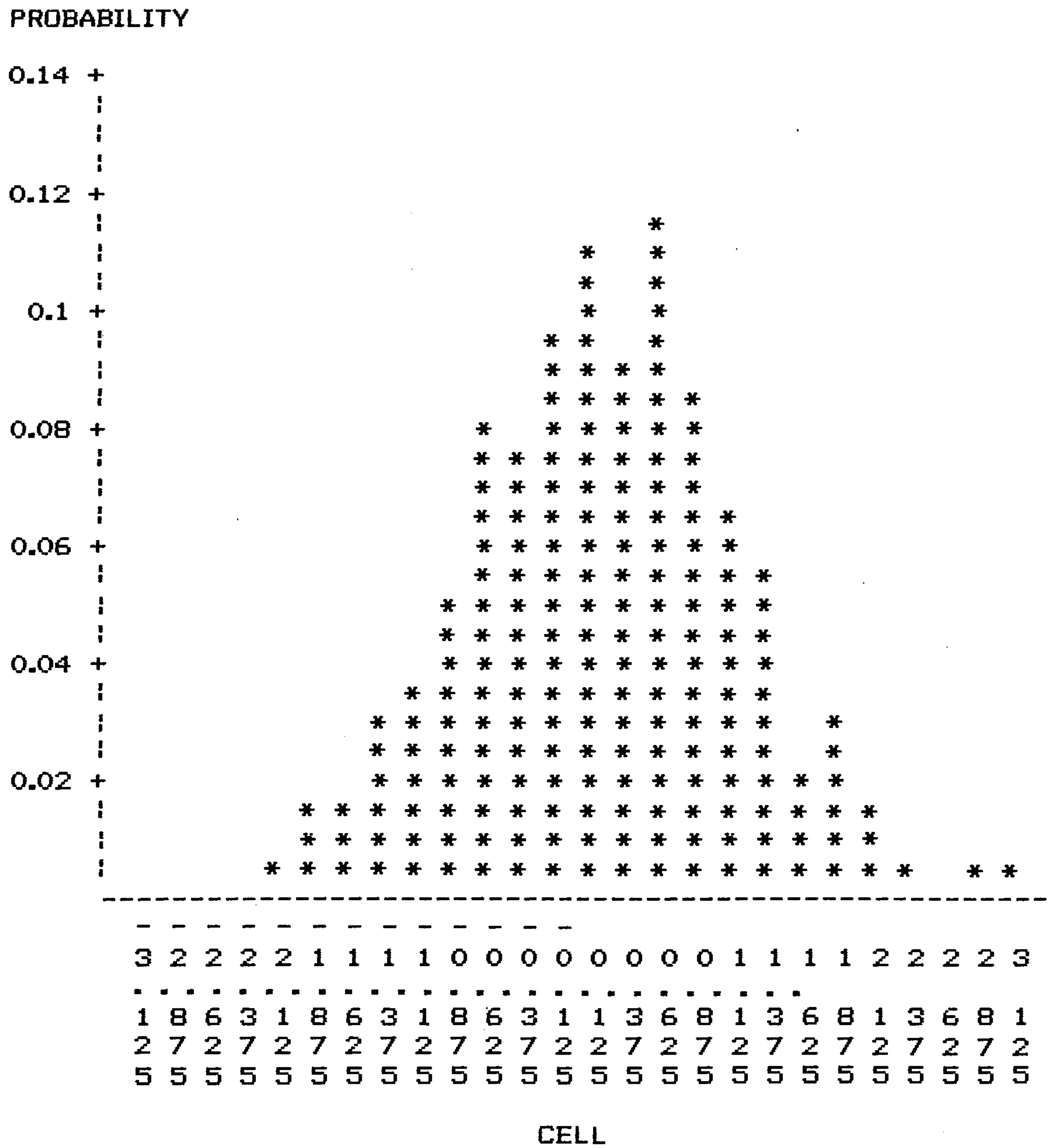
SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (40,40), (X11,X21) = (L.9,L.95)
FIGURE 39 FREQUENCY DISTRIBUTION OF L75HAT



SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (40,40), (X11,X21) = (L.9,L.95)
FIGURE 40 FREQUENCY DISTRIBUTION OF L75HAT



SIMULATION RESULTS FOR TWO STAGE METHOD
(m,n) = (40,40), (X11,X21) = (L.9,L.95)
FIGURE 41 FREQUENCY DISTRIBUTION OF L90HAT



SIMULATION RESULTS FOR SEQUENTIAL METHOD
(m,n) = (40,40), (X11,X21) = (L.9,L.95)
FIGURE 42 FREQUENCY DISTRIBUTION OF L90HAT

APPENDIX A

EXTENSION TO OTHER BINARY MODELS

Let Y be a dichotomous random variable such that $Pr\{Y(x)=1\} = G(\theta_1+\theta_2 x)$. The expectation of Y is given by $E(Y)=Pr\{Y(x)=1\}(1)+Pr\{Y(x)=0\}(0)=Pr\{Y(x)=1\}=G(\theta_1+\theta_2 x)$. Let p be any value $0<p<1$, then solving $p=G(\theta_1+\theta_2 x)$ for x we obtain $L_p = (-\theta_1+G^{-1}(p))/\theta_2$, where L_p is the value of x such that $Pr\{Y(x)=1|x\}=p$. Hence the TSE of L_p is the MLE of L_p based on $(2m+n_1+n_2)$ samples:

$$\hat{L}_p = (-\hat{\theta}_1+G^{-1}(p))/\hat{\theta}_2$$

where $\hat{\theta}_1$, and $\hat{\theta}_2$ are the MLE of θ_1 and θ_2, respectively.

For notational convenience let $G(\theta_1+\theta_2 x) = G$ and $G(\theta_1+\theta_2 x_i^*) = G^*$. Also let $\partial/\partial\theta_1\{G(\theta_1+\theta_2 x_{ij})\} = G_1$, $\partial/\partial\theta_2\{G(\theta_1+\theta_2 x_{ij})\}= G_2$, $\partial^2/\partial\theta_1^2\{G(\theta_1+\theta_2 x_{ij})\} = G_{11}$, $\partial^2/\partial\theta_2\partial\theta_1\{G(\theta_1+\theta_2 x)_{ij}\} = G_{12}$ and $\partial^2/\partial\theta_2^2\{G(\theta_1+\theta_2 x_{ij})\} = G_{22}$. Moreover, let these notations also hold for G^*.

The likelihood function is given by

$$L(\theta_1,\theta_2) = \prod_{i=1}^{2}\prod_{j=1}^{m} G^{y_{ij}}[1-G]^{1-y_{ij}} \prod_{i=1}^{2}\prod_{j=m+1}^{m+n_i} G^{*y_{ij}}[1-G^*]^{1-y_{ij}}$$

The logarithm of the likelihood function is

$$\ell = \sum_{i=1}^{2}\sum_{j=1}^{m}\left\{y_{ij}\log[G] + (1-y_{ij})\log[1-G]\right\}$$
$$+ \sum_{i=1}^{2}\sum_{j=m+1}^{m+n_i}\left\{y_{ij}\log[G^*] + (1-y_{ij})\log[1-G^*]\right\}$$

$$\frac{\partial \ell}{\partial \theta_1} = \sum_{i=1}^{2}\sum_{j=1}^{m}\left\{y_{ij}\,\frac{G_1}{G} + (1-y_{ij})\frac{-G_1}{[1-G]}\right\}$$

$$+\sum_{i=1}^{2}\sum_{j=m+1}^{m+n_i}\left\{y_{ij}\,\frac{G_1^*}{G^*} + (1-y_{ij})\frac{-G_1^*}{[1-G^*]}\right\}$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{m}\frac{(y_{ij}-G)G_1}{G(1-G)} + \sum_{i=1}^{2}\sum_{j=m+1}^{m+n_i}\frac{(y_{ij}-G^*)G_1^*}{G^*(1-G^*)}$$

$$\frac{\partial^2 \ell}{\partial \theta_1^2} = \sum_{i=1}^{2}\sum_{j=1}^{m}\frac{G(1-G)\{y_{ij}G_{11}-GG_{11}-G_1^2\}-(y_{ij}-G)G_1(G_1-2GG_1)}{[G(1-G)]^2}$$

$$+\sum_{i=1}^{2}\sum_{j=m+1}^{m+n_i}\frac{G^*(1-G^*)\{y_{ij}G_{11}^*-G^*G_{11}^*-G_1^{*2}\}-(y_{ij}-G^*)G_1^*(G_1^*-2G^*G_1^*)}{[G^*(1-G^*)]^2}$$

$$\frac{\partial^2 \ell}{\partial \theta_2 \partial \theta_1} = \sum_{i=1}^{2}\sum_{j=1}^{m}\frac{G(1-G)\{y_{ij}G_{12}-GG_{12}-G_1G_2\}-(G_2-2GG_2)(y_{ij}-G)G_1}{[G(1-G)]^2}$$

$$+\sum_{i=1}^{2}\sum_{j=m+1}^{m+n_i}\frac{G^*(1-G^*)\{y_{ij}G_{12}^*-G^*G_{12}^*-G_1^*G_2^*\}-(G_2^*-2G^*G_2^*)(y_{ij}-G^*)G_1^*}{[G^*(1-G^*)]^2}$$

$$\frac{\partial \ell}{\partial \theta_2} = \sum_{i=1}^{2}\sum_{j=1}^{m}\frac{(y_{ij}-G)G_2}{G(1-G)} + \sum_{i=1}^{2}\sum_{j=m+1}^{m+n_i}\frac{(y_{ij}-G^*)G_2^*}{G^*(1-G^*)} \quad \text{and}$$

$$\frac{\partial^2 \ell}{\partial \theta_2^2} = \sum_{i=1}^{2}\sum_{j=1}^{m}\frac{G(1-G)\{y_{ij}G_{22}-G_2^2-GG_{22}\}-(G_2-2GG_2)(y_{ij}-G)G_2}{[G(1-G)]^2}$$

$$+\sum_{i=1}^{2}\sum_{j=1}^{m+n_i}\frac{G^*(1-G^*)\{y_{ij}G_{22}^*-G_2^{*2}-G^*G_{22}^*\}-(G_2^*-2G^*G_2^*)(y_{ij}-G^*)G_2^*}{[G^*(1-G^*)]^2}$$

Let $a = -E\dfrac{\partial^2 \ell}{\partial \theta_1^2}$, $G = -E\dfrac{\partial^2 \ell}{\partial \theta_2 \partial \theta_1}$, and $c = -E\dfrac{\partial^2 \ell}{\partial \theta_2^2}$.

We note that

$$E\left\{G(1-G)\left[y_{ij}G_{11}-GG_{11}-G_1^2\right]\right\} = -G(1-G)G_1^2,$$

$$E\{(y_{ij}-G)G_1(G_1-2GG_1)\} = 0,$$

$$E\left\{G(1-G)\left[y_{ij}G_{12}-GG_{12}-G_1G_2\right]\right\} = -G(1-G)G_1G_2,$$

$$E\{(y_{ij}-G)G_1(G_2-2GG_2)\} = 0,$$

$$E\left\{G(1-G)\left[y_{ij}G_{22}-GG_{22}-G_2^2\right]\right\} = -G(1-G)G_2^2,$$

$$E\{(y_{ij}-G)G_2(G_2-2GG_2)\} = 0.$$

Similarly, the above relations are true for the parts involving G^*. Hence we have the following simplified results:

$$a = \sum_{i=1}^{2}\sum_{j=1}^{m}\frac{G_1^2}{G(1-G)} + \sum_{i=1}^{2} n_i \frac{G_1^{*2}}{G^*(1-G^*)}$$

$$b = \sum_{i=1}^{2}\sum_{j=1}^{m}\frac{G_1G_2}{G(1-G)} + \sum_{i=1}^{2} n_i \frac{G_1^*G_2^*}{G^*(1-G^*)}$$

$$c = \sum_{i=1}^{2}\sum_{j=1}^{m}\frac{G_2^2}{G(1-G)} + \sum_{i=1}^{2} n_i \frac{G_2^{*2}}{G^*(1-G^*)}$$

Then the variance of $\begin{bmatrix}\hat{\theta}_1 \\ \hat{\theta}_2\end{bmatrix}$ and $\hat{L}_p$ is given by

(i) $$\frac{1}{(ac-b^2)}\begin{bmatrix} c & -b \\ -b & a \end{bmatrix}$$

and

(ii) $$\frac{aL_p^2 - 2bL_p + c}{(\theta_2)^2(ac-b^2)}$$

,respectively.

(2) For the logit model

$$G(\theta_1+\theta_2x_{ij})=G=P_{ij}=\exp(\theta_1+\theta_2x_{ij})/(1+\exp(\theta_1+\theta_2x_{ij})),$$

$$G(\theta_1+\theta_2x_i^*)=G^*=P_i=\exp(\theta_1+\theta_2x_i^*)/(1+\exp(\theta_1+\theta_2x_i^*)),$$

$$G_1(\theta_1+\theta_2x_{ij})=\frac{(1+\exp(\theta_1+\theta_2x_{ij}))\exp(\theta_1+\theta_2x_{ij})-[\exp(\theta_1+\theta_2x_{ij})]^2}{[1+\exp(\theta_1+\theta_2x_{ij})]^2}$$

$$=p_{ij}(1-p_{ij})$$

and $G_2(\theta_1+\theta_2x_{ij})=p_{ij}x_{ij}-[p_{ij}]^2x_{ij}$

$$=p_{ij}(1-p_{ij})x_{ij}.$$

Similarly, $G_1^* = P_i(1-P_i)$ and $G_2^* = P_i(1-P_i)x_i^*$.

$$\therefore\ a = \sum_{i=1}^{2}\sum_{j=1}^{m}\frac{[P_{ij}(1-P_{ij})]^2}{P_{ij}(1-P_{ij})} + \sum_{i=1}^{2}n_i\frac{[P_i(1-P_i)]^2}{P_i(1-P_i)}$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{m}P_{ij}(1-P_{ij}) + \sum_{i=1}^{2}n_iP_i(1-P_i)$$

$$b = \sum_{i=1}^{2}\sum_{j=1}^{m}\frac{[P_{ij}(1-P_{ij})]^2x_{ij}}{P_{ij}(1-P_{ij})} + \sum_{i=1}^{2}n_i\frac{[P_i(1-P_i)]^2x_i^*}{P_i(1-P_i)}$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{m}P_{ij}(1-P_{ij})x_{ij} + \sum_{i=1}^{2}n_iP_i(1-P_i)x_i^*$$

$$c = \sum_{i=1}^{2}\sum_{j=1}^{m}\frac{[P_{ij}(1-P_{ij})x_{ij}]^2}{P_{ij}(1-P_{ij})} + \sum_{i=1}^{2}n_i\frac{[P_i(1-P_i)x_i^*]^2}{P_i(1-P_i)}$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{m}P_{ij}(1-P_{ij})x_{ij}^{\ 2} + \sum_{i=1}^{2}n_iP_i(1-P_i)x_i^{*2}$$

Hence the generated result above agrees with the result obtained in equations (3.2.13), (3.2.14) and (3.2.15), respectively.

(3) Now consider the log-log model,

$Pr\{y(x)=1\}=1-\exp(-\exp[\theta_1+\theta_2x])=G(\theta_1+\theta_2x)$. Then

$$\hat{L}_p=(-\hat{\theta}_1+\log(-\log(1-p)))/\hat{\theta}_2$$

where L_p is the value of p such that $p=Pr\{Y(x)=1|x)$ and $\hat{\theta}_1$ and $\hat{\theta}_2$ are the MLE's of θ_1 and θ_2 based on $(2m+n_1+n_2)$ two stage samples. Taking the partial derivatives with respect to θ_1 and θ_2, we obtain

$$G_1=\exp(-\exp[\theta_1+\theta_2x_{ij}])\exp(\theta_1+\theta_2x_{ij})=-(1-p_{ij})\log(1-p_{ij}).$$

$$G_2=-(1-p_{ij})\log(1-p_{ij})x_{ij}.$$

Using P_{ij} for G in the equations derived in part 1, and replacing G by G^* whenever x_{ij} is replaced by x_i^* ,we have:

$$a=\sum_{i=1}^{2}\sum_{j=1}^{m}\frac{[(1-P_{ij})\log(1-P_{ij})]^2}{P_{ij}(1-P_{ij})}+\sum_{i=1}^{2}n_i\frac{[(1-P_i)\log(1-P_i)]^2}{P_i(1-P_i)}$$

$$=\sum_{i=1}^{2}\sum_{j=1}^{m}\frac{(1-P_{ij})}{P_{ij}}[\log(1-P_{ij})]^2+\sum_{i=1}^{2}n_i\frac{(1-P_i)}{P_i}[\log(1-P_i)]^2$$

$$b=\sum_{i=1}^{2}\sum_{j=1}^{m}\frac{[(1-P_{ij})\log(1-P_{ij})]^2x_{ij}}{P_{ij}(1-P_{ij})}+\sum_{i=1}^{2}n_i\frac{[(1-P_i)\log(1-P_i)]^2x_i^*}{P_i(1-P_i)}$$

$$=\sum_{i=1}^{2}\sum_{j=1}^{m}\frac{[(1-P_{ij})\log(1-P_{ij})]^2x_{ij}}{P_{ij}(1-P_{ij})}+\sum_{i=1}^{2}n_i\frac{[(1-P_i)\log(1-P_i)]^2x_i^*}{P_i(1-P_i)}$$

$$=\sum_{i=1}^{2}\sum_{j=1}^{m}\frac{(1-P_{ij})}{P_{ij}}[\log(1-P_{ij})]^2x_{ij}+\sum_{i=1}^{2}n_i\frac{(1-P_i)}{P_i}[\log(1-P_i)]^2x_i^*$$

$$c=\sum_{i=1}^{2}\frac{[(1-P_{ij})\log(1-P_{ij})x_{ij}]^2}{P_{ij}(1-P_{ij})}+\sum_{i=1}^{2}n_i\frac{[(1-P_i)\log(1-P_i)x_i^*]^2}{P_i(1-P_i)}$$

$$=\sum_{i=1}^{2}\sum_{j=1}^{m}\frac{(1-P_{ij})}{P_{ij}}[\log(1-P_{ij})]^2x_{ij}^{\ 2}+\sum_{i=1}^{2}n_i\frac{(1-P_i)}{P_i}[\log(1-P_i)]^2x_i^{*2}$$

Then the variance of the TSE of $(\theta_1,\theta_2)'$ and L_p for the log-log model is given by (i) and (ii), respectively.

APPENDIX B

SAS CODE FOR THE SIMULATION STUDY

```
FILENAME ----- 'A:TSSQSM--.DAT';
DATA ONE;
OPTIONS PS = 80;

ARRAY X1(21);
ARRAY X2(21);
ARRAY Y1(20);
ARRAY Y2(20);
ARRAY P10CL(26);
ARRAY P25CL(26);
ARRAY P50CL(26);
ARRAY P75CL(26);
ARRAY P90CL(26);
ARRAY P101CL(26);
ARRAY P251CL(26);
ARRAY P501CL(26);
ARRAY P751CL(26);
ARRAY P901CL(26);
*
* INPUT THE SEQUENTIAL AND FIXED SAMPLE SIZES
*;
M = --;
N = --;
MP1 = M + 1;
MP2 = M + 2;
MPNM1 = M + N - 1;
MPN = M + N;
*
* INPUT THE NUMBER OF SIMULATION RUNS
*;
SIMNUM = 1000;
*
* INPUT STARTING VALUES FOR MU AND BETA
*;
MU = 0;
BETA = 1;
*
* INPUT STARTING VALUES FOR X1(1) AND X2(1)
*;
X1(1) = INITIAL 1;
X2(1) = INITIAL 2;
*
* INPUT VALUES FOR P1 AND P2
*;
P1 = 0.2;
P2 = 0.8;
*
* INPUT VALUES OF C1 AND C2
*;
```

```
C1  = 0.44;
C2  = 0.44;

 *
 * INITIALIZE CELL PROBABILITIES
 *;
 DO S = 1 TO 26;
 P10CL(S) = 0;
 P25CL(S) = 0;
 P50CL(S) = 0;
 P75CL(S) = 0;
 P90CL(S) = 0;
 P101CL(S) = 0;
 P251CL(S) = 0;
 P501CL(S) = 0;
 P751CL(S) = 0;
 P901CL(S) = 0;
 END;
 *
 * INITIALIZE SUMS FOR MEAN, VARIANCE AND BIAS CALCULATIONS
 *;
 SUML102 = 0;
 SUML10 = 0;
 SUML252 = 0;
 SUML25 = 0;
 SUML502 = 0;
 SUML50 = 0;
 SUML752 = 0;
 SUML75 = 0;
 SUML902 = 0;
 SUML90 = 0;
 SUM1L102 = 0;
 SUM1L10 = 0;
 SUM1L252 = 0;
 SUM1L25 = 0;
 SUM1L502 = 0;
 SUM1L50 = 0;
 SUM1L752 = 0;
 SUM1L75 = 0;
 SUM1L902 = 0;
 SUM1L90 = 0;
 *
 * THIS IS WHERE THE SIMULATION PORTION OF
 * THE PROGRAM BEGINS
 *;
 DO P = 1 TO SIMNUM;
 *
 * THIS IS WHERE THE TWO STAGE ESTIMATION
 * PORTION OF THE PROGRAM BEGINS
 *;
 REDO1:
 DO RR=1 TO M;
 RRP1 = RR + 1;
```

```
Y1(RR) = 0;
Y2(RR) = 0;
PR1 = 1/(1 + EXP(-MU - BETA*X1(RR)));
PR2 = 1/(1 + EXP(-MU - BETA*X2(RR)));
IF PR1 > UNIFORM(SEED1) THEN Y1(RR) = 1;
IF PR2 > UNIFORM(SEED2) THEN Y2(RR) = 1;
X1(RRP1) = X1(RR) - (1/RR)*(X2(RR)
              - X1(RR))*(Y1(RR) - P1)/C1;
X2(RRP1) = X2(RR) - (1/RR)*(X2(RR)
              - X1(RR))*(Y2(RR) - P2)/C2;
DIFF1 = X1(RRP1) - X1(RR);
DIFF2 = X2(RRP1) - X2(RR);
IF DIFF1 > 5 THEN X1(RRP1) = X1(RR) + 5;
IF DIFF2 > 5 THEN X2(RRP1) = X2(RR) + 5;
IF DIFF1 < -5 THEN X1(RRP1) = X1(RR) - 5;
IF DIFF2 < -5 THEN X2(RRP1) = X2(RR) - 5;

IF X1(RRP1) > X2(RRP1) THEN X1(RRP1) = (X1(RR)
                                        + X2(RR))/2 - .5;
IF X1(RRP1) > X2(RRP1) THEN X2(RRP1) = (X1(RR)
                                        + X2(RR))/2 + .5;
END;
* FIND X1MAX, X1MIN, XOMAX AND XOMIN, THE LARGEST
* AND SMALLEST VALUES OF X THAT YIELD A ONE & THE LARGEST
* AND SMALLEST VALUES OF X THAT YIELD A ZERO, RESPECTIVELY.
*;
XOMIN =  50;
X1MIN =  50;
XOMAX = -50;
X1MAX = -50;
DO AAA = 1 TO M;
   IF Y1(AAA) = 0 THEN DO;
      IF X1(AAA) GT XOMAX THEN XOMAX = X1(AAA);
      IF X1(AAA) LT XOMIN THEN XOMIN = X1(AAA);
   END;
   IF Y1(AAA) = 1 THEN DO;
      IF X1(AAA) GT X1MAX THEN X1MAX = X1(AAA);
      IF X1(AAA) LT X1MIN THEN X1MIN = X1(AAA);
   END;
   IF Y2(AAA) = 0 THEN DO;
      IF X2(AAA) GT XOMAX THEN XOMAX = X2(AAA);
      IF X2(AAA) LT XOMIN THEN XOMIN = X2(AAA);
   END;
   IF Y2(AAA) = 1 THEN DO;
      IF X2(AAA) GT X1MAX THEN X1MAX = X2(AAA);
      IF X2(AAA) LT X1MIN THEN X1MIN = X2(AAA);
   END;

END;
PR1 = 1/(1 + EXP(-MU - BETA*X1(RRP1)));
PR2 = 1/(1 + EXP(-MU - BETA*X2(RRP1)));
K1 = RANBIN(SEED3,N,PR1);
K2 = RANBIN(SEED4,N,PR2);
```

```
    IF K1 NE N AND X1(RRP1) GT XOMAX THEN XOMAX = X1(RRP1);
    IF K1 NE N AND X1(RRP1) LT XOMIN THEN XOMIN = X1(RRP1);
    IF K2 NE O AND X2(RRP1) GT X1MAX THEN X1MAX = X2(RRP1);
    IF K2 NE O AND X2(RRP1) LT X1MIN THEN X1MIN = X2(RRP1);
    IF K1 NE O AND X1(RRP1) GT X1MAX THEN X1MAX = X1(RRP1);
    IF K1 NE O AND X1(RRP1) LT X1MIN THEN X1MIN = X1(RRP1);
    IF K2 NE N AND X2(RRP1) GT XOMAX THEN XOMAX = X2(RRP1);
    IF K2 NE N AND X2(RRP1) LT XOMIN THEN XOMIN = X2(RRP1);
    IF XOMAX LE X1MIN OR X1MAX LE XOMIN THEN GO TO REDO1;


MUHAT = O;
BETAHAT = O;
DO B = 1 TO 10;
H11 = O;
H12 = O;
H22 = O;
G1 = O;
G2 = O;
DO A = 1 TO M;
    Z1 = MUHAT + BETAHAT*X1(A);
    IF Z1 GT 15 THEN PRED1 = 1;
    IF Z1 GT 15 THEN GO TO DE3;
    IF Z1 LT -15 THEN PRED1 = O;
    IF Z1 LT -15 THEN GO TO DE3;
    PRED1 = EXP(Z1)/(1 + EXP(Z1));
DE3: Z2 = MUHAT + BETAHAT*X2(A);
    IF Z2 GT 15 THEN PRED2 = 1;
    IF Z2 GT 15 THEN GO TO DE4;
    IF Z2 LT -15 THEN PRED2 = O;
    IF Z2 LT -15 THEN GO TO DE4;
    PRED2 = EXP(Z2)/(1 + EXP(Z2));
DE4: H11 = H11 - PRED1*(1-PRED1) - PRED2*(1-PRED2);
    H12 = H12 - X1(A)*PRED1*(1-PRED1) - X2(A)*PRED2*(1-PRED2);
    H22 = H22 - X1(A)*X1(A)*PRED1*(1-PRED1)
                  - X2(A)*X2(A)*PRED2*(1-PRED2);
    G1 = G1 + (Y1(A) - PRED1) + (Y2(A) - PRED2);
    G2 = G2 + X1(A)*(Y1(A) - PRED1) + X2(A)*(Y2(A) - PRED2);
END;
    Z1 = MUHAT + BETAHAT*X1(RRP1);
    IF Z1 GT 15 THEN PRED1 = 1;
    IF Z1 GT 15 THEN GO TO DE5;
    IF Z1 LT -15 THEN PRED1 = O;
    IF Z1 LT -15 THEN GO TO DE5;
    PRED1 = EXP(Z1)/(1 + EXP(Z1));
DE5:  Z2 = MUHAT + BETAHAT*X2(RRP1);
    IF Z2 GT 15 THEN PRED2 = 1;
    IF Z2 GT 15 THEN GO TO DE6;
    IF Z2 LT -15 THEN PRED2 = O;
    IF Z2 LT -15 THEN GO TO DE6;
    PRED2 = EXP(Z2)/(1 + EXP(Z2));
DE6: H11 = H11 - N*PRED1*(1-PRED1) - N*PRED2*(1-PRED2);
```

```
 H12 = H12 - N*X1(RRP1)*PRED1*(1-PRED1)
           - N*X2(RRP1)*PRED2*(1-PRED2);
 H22 = H22 - N*X1(RRP1)*X1(RRP1)*PRED1*(1-PRED1) -
             N*X2(RRP1)*X2(RRP1)*PRED2*(1-PRED2);
 G1 = G1 + K1*(1 - PRED1) + (N - K1)*(0 - PRED1)
         + K2*(1 - PRED2) + (N - K2)*(0 - PRED2);
G2 = G2 + X1(RRP1)*K1*(1 - PRED1)
        + X1(RRP1)*(N - K1)*(0 - PRED1)
        + X2(RRP1)*K2*(1 - PRED2)
        + X2(RRP1)*(N - K2)*(0 - PRED2);
DET = (H11*H22) - (H12**2);
HINV11 = H22/DET;
HINV22 = H11/DET;
HINV12 = -(H12/DET);
MUHAT = MUHAT - ((HINV11*G1) + (HINV12*G2));
BETAHAT = BETAHAT - ((HINV12*G1) + (HINV22*G2));
IF (G1**2 + G2**2) LT .0001 THEN B = 10;
END;
L10HAT = -(MUHAT + 2.197225)/BETAHAT;
SL10HAT = (L10HAT + 2.197225)/(3.312854/((M+N)**.5));
L25HAT = -(MUHAT + 1.098612)/BETAHAT;
SL25HAT =(L25HAT + 1.098612)/(2.256103/((M+N)**.5));
L50HAT = -MUHAT/BETAHAT;
SL50HAT = L50HAT/(1.767767/((M+N)**.5));
L75HAT = (-MUHAT + 1.098612)/BETAHAT;
SL75HAT =(L75HAT - 1.098612)/(2.256103/((M+N)**.5));
L90HAT = (-MUHAT + 2.197225)/BETAHAT;
SL90HAT =(L90HAT - 2.197225)/(3.312854/((M+N)**.5));
N10 = INT(4*(SL10HAT + 3) +2);
IF SL10HAT < -3 THEN N10 = 1;
IF SL10HAT >  3 THEN N10 = 26;
P10CL(N10) = P10CL(N10) + 1/SIMNUM;
N25 = INT(4*(SL25HAT + 3) +2);
IF SL25HAT < -3 THEN N25 = 1;
IF SL25HAT >  3 THEN N25 = 26;
P25CL(N25) = P25CL(N25) + 1/SIMNUM;
N50 = INT(4*(SL50HAT + 3) +2);
IF SL50HAT < -3 THEN N50 = 1;
IF SL50HAT >  3 THEN N50 = 26;
P50CL(N50) = P50CL(N50) + 1/SIMNUM;
N75 = INT(4*(SL75HAT + 3) +2);
IF SL75HAT < -3 THEN N75 = 1;
IF SL75HAT >  3 THEN N75 = 26;
P75CL(N75) = P75CL(N75) + 1/SIMNUM;
N90 = INT(4*(SL90HAT + 3) +2);
IF SL90HAT < -3 THEN N90 = 1;
IF SL90HAT >  3 THEN N90 = 26;
P90CL(N90) = P90CL(N90) + 1/SIMNUM;
SUML102 = SUML102 + L10HAT**2;
SUML10 = SUML10 + L10HAT;
SUML252 = SUML252 + L25HAT**2;
SUML25 = SUML25 + L25HAT;
```

```
SUML502 = SUML502 + L50HAT**2;
SUML50 = SUML50 + L50HAT;
SUML752 = SUML752 + L75HAT**2;
SUML75 = SUML75 + L75HAT;
SUML902 = SUML902 + L90HAT**2;
SUML90 = SUML90 + L90HAT;
*
* THIS IS WHERE THE SEQUENTIAL ESTIMATION
* PORTION OF THE PROGRAM BEGINS
*;
DO DD=MP1 TO MPN;
DDP1 = DD + 1;
Y1(DD) = 0;
Y2(DD) = 0;
PR1 = 1/(1 + EXP(-MU - BETA*X1(DD)));
PR2 = 1/(1 + EXP(-MU - BETA*X2(DD)));
IF PR1 > UNIFORM(45305) THEN Y1(DD) = 1;
IF PR2 > UNIFORM(SEED1) THEN Y2(DD) = 1;
X1(DDP1) = X1(DD) - (1/DD)*(X2(DD)
           - X1(DD))*(Y1(DD) - P1)/C1;
X2(DDP1) = X2(DD) - (1/DD)*(X2(DD)
           - X1(DD))*(Y2(DD) - P2)/C2;
DIFF1 = X1(DDP1) - X1(DD);
DIFF2 = X2(DDP1) - X2(DD);
IF DIFF1 > 5 THEN X1(DDP1) = X1(DD) + 5;
IF DIFF2 > 5 THEN X2(DDP1) = X2(DD) + 5;
IF DIFF1 < -5 THEN X1(DDP1) = X1(DD) - 5;
IF DIFF2 < -5 THEN X2(DDP1) = X2(DD) - 5;

IF X1(DDP1) > X2(DDP1) THEN X1(DDP1) = (X1(DD)
                                        + X2(DD))/2 - .5;
IF X1(DDP1) > X2(DDP1) THEN X2(DDP1) = (X1(DD)
                                        + X2(DD))/2 + .5;

END;

*
* THE DEFINITIONS OF MUHAT AND BETAHAT
* FOR THE SEQUENTIAL PORTION
*         ARE GIVEN IN MOSER AND FEI (METRIKA, 1991)
*;

MUHAT = (X2(DDP1)*LOG(P1/P2)
        - X1(DDP1)*LOG(P2/P1))/(X2(DDP1) - X1(DDP1));
BETAHAT = (LOG(P2/P1) - LOG(P1/P2))/(X2(DDP1) - X1(DDP1));

L10HAT1 = -(MUHAT + 2.197225)/BETAHAT;
SL10HAT1 = (L10HAT1 + 2.197225)/(3.312854/((M+N)**.5));
L25HAT1 = -(MUHAT + 1.098612)/BETAHAT;
SL25HAT1 =(L25HAT1 + 1.098612)/(2.256103/((M+N)**.5));
L50HAT1 = -MUHAT/BETAHAT;
SL50HAT1 = L50HAT1/(1.767767/((M+N)**.5));
L75HAT1 = (-MUHAT + 1.098612)/BETAHAT;
SL75HAT1 =(L75HAT1 - 1.098612)/(2.256103/((M+N)**.5));
L90HAT1 = (-MUHAT + 2.197225)/BETAHAT;
```

```
SL90HAT1 =(L90HAT1 - 2.197225)/(3.312854/((M+N)**.5));

N10 = INT(4*(SL10HAT1 + 3) +2);
IF SL10HAT1 < -3 THEN N10 = 1;
IF SL10HAT1 >  3 THEN N10 = 26;
P101CL(N10) = P101CL(N10) + 1/SIMNUM;

N25 = INT(4*(SL25HAT1 + 3) +2);
IF SL25HAT1 < -3 THEN N25 = 1;
IF SL25HAT1 >  3 THEN N25 = 26;
P251CL(N25) = P251CL(N25) + 1/SIMNUM;

N50 = INT(4*(SL50HAT1 + 3) +2);
IF SL50HAT1 < -3 THEN N50 = 1;
IF SL50HAT1 >  3 THEN N50 = 26;
P501CL(N50) = P501CL(N50) + 1/SIMNUM;

N75 = INT(4*(SL75HAT1 + 3) +2);
IF SL75HAT1 < -3 THEN N75 = 1;
IF SL75HAT1 >  3 THEN N75 = 26;
P751CL(N75) = P751CL(N75) + 1/SIMNUM;

N90 = INT(4*(SL90HAT1 + 3) +2);
IF SL90HAT1 < -3 THEN N90 = 1;
IF SL90HAT1 >  3 THEN N90 = 26;
P901CL(N90) = P901CL(N90) + 1/SIMNUM;

SUM1L102 = SUM1L102 + L10HAT1**2;
SUM1L10 = SUM1L10 + L10HAT1;
SUM1L252 = SUM1L252 + L25HAT1**2;
SUM1L25 = SUM1L25 + L25HAT1;
SUM1L502 = SUM1L502 + L50HAT1**2;
SUM1L50 = SUM1L50 + L50HAT1;
SUM1L752 = SUM1L752 + L75HAT1**2;
SUM1L75 = SUM1L75 + L75HAT1;
SUM1L902 = SUM1L902 + L90HAT1**2;
SUM1L90 = SUM1L90 + L90HAT1;
END;
ML10HAT = SUML10/SIMNUM;
VL10HAT = (1/(SIMNUM - 1))*(SUML102 - SIMNUM*ML10HAT**2);
BIASL10 = ML10HAT - LOG(.1/.9);
ML25HAT = SUML25/SIMNUM;
VL25HAT = (1/(SIMNUM - 1))*(SUML252 - SIMNUM*ML25HAT**2);
BIASL25 = ML25HAT - LOG(.25/.75);
ML50HAT = SUML50/SIMNUM;
VL50HAT = (1/(SIMNUM - 1))*(SUML502 - SIMNUM*ML50HAT**2);
BIASL50 = ML50HAT;
ML75HAT = SUML75/SIMNUM;
VL75HAT = (1/(SIMNUM - 1))*(SUML752 - SIMNUM*ML75HAT**2);
BIASL75 = ML75HAT - LOG(.75/.25);
ML90HAT = SUML90/SIMNUM;
VL90HAT = (1/(SIMNUM - 1))*(SUML902 - SIMNUM*ML90HAT**2);
BIASL90 = ML90HAT - LOG(.9/.1);
```

```
ML10HAT1  = SUM1L10/SIMNUM;
VL10HAT1  = (1/(SIMNUM - 1))*(SUM1L102 - SIMNUM*ML10HAT1**2);
BIASL101  = ML10HAT1 - LOG(.1/.9);
ML25HAT1  = SUM1L25/SIMNUM;
VL25HAT1  = (1/(SIMNUM - 1))*(SUM1L252 - SIMNUM*ML25HAT1**2);
BIASL251  = ML25HAT1 - LOG(.25/.75);
ML50HAT1  = SUM1L50/SIMNUM;
VL50HAT1  = (1/(SIMNUM - 1))*(SUM1L502 - SIMNUM*ML50HAT1**2);
BIASL501  = ML50HAT1;
ML75HAT1  = SUM1L75/SIMNUM;
VL75HAT1  = (1/(SIMNUM - 1))*(SUM1L752 - SIMNUM*ML75HAT1**2);
BIASL751  = ML75HAT1 - LOG(.75/.25);
ML90HAT1  = SUM1L90/SIMNUM;
VL90HAT1  = (1/(SIMNUM - 1))*(SUM1L902 - SIMNUM*ML90HAT1**2);
BIASL901  = ML90HAT1 - LOG(.9/.1);

DO T = 1 TO 26;
P10CELL  = P10CL(T);
P25CELL  = P25CL(T);
P50CELL  = P50CL(T);
P75CELL  = P75CL(T);
P90CELL  = P90CL(T);
P10CELL1  = P101CL(T);
P25CELL1  = P251CL(T);
P50CELL1  = P501CL(T);
P75CELL1  = P751CL(T);
P90CELL1  = P901CL(T);
OUTPUT;
END;
KEEP ML10HAT ML25HAT ML50HAT ML75HAT ML90HAT
     ML10HAT1 ML25HAT1 ML50HAT1 ML75HAT1 ML90HAT1
     VL10HAT VL25HAT VL50HAT VL75HAT VL90HAT
     VL10HAT1 VL25HAT1 VL50HAT1 VL75HAT1 VL90HAT1
     BIASL10 BIASL25 BIASL50 BIASL75 BIASL90
     BIASL101 BIASL251 BIASL501 BIASL751 BIASL901
     P10CELL P25CELL P50CELL P75CELL P90CELL
     P10CELL1 P25CELL1 P50CELL1 P75CELL1 P90CELL1
     M N X11 X21;
DATA _NULL_; SET ONE;
FORMAT M 5. N 5. X11 9.5 X21 9.5
   P10CELL 8.4 P10CELL1 8.4 ML10HAT 12.8 ML10HAT1 12.8
   VL10HAT 12.8 VL10HAT1 12.8 BIASL10 12.8 BIASL101 12.8
   P25CELL 8.4 P25CELL1 8.4 ML25HAT 12.8 ML25HAT1 12.8
   VL25HAT 12.8 VL25HAT1 12.8 BIASL25 12.8 BIASL251 12.8
   P50CELL 8.4 P50CELL1 8.4 ML50HAT 12.8 ML50HAT1 12.8
   VL50HAT 12.8 VL50HAT1 12.8 BIASL50 12.8 BIASL501 12.8
   P75CELL 8.4 P75CELL1 8.4 ML75HAT 12.8 ML75HAT1 12.8
   VL75HAT 12.8 VL75HAT1 12.8 BIASL75 12.8 BIASL751 12.8
   P90CELL 8.4 P90CELL1 8.4 ML90HAT 12.8 ML90HAT1 12.8
   VL90HAT 12.8 VL90HAT1 12.8 BIASL90 12.8 BIASL901 12.8;
FILE -----;
IF _N_ = 1 THEN DO;
 PUT '  SIMULATION RESULTS FOR THE TWO STAGE    ';
```

```
PUT '   AND SEQUENTIAL ESTIMATION METHODS.      ';
PUT ' OUTPUTS INCLUDE MEAN, VARIANCE AND BIAS  ';
PUT 'PLUS FREQENCY DISTRIBUTIONS FOR ESTIMATED ';
PUT '    ROOTS  L10, L25, L50, L75 AND L90.    ' ///;
PUT @1 'M' @3 'N' @9 'X11' @17 'X21';
PUT M N X11 X21 ///;
PUT 'MEANS, VARIANCES AND BIASES FOR
THE TWO STAGE METHOD'//;
PUT @4 'ML10' @16 'ML25' @28 'ML50'
@39 'ML75' @50 'ML90';
PUT ML10HAT ML25HAT ML50HAT ML75HAT ML90HAT /;
PUT @3 'VARL10' @14 'VARL25'
@25 'VARL50' @36 'VARL75' @47 'VARL90';
PUT VL10HAT VL25HAT VL50HAT VL75HAT VL90HAT /;
PUT @4 'BIASL10' @16 'BIASL25' @28 'BIASL50'
@39 'BIASL75' @50 'BIASL90';
PUT BIASL10 BIASL25 BIASL50 BIASL75 BIASL90 ///;
PUT 'MEANS, VARIANCES AND BIASES
FOR THE SEQUENTIAL METHOD'//;
PUT @4 'ML10' @16 'ML25' @28 'ML50' @39 'ML75' @50 'ML90';
PUT ML10HAT1 ML25HAT1 ML50HAT1 ML75HAT1 ML90HAT1 /;
PUT @3 'VARL10' @14 'VARL25' @25 'VARL50'
@36 'VARL75' @47 'VARL90';
PUT VL10HAT1 VL25HAT1 VL50HAT1 VL75HAT1 VL90HAT1 /;
PUT @4 'BIASL10' @16 'BIASL25' @28 'BIASL50'
@39 'BIASL75' @50 'BIASL90';
PUT BIASL101 BIASL251 BIASL501 BIASL751 BIASL901 ///;
PUT '  THE FREQUENCY DISTRIBUTIONS OF THE
ESTIMATED L10, L25,  ';
PUT '  L50, L75, L90 FOR THE TWO STAGE
AND SEQUENTIAL METHODS. ' /;
PUT 'THE FREQUENCY DISTRIBUTION BELOW
BELOW IS DIVIDED INTO 26 CELLS.';
PUT 'CELL1 RANGES FROM -INFINITY TO -3,
CELL2 FROM -3 TO -2.75 ';
PUT 'CELL3 FROM -2.75 TO -2.5,
CELL4 FROM -2.5 TO -2.25, ... , ';
PUT '  CELL25 FROM 2.75 TO 3
AND CELL26 FROM 3 TO INFINITY.     ' //;
PUT @14 'TWO STAGE' @51 'SEQUENTIAL';
PUT @1 'L10HAT' @8 'L25HAT'
@15 'L50HAT' @22 'L75HAT' @29 'L90HAT'
@39 'L10HAT' @46 'L25HAT'
@53 'L50HAT' @60 'L75HAT' @67 'L90HAT';
END;
PUT P10CELL P25CELL P50CELL P75CELL P90CELL
    @39 P10CELL1 P25CELL1 P50CELL1 P75CELL1 P90CELL1;
RUN;
```

VITA

Tesfaye Liranso

Candidate for the Degree of
Doctor of Philosophy

Thesis: TWO STAGE ESTIMATION OF THE QUANTILES OF THE LOGIT MODEL

Major Field: Statistics

Biographical:

Personal Data: Born in Hosanna, Ethiopia, September 2, 1958.

Education: Graduated from the Laboratory School, Addis Ababa, Ethiopia, in June 1973; Received Bachelor of Science in Mathematics from Addis Ababa University, Addis Ababa, Ethiopia, in July 1982; Received Master of Science in Statistics from Addis Ababa University, Addis Ababa, Ethiopia, in July 1987; Completed the requirements for the Doctor of Philosophy Degree in Statistics at Oklahoma State University in December, 1993.

Professional Experience: Teaching Assistant, Department of Statistics, Oklahoma State University, Fall 1993; Lecturer and Statistical Consultant, Awassa College of Agriculture, Ethiopia, 1987-1989; Graduate Assistant, Awassa College of Agriculture, Ethiopia, 1982-1983.

Scholarly Organizations: Member of the American Statistical Association.

Honors: Member of Mu Sigma Rho National Statistical honor Fraternity.