STOCHASTIC APPROXIMATION AND NON-PARAMETRIC INTERVAL ESTIMATION IN SENSITIVITY TESTING WHICH INVOLVES QUANTAL RESPONSE DATA

By

PATRICK LOWRY ODELL

Bachelor of Science
University of Texas
Austin, Texas
1952

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1958

Submitted to the Faculty of the Graduate School of
the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY
May, 1962

STOCHASTIC APPROXIMATION AND NON-PARAMETRIC INTERVAL ESTIMATION IN SENSITIVITY TESTING WHICH INVOLVES QUANTAL RESPONSE DATA

Thesis Approved:

Robert D Morrison
Thesis Adviser

Earl E Manhart

David L. Weeks

O.H. Hamilton

W Harrington

Dean of the Graduate School

ACKNOWLEDGMENTS

The author wishes to express his gratitude to William Gordon of United States Naval Weapons Evaluation Facility (USNWEF) for suggesting that the study be made.

The author is also deeply indebted to Dr. L. Wayne Johnson and Dr. Carl E. Marshall for providing financial assistance in the form of a graduate assistantship. Also, I wish to acknowledge the financial aid received from the USNWEF in partial support for this research.

Indebtedness is acknowledged to Dr. Robert D. Morrison for guiding my research and for assisting in the preparation of this thesis; to Drs. David L. Weeks, Olin H. Hamilton, and Harold E. Harrington, for serving as members of my committee. Also, I am especially indebted to Charles E. Antle, USNWEF and Dr. Vanamamalai Seshadri, Southern Methodist University, for their suggestions and their technical review of the first six chapters during the summer of 1961; to Paul E. Pulley and John H. Blankenship, of Oklahoma State University Computing Laboratory, and to Wallace Franck, USNWEF, for programming various exploratory mathematical experiments; and to Miss Jane Duncan for typing the final draft.

Special thanks are reserved for my wife Norma, whose patience and encouraging spirit have made major contributions to the successful completion of this dissertation.

TABLE OF CONTENTS

Chapter		Page
I.	INTRODUCTION	1
II.	THREE STOCHASTIC APPROXIMATION ESTIMATORS	5
III.	APPLICATION OF STOCHASTIC APPROXIMATION METHODS TO QUANTAL RESPONSE PROBLEMS	13
IV.	THE MONTE CARLO SAMPLING PLAN TO STUDY THE RATES OF CONVERGENCE OF THE ESTIMATORS	18
V.	PROPERTIES OF THE SEQUENCE $[c_j]$ ASSOCIATED WITH ESTIMATOR III	29
VI.	RESULTS OF THE EMPIRICAL STUDY	40
VII.	AN APPLICATION OF FINITE MARKOV CHAINS TO UP-AND-DOWN TESTING	46
	1. Preliminary Notions and Notation	47
	2. Theory	49
	3. Numerical Example	62
BIBLIOGRAPHY		65
APPENDIX A		66
	1. Fundamental Definitions	66
	2. Basic Theorems	67

LIST OF TABLES

Table		Page
I.	Comparison of the Mean and Median for Triangular Density Function.	10
II.	Comparison of the Mean and Median for the Gamma Density Function	10
III.	Data for Sampling Procedure	22
IV.	Average Value of x_j for 100 Trials for Estimators I, II, and III When $\alpha=0.05$ and $\Delta k=0$.	24
V.	Average Value of x_j for Estimators I, II, and III When $\alpha=0.10$ and $\Delta k=0$	25
VI.	Average Value of x_j for Estimators I, II, and III When $\alpha=0.30$ and $\Delta k=0$	26
VII.	Average Value of x_j for 100 Trials for Estimators I, II, and III When $\alpha=.50$ and $\Delta k=0$	27
VIII.	Average Value of x_j for 100 Trials for Estimators I, II, and III When $k=1$ and $\Delta k=5$.	28

LIST OF FIGURES

Figure		Page
1.	The Regression Function and the Associated Curve Illustrating the Probability of Making an Incorrect Decision in Direction	16
2.	The Five Forms of M(x)	20
3.	The Effect of the Selection of [c/j] on the Rate of Convergence (Estimator I)	21
4.	The Variation of the Estimator for Various Values of c (Estimator I)	21
5.	Comparisons of Rates of Convergence of Estimators I, II, and III	41
6.	The Effect of Sample Size on the Rate of Convergence When $\alpha = 0.10$	42
7.	Sketch of M(x) in Neighborhood of $x = \theta$	58

CHAPTER I

INTRODUCTION

Sensitivity testing deals with a continuous variable which cannot be determined in practice. For example, suppose it is desirable to know the amount of mass of a high explosive such that the probability that an explosive response will occur when the mass is subjected to a jet-fuel fire is less than some specified level, say ν. There are levels of mass at which the probability of response is less than ν and levels at which the probability of response exceeds ν. Clearly, the critical value of mass at which the probability of response is exactly ν cannot be measured. All one can do is select a sample arbitrarily and determine whether the critical value for a sample is less than or greater than the mass of each element of the sample.

This situation arises in many fields of research. In selecting insecticides, a critical dose is associated with each insect but cannot be measured. One can only try some dose and observe whether or not the preassigned percentage of insects are killed, i.e., observe whether or not the desired dose for the insect is less than the chosen dose. The same difficulty arises in pharmaceutical research dealing with germicides, anaesthetics, and other drugs, in testing strengths of materials, and in several areas of engineering and developmental research.

In true sensitivity experiments, it is not possible to make more than one observation on a given specimen. Once a test has been made, the specimen is altered (e.g., the explosive is destroyed, the insect weakened) so that a bona fide result cannot be obtained from a second test on the same specimen. The common procedure in experiments of this kind is to divide the sample of specimens into several groups (usually, but not necessarily, of the same size) and to test one group at a chosen level, and a second group at a second level, etc. The data consist of the numbers affected and not affected at each level. Several methods of analyzing such data (variously called sensitivity data, all-or-none data, or quantal responses) are available. (1), (2).

Most of the methods commonly used are applicable only in special cases, most of which are based on various assumptions concerning the distributions of the estimators, expecially if confidence limits are desired. A method, devised relatively recently (and seldom used for various reasons), is available to the experimenter in which he may estimate any critical value. The method, called a stochastic approximation method, was formulated by Robbins and Monro (3).

Briefly stated, stochastic approximation is concerned with the regression of a variable y on a variable x, and seeks the value $x = \theta$ for which the regression value of y is some preassigned number, $y = \alpha$. The estimation procedure for θ is sequential and distribution-free. Despite its extreme simplicity in application and the wide variety of the situations in which it may be useful, the technique has not been taken advantage of by empirical research workers. One reason for this may be that the existing literature is addressed primarily to the professional mathematician. Another reason may be that the mathematical

theory itself is not yet complete for relatively small samples.

A desirable feature of stochastic approximation is the lack of assumptions required. In many problems, the researcher has no clear picture of the structure of the relationship he wishes to study and would prefer, if possible, not to commit himself to hypothesize the precise shapes of the regression or other distribution features. In such cases, he needs a procedure which is distribution-free.

Theoretically, the problem reduces to solving the regression equation

$$M(x) = \alpha \tag{1}$$

This problem has been studied by Robbins and Monro (3), Blum (4), Keston (5), and others (6, 7, and 8). Using the notation of Robbins and Monro, $M(x)$ denotes the expected value at level x of the response, say Y, of a certain experiment. $M(x)$ is assumed to be a continuous monotone function of x, but is unknown to the experimenter, and it is desired to find the solution $X = \theta$ of the equation $M(x) = \alpha$ where α is a given constant. The Robbins and Monro method is one in which successive experiments are performed at levels X_1, X_2, ... in such a way that X_j will tend to θ in probability.

Except for an unpublished study by Teichrow (9) and an application of the Robbins and Monro technique described by Louis and Ruth Guttman (10), little is available to the experimenter to guide him in the use of stochastic approximation methods.

The thesis is divided into two parts. The first is a discussion and description of three stochastic approximation estimators and an empirical comparison of the convergence properties of the three estimators. The second part (Chapter 7) is a formulation of a method to determine confidence limits on the solution of (1) when $\alpha = 0.50$ using theory

developed for finite Markov Chains.

Since the form of M(x) is not known to the experimenter, the means used here to study the convergence properties is to employ a Monte Carlo sampling scheme to simulate a test in which stochastic approximation methods will be used. Upon repeated simulations of trials for various forms of M(x), various convergence properties of each of the three estimators can be observed.

The primary interest here lies in sensitivity testing, sometimes called quantal response testing; therefore, the empirical study made is a simulation of this type of testing. A similar study could be made by assigning a continous distribution function to the observed random variable Y(x).

CHAPTER II

THREE STOCHASTIC APPROXIMATION ESTIMATORS

For each real number x, let $Y(x)$ be a random variable such that $E[Y(x)] = M(x)$ exists. Assume that the regression equation $M(x) = \alpha$ has a single root at $x = \theta$, which is to be estimated, and that $(x - \theta)[M(x) - \alpha] > 0$ for all $x \neq \theta$. An initial value x_1 and a sequence $[c_j]$ of positive numbers are selected. The $(j+1)$st approximation to θ is defined inductively by the recursive formula

$$x_{j+1} = x_j + c_j(\alpha - y_j) \tag{2}$$

where y_j is the observed value of the random variable at $x = x_j$. The letter j denotes the trial number.

Each of the three estimators can be written in the form of equation (2). However, the difference lies in the way the sequence $[c_j]$ is defined. The proofs that estimators I and II converge with probability one to the desired value, $x = \theta$, are available in statistical literature and will not be discussed here.

The sequence $[c_j]$ which defines estimator I (the Robbins-Monro estimator) is a fixed sequence of positive elements with the following properties:

$$\text{(a)} \qquad \sum_{j=1}^{\infty} c_j = \infty$$

$$\text{(b)} \qquad \sum_{j=1}^{\infty} c_j^2 < \infty$$

The sequence $[1/j]$ has these properties.

The second estimator (estimator II proposed by Keston) is defined by equation (2) where the sequence $[c_j]$ is defined in the following way from the sequence

$$c_1 = a_1$$

$$c_2 = a_2$$

$$c_j = a_{t(j)}$$

where $$t(j) = 2 + \sum_{i=3}^{j} \delta\,[(x_i - x_{i-1})(x_{i-1} - x_{i-2})]$$

and $$\delta(x) = 1 \quad \text{if } x \leq 0$$

$$= 0 \quad \text{if } x > 0$$

Thus every time $(x_j - x_{j-1})$ differs in sign from $(x_{j-1} - x_{j-2})$, another a_k is taken. A further restriction on the sequence $[a_k]$ other than the properties (a) and (b) is

(c) $$a_{k+1} \leq a_k$$

It is important to note that the elements of $[c_j]$ for $j > 2$ are random variables.

Keston's rule for selecting the members of $[c_j]$ is based on the conjecture that in the neighborhood of $x = \theta$, θ being the solution of equation (1), it seemed likely that frequent fluctuations in the sign of $(x_j - \theta) - (x_{j+1} - \theta) = x_j - x_{j+1}$ indicate that $|x_j - \theta|$ is small where a few fluctuations in the sign of $x_j - x_{j+1}$ indicate that x_j is far away from θ.

It can be shown that there exists a θ', not necessarily identical with θ, where fluctuations in the sign occur more frequently in a finite number of trials. The value $x = \theta'$ is defined by the intersection of the

line $Y(x) = a$ and the locus of the medians of the densities $dH(y \mid x) / dy$ for any x. It should be noted that if the density $dH(y \mid x) / dy$ is symmetric, then Keston's conjecture is obviously correct. Even though the fluctuation would be expected to occur at θ' instead of θ, this does not affect the convergence in probability of

$$x_{j+1} = x_j + c_j (a - y_j) \tag{3}$$

to θ, as Keston has proved.

Let x_j be the value such that the variation in the algebraic sign of $x_j - x_{j+1}$ is maximum. Suppose that $x_{j-1} < x_j$. In order for a variation in the sign to occur, $x_{j+1} < x_j$; where x_{j+1} is defined by equation (3).

Let U denote a random variable whose density is the point binomial. The variable U takes on the value unity with the probability P_x where

$$P_x = Pr[X_{j+1} < x_j \mid x_{j-1} < x_j] \tag{4}$$

From equation (3), it follows that

$$P_x = Pr[Y(x_j) > a] \tag{5}$$

Clearly, U has maximum variance at $P_x = 1/2$. Therefore, that value of x such that

$$Pr[Y(x_j) > a] = 1/2 \tag{6}$$

is the desired value of θ'.

If $x_{j-1} > x_j$, a similar argument leads to the conclusion that the value of x such that

$$Pr[Y(x_j) < a] = 1/2 \tag{7}$$

is the desired θ'. Hence, θ' is the value of x defined by the intersection of the line $M(x) = a$ and the locus of the medians of $dH(Y \mid x) /dy$.

Since the sequence $[x_j]$ converges to θ with probability one, there

exists a J such that for all j > J

$$\Pr\left[\sup_{x_j} |X_j - \theta| < |\theta - \theta'|\right] = 1 - e \qquad \theta' \neq \theta \text{ and } e > 0$$

That is, there exists a neighborhood of θ which does not contain θ' such that after some trial number N almost surely all x_j will lie inside the neighborhood. Hence, there will exist almost surely only a finite number of sign changes in a neighborhood of θ' if θ is not in the neighborhood of θ'. But, for a finite number of trials, the experimenter cannot be assured that the sign changes are occurring in the neighborhood of θ or θ'.

The terminology "almost surely" is used to denote that the probability that an event occurring is one except for a set of probability measure zero.

In order to obtain an indication of how this fact would affect the sequence $[c_j]$, consider the difference between the median and means of two rather common skewed densities: the triangular and the gamma.

Consider first the following form of the triangular distribution:

$$f(x) = \begin{cases} \frac{2}{cb}\, x & 0 \leq x \leq b \\ \\ \frac{2}{c(c-b)}\,(c - x) & b \leq x \leq c \end{cases}$$

Table I presents values of the ratio of the median to c, the ratio of the mean to c, and their difference for various values of b / c. Note that for small values of b / c, the difference between the median and the mean can be slight.

Table II presents the ratio of the median to β, the ratio of the mean to β, and their difference for various values of α, when the

gamma density is of the following form:

$$f(x) = \frac{1}{\beta^{\alpha+1}\,\Gamma(\alpha+1)}\, x^{\alpha} e^{-x/\beta} \qquad x > 0$$

From the data in Table II, it appears that even for small values of β, the difference between the median and the mean can be relatively large.

It should be noted that the mean and the median are identical in the binomial distribution if, and only if, $p = 1$, 0, or $1/2$ where $p + q = 1$. The importance of the binomial distribution is that it is the basic distribution for quantal response problems.

It is hard to justify the use of an estimator computed from a small number of trials simply because it is known to converge to the desired value as the number of trials increases without bound. The fact that no other estimators have been proposed and found better, in some sense could be a just reason for using the stochastic approximation estimator. Therefore, it seems desirable to compare the two stochastic approximation estimators previously described with an estimator (estimator III) which seems to be the one which would be most naturally proposed by an experimenter who had no knowledge of the Robbins-Monro or the Keston estimators.

An experimenter who wishes to determine an x such that $M(x) = \alpha$ would most logically select an x_1 which he would consider as being close to the desired value and then compare the random variable $Y(x_1)$ with α.

If $Y(x_1)$ exceeded α, then $x_2 < x_1$ would be selected according to the magnitude of $\alpha - Y(x_1)$. Similarly, if $Y(x_1)$ were less than α,

TABLE I

COMPARISON OF THE MEAN AND MEDIAN FOR THE TRIANGULAR DENSITY FUNCTION

b/c	Median/c	Mean/c	Difference/c
.5	.500	.500	.000
.6	.548	.533	.015
.7	.592	.567	.025
.8	.632	.600	.032
.9	.671	.633	.038
1.0	.707	.667	.040

TABLE II

COMPARISON OF THE MEAN AND MEDIAN FOR THE GAMMA DENSITY FUNCTION

α	Median/β	Mean/β	Difference/β
0	.693	1.000	.307
1	1.678	2.000	.322
2	2.674	3.000	.326
3	3.672	4.000	.328
4	4.671	5.000	.329
5	5.670	6.000	.330
6	6.670	7.000	.330
7	7.669	8.000	.331
8	8.669	9.000	.331
9	9.669	10.000	.331
10	10.669	11.000	.331

the experimenter would continue testing at x_1. If after j tests $Y(x_{j-1}) < \alpha$ and $Y(x_j) > \alpha$ or $Y(x_{j-1}) > \alpha$ and $Y(x_j) < \alpha$, then it seems logical that the experimenter would interpolate in order to obtain x_{j+1}. Also, it seems a desirable procedure to shorten the steps that one takes after each trial in a small neighborhood of the desired value of x. A modification of Keston's procedure for shortening the step length seems intuitively adequate.

Mathematically, this procedure can be described by the recursive formula, equation (2), where c_j is an element of a sequence $[c_j]$ defined by the following rule:

$$c_1 = a_1$$

$$c_2 = a_2$$

If $c_{j-1} = a_k$ for $k \geq 2$, then

$$c_j = \begin{cases} a_k & \text{when } \alpha \notin (y_j, y_{j-1}) \\ (x_j - x_{j-1}) / (y_j - y_{j-1}) & \text{when } \alpha \in (y_j, y_{j-1}) \end{cases}$$

$$c_{j+1} = \begin{cases} a_k & \text{when } c_j = a_k \text{ and } \alpha \notin (y_{j+1}, y_j) \\ (x_{j+1} - x_j) / (y_{j+1} - y_j) & \text{when } \alpha \in (y_{j+1}, y_j) \\ a_{k+1} & \text{when } \alpha \notin (y_{j+1}, y_j) \\ & \text{and } c_j = (x_j - x_{j-1}) / (y_j - y_{j-1}) \end{cases}$$

where a_k is an element of a sequence $[a_k]$ having the following properties:

(a) $a_k > 0$ for $k = 1, 2, \ldots$

(b) $a_k > a_{k+1}$ for $k = 1, 2, \ldots$

(c) $\sum_{1}^{\infty} a_k = \infty$

(d) $\sum_{1}^{\infty} a_k^2 < \infty$

That is, if $\alpha \in (y_j, y_{j-1})$, then x_{j+1} is obtained by linear interpolation. A new a_k is selected after each period of linear interpolation. An end of a period occurs if $\alpha \in (y_j, y_{j-1})$ but $\alpha \notin (y_{j+1}, y_j)$; hence, c_{j+1} is the next unused element of the sequence $[a_k]$ The symbol $\in$ denotes "contained between." while $\notin$ denotes "is not contained between."

CHAPTER III

APPLICATION OF STOCHASTIC APPROXIMATION METHODS TO QUANTAL RESPONSE PROBLEMS

Let the random variable Y take on only two values, unity with the probability M(x) and zero with the probability 1 - M(x). This type of a response has been called quantal response. Let there be two real numbers, a and b (a < b), such that

$$Y(x) = 0 \qquad \text{for all } x \leq a$$

and

$$Y(x) = 1 \qquad \text{for all } x \geq b$$

Assume that a = 0 and b = 1. Then the regression function M(x) will have the following properties:

(8)
$$M(x) = 0 \qquad \text{for } x \leq 0$$
$$= f(x) \qquad \text{for } 0 \leq x \leq 1$$
$$= 1 \qquad \text{for } x \geq 1$$

In a neighborhood of $x = \theta$, the root of the regression equation $M(x) = \alpha$, we know that there exists some neighborhood of θ in which

(9) $\Pr[\ x_{j+1} - \theta \geq x_j - \theta \text{ and } (x_{j+1} - \theta)(x_j - \theta) \geq 0\ ]$,

the probability of making an incorrect decision at x_j, is an increasing function of x as x tends toward θ.

Since $Y(x_j)$ can take on only the values of zero and unity, and assuming $\alpha \neq 1$ or $\alpha \neq 0$, then $\Pr[Y(x_j) = \alpha \mid x_j = \theta] = 0$, or the value of the probability statement 9 is unity.

Suppose, however, that at each level x_j a sample of $k > 1$ Y's are taken. Since the sample mean

$$\overline{Y}(x_j) = \frac{1}{k} \sum_{1}^{k} Y_i(x_j) \qquad \text{where } Y_i = \begin{cases} 0 \text{ if no response occurs} \\ 1 \text{ if a response occurs} \end{cases}$$

has the same expected value as the random variable Y(x), the recursive formula $x_{j+1} = x_j + c_j \left[\alpha - \overline{y}(x_j) \right]$ will also converge in probability to θ for estimators I and II.

Let us consider a special application of the general stochastic approximation technique, that is, the problem to which stochastic approximations would be most applicable: the quantal response problem or sensitivity testing. This is a test in which the experimenter wants to determine a level of x such that the probability of a response as defined by the problem will be some preassigned value, say α. Let M(x) be defined by equation (8) where f(x) is monotonically increasing in its range. Let us now consider the upper and lower tolerance equations, L_1 and L_2, respectively, such that $1 - 2\gamma$ percent of the observed $\overline{Y}(x)$ will be expected to fall between them. Let us represent these by $L_1(x)$ and $L_2(x)$.

In order to show that $L_1(x)$ and $L_2(x)$ are monotone increasing functions consider the following argument: We know that

$$\Pr[Y = 1] = M(x) \qquad \text{for each } x$$

therefore

$$\Pr\left[\sum_{i=1}^{k} Y_i \geq kL_1 \mid M(x_1) \right] = \gamma$$

and if $x_2 > x_1$, then $M(x_2) > M(x_1)$ and

$$\Pr\left[\sum_{i=1}^{k} Y_i \geq kL_1(x_1) \mid M(x_2) \right] \geq \gamma$$

Now selecting $kL_1(x_2)$ such that

$$\Pr\left[\sum_{i=1}^{k} Y_i \geq kL_1(x_2) \mid M(x_2)\right] = \gamma$$

implies that $L_1(x_2) > L_1(x_1)$.

Hence, we can conclude that if $x_2 > x_1$ then $L_1(x_2) \geq L_1(x_1)$ and the tolerance function is non-decreasing. It follows similarly that $L_2(x)$ is also a non-decreasing function. The ranges of $L_1(x)$ and $L_2(x)$ are from zero to unity.

In order to gain further insight, consider Fig. 1. A desirable quality of a test would be conditions such that the length of the interval $I(\theta) = [x(L_1), x(L_2)]$ be minimized. The length of $I(\theta)$ depends upon slope and curvature of $f(x)$ in the neighborhood of θ and the distribution function of $\overline{Y}$, say $G(\overline{y} \mid x)$.

Since $k\overline{Y}$ is distributed as

$$\binom{k}{k\overline{y}} M(x)^{k\overline{y}} [1 - M(x)]^{k-k\overline{y}}$$

increasing the sample size k decreases the variance

$$\mathrm{Var}(\overline{Y} \mid x) = \frac{M(x)[1 - M(x)]}{k}$$

We note that $\lim_{k\to\infty} I(\theta) = 0$ and that the density $g(\overline{y} \mid x)$ becomes symmetric as k increases. Hence, for large samples, we are assured that as the trials proceed we will move toward θ with a probability of at least $1 - \gamma$ at each trial when $x \notin I(\theta)$. It is only in those trials at levels of x which are contained in $I(\theta)$ that the probability the next step will be toward θ is less than $1 - \gamma$.

Figure 1 illustrates that each sample size fixes the tolerance

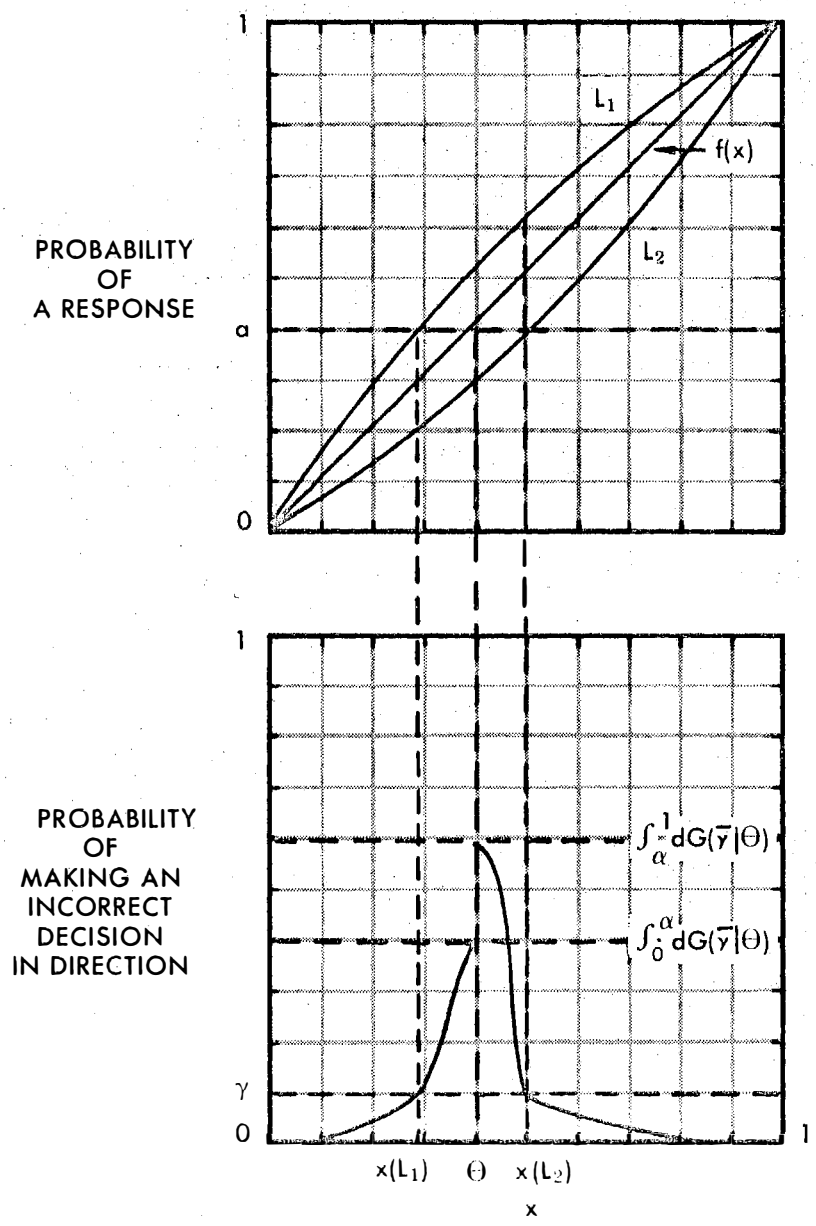

Fig. 1. The Regression Function and the Associated Curve Illustrating the Probability of Making an Incorrect Decision in Direction.

equations $L_1(x)$ and $L_2(x)$. Note that the probability of moving toward θ at each x_j exceeds or is equal to $1 - \gamma$ if $x \notin I(\theta)$. Since cost and sample size are usually directly related, it would be desirable to minimize k, the sample size. If $|x_j - \theta|$ is relatively large, a small sample size seems to be desirable. When $|x_j - \theta|$ is relatively small, a larger sample size requires the length of $I(\theta)$ to decrease and the likelihood that $x \notin I(\theta)$ to increase for a given probability level.

The effect of increasing sample size with number of trials has been studied empirically. (See Tables IV and VIII.)

CHAPTER IV

THE MONTE CARLO SAMPLING PLAN TO STUDY THE RATES OF CONVERGENCE OF THE ESTIMATORS

Due to the number of uncontrollable parameters involved, perhaps the most practical means available at this time to study convergence properties of the three estimators is a Monte Carlo procedure. The procedure used is as follows:

1. Define $M(x)$, a, Δk, and k, where k is the size of the sample taken at each level of x, and Δk is an increment which will be added to k with increasing trials.

2. Letting $x_1 = a$, compute $M(x_1)$.

3. Generate k random numbers (r_i, $i = 1, 2 \ldots k$) from a uniform density.

4. Compare each random number r_i with $M(x_1)$. If $r_i > M(x_1)$, assign the value of zero to Y_i. If $r_i \leq M(x_1)$, assign the value of unity to Y_i.

5. Compute $y_1 = \frac{1}{k} \sum_{i=1}^{k} y_i$

6. Substitute $\bar{y}_1$ into the recursive formula to determine x_2.

7. If $(x_j - x_{j-1})(x_{j-1} - x_{j-2}) < 0$, an increment of Δk is added to the sample size.

This procedure was programmed for the IBM 704 and continued for a desired number of trials.

In the study, each test was composed of a simulation of forty-nine trials. Each test was repeated one hundred times. Average value of 100 trials for x_7, x_{14}, x_{21}, x_{28}, x_{35}, x_{42}, and x_{49} were tabulated (Tables IV-VIII) for various values of α, k, and Δk.

In practice, the form of M(x) is unknown to the experimenter, but it was necessary to define the form of M(x) to perform the sampling plan. In this study, five forms of M(x) were selected in order for a relatively complete grid to be placed over the unit square. A sketch of these three forms are shown in Figure 2.

These were

$$M_1(x) = x^{1/4} \qquad 0 \leq x \leq 1$$

$$M_2(x) = \begin{cases} 4x^2 & 0 \leq x \leq 1/4 \\ 1 - 4(1 - x)^2/3 & 1/4 \leq x \leq 1 \end{cases}$$

$$M_3(x) = \begin{cases} 2x^2 & 0 \leq x \leq 1/2 \\ 1 - 2(1 - x)^2 & 1/2 \leq x \leq 1 \end{cases}$$

$$M_4(x) = \begin{cases} 4x^2/3 & 0 \leq x \leq 3/4 \\ 1 - 4(1 - x)^2 & 3/4 \leq x \leq 1 \end{cases}$$

$$M_5(x) = x^4 \qquad 0 \leq x \leq 1$$

The form of $dH(y \mid x) / dy$ is defined by the quantal response property as the point binomial.

The values of α considered here with their associated θ_i for i = 1, 2, 3, 4, 5, where θ_i is the x value of the intersection of $M_i(x) = \alpha$ and $M_i(x)$, are tabulated in Table III.

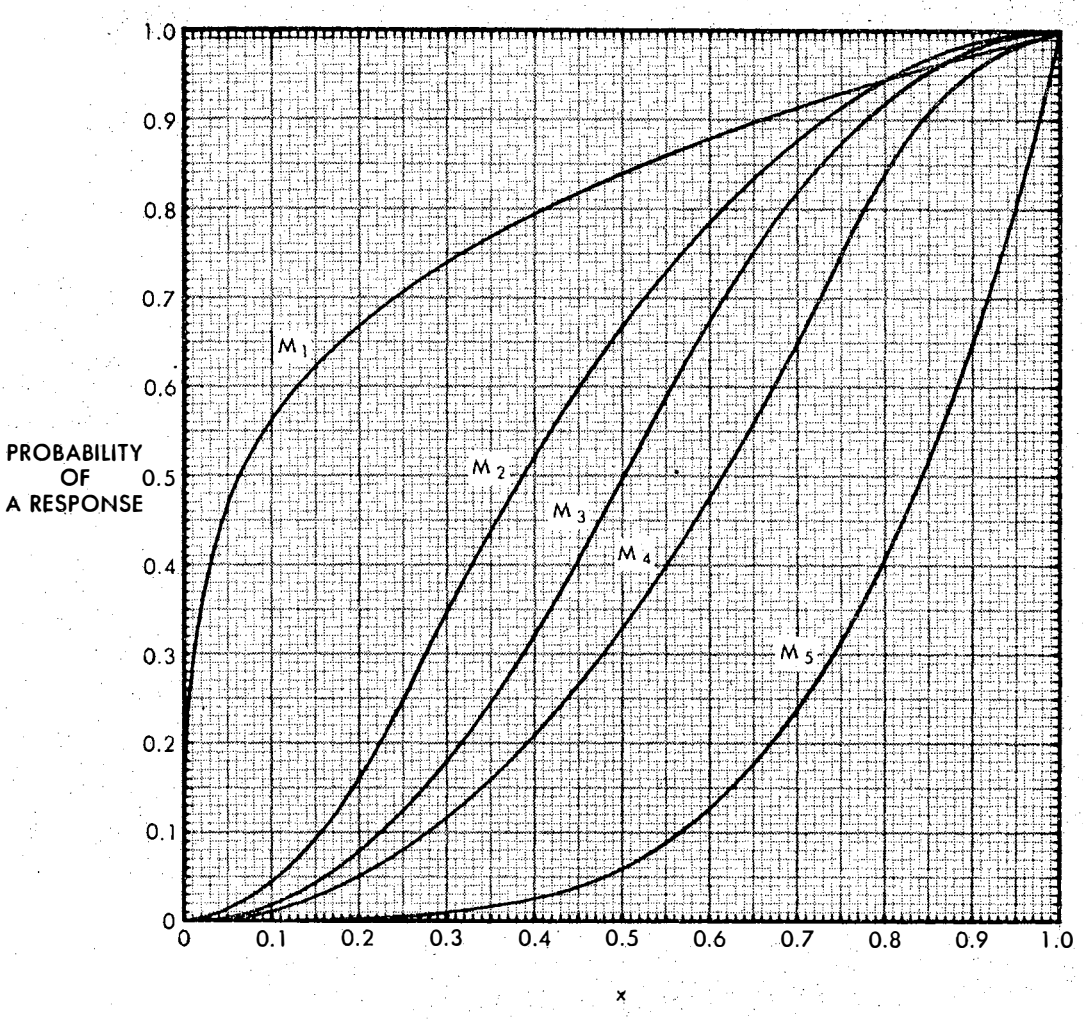

Fig. 2. The Five Forms of M(x).

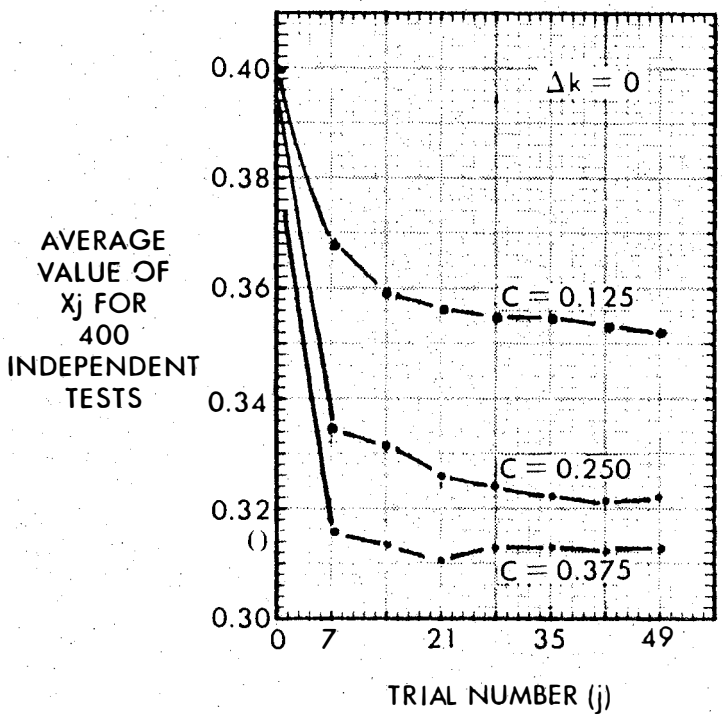

Fig. 3. The Effect of the Selection of [c/j] on the Rate of Convergence. (Estimator I)

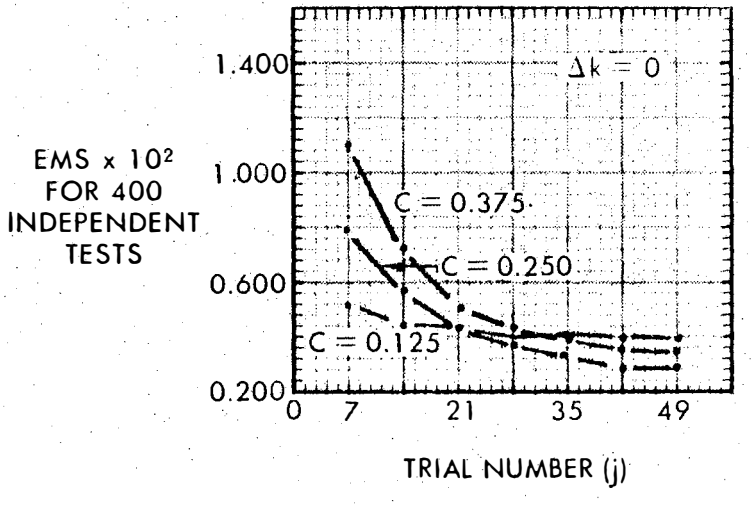

Fig. 4. The Variation of the Estimator for Various Values of c. (Estimator I)

TABLE III.

DATA FOR SAMPLING PROCEDURE

α	θ_1	θ_2	θ_3	θ_4	θ_5
.05	.00006	.11180	.15811	.19365	.47287
.10	.00010	.15811	.22361	.27386	.56234
.30	.00810	.27543	.38730	.47434	.74008
.50	.06250	.38763	.50000	.61237	.84090

Various sample sizes, ranging from one to twenty, were used in simulating the test. Also, a scheme in which the sample size increases by an increment of five as the number of trials increased was considered. When $(x_j - x_{j-1})(x_{j-1} - x_{j-2}) < 0$, the sample size was increased.

The sequence $[c_j]$ for the empirical study was $[c/j]$ where $c = 0.250$ and j the trial number. The choice of 0.250 is arbitrary and is not optimum for all forms of $M(x)$.

The selection of $c = 0.250$ was based on the data summarized in Fig. 3 and 4. Three choices of c($c = 0.125, 0.250, 0.375$) were studied empirically using estimator I. From Fig. 3, a "good" value of c in terms of minimum bias, in a sequence of form $[c/j]$, would be in the range of from 0.250 to 0.375. Figure 4 shows that the greater variability of the estimator for a small sample size for $c = 0.375$ may offset its value as an estimator even though it is associated with the minimum bias of the three cases studied here.

The results of the Monte Carlo simulation are tabulated in Tables IV - VII.

TABLE IV

AVERAGE x_j FOR 100 TRIALS FOR ESTIMATORS I, II, AND III, $\alpha = 0.05$, $\Delta k = 0$

Sample size (k)	Trial number (j)	θ_1 = .00006			θ_2 = .112			θ_3 = .158			θ_4 = .194			θ_5 = .473		
		I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
1	7	.084	.099	.092	.067	.087	.086	.077	.085	.091	.076	.091	.093	.081	.094	.094
	14	.080	.075	.064	.073	.086	.121	.084	.099	.130	.083	.122	.134	.090	.137	.137
	21	.075	.058	.037	.075	.073	.148	.088	.101	.164	.087	.141	.171	.095	.181	.181
	28	.072	.058	.016	.077	.069	.167	.090	.113	.188	.090	.146	.200	.099	.225	.224
	35	.070	.037	.005	.078	.074	.175	.092	.110	.208	.092	.148	.226	.101	.261	.268
	42	.068	.015	.007	.079	.081	.183	.093	.112	.224	.094	.152	.246	.104	.298	.310
	49	.066	.020	.009	.080	.087	.187	.095	.120	.231	.095	.153	.260	.106	.331	.351
5	7	.052	.050	.041	.073	.079	.085	.076	.086	.088	.078	.087	.091	.081	.094	.094
	14	.043	.034	.019	.078	.091	.105	.083	.106	.118	.085	.111	.123	.090	.137	.137
	21	.039	.016	.010	.081	.094	.114	.087	.117	.135	.089	.127	.147	.095	.180	.180
	28	.035	.010	.004	.082	.097	.120	.090	.123	.147	.092	.136	.164	.099	.220	.222
	35	.033	.004	.003	.083	.099	.125	.091	.127	.152	.094	.142	.174	.101	.258	.263
	42	.032	.005	.002	.085	.101	.127	.093	.130	.158	.096	.148	.182	.104	.291	.301
	49	.031	.004	.001	.085	.103	.128	.094	.132	.162	.097	.151	.187	.106	.318	.335
10	7	.041	.035	.032	.072	.076	.078	.075	.085	.085	.078	.086	.086	.081	.094	.094
	14	.034	.018	.009	.077	.087	.091	.082	.102	.106	.085	.109	.112	.090	.137	.137
	21	.030	.008	.003	.080	.091	.098	.085	.111	.118	.089	.120	.128	.095	.179	.179
	28	.027	.004	.002	.082	.094	.103	.088	.117	.125	.092	.128	.138	.099	.219	.219
	35	.025	.002	.001	.083	.096	.105	.090	.121	.130	.094	.132	.145	.101	.255	.256
	42	.023	.002	.001	.084	.097	.107	.091	.123	.134	.096	.136	.151	.104	.284	.288
	49	.022	.001	.000	.085	.098	.108	.093	.125	.137	.097	.140	.155	.106	.306	.314
20	7	.040	.036	.024	.072	.075	.078	.076	.082	.086	.078	.086	.089	.081	.094	.094
	14	.032	.016	.005	.077	.082	.093	.083	.097	.111	.085	.104	.120	.090	.137	.137
	21	.027	.006	.002	.080	.086	.100	.087	.103	.127	.089	.114	.142	.095	.180	.180
	28	.024	.002	.001	.081	.088	.105	.089	.108	.134	.092	.119	.155	.099	.221	.223
	35	.022	.002	.000	.083	.090	.107	.091	.111	.141	.094	.123	.163	.101	.256	.264
	42	.020	.001	.000	.084	.091	.108	.093	.113	.145	.096	.126	.169	.104	.280	.302
	49	.019	.001	.000	.085	.092	.110	.094	.115	.148	.098	.128	.175	.106	.296	.335

TABLE V

AVERAGE x_j FOR 100 TRIALS FOR 100 TRIALS FOR ESTIMATORS I, II, AND III WHEN $\alpha = .10$ AND $\Delta k = 0$

Sample size (k)	Trial number (j)	$\theta_1 = .0001$			$\theta_2 = .158$			$\theta_3 = .224$			$\theta_4 = .274$			$\theta_5 = .562$		
		I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
1	7	.042	.048	.018	.129	.128	.165	.149	.151	.168	.150	.174	.180	.161	.187	.187
	14	.036	.014	.012	.134	.117	.207	.157	.164	.227	.161	.212	.248	.180	.272	.275
	21	.031	.023	.001	.135	.121	.215	.161	.175	.267	.168	.212	.297	.190	.348	.359
	28	.027	.010	.007	.136	.130	.228	.164	.180	.286	.172	.220	.327	.197	.416	.440
	35	.024	.010	.009	.137	.138	.234	.166	.186	.295	.176	.224	.348	.203	.465	.508
	42	.021	.006	.008	.137	.139	.239	.168	.189	.306	.178	.228	.359	.207	.482	.556
	49	.020	.006	.006	.138	.140	.238	.170	.193	.311	.181	.239	.370	.211	.495	.588
5	7	.017	.015	.013	.129	.132	.134	.144	.151	.159	.151	.164	.163	.161	.186	.187
	14	.009	.004	.001	.135	.139	.148	.154	.168	.183	.163	.190	.200	.179	.269	.271
	21	.005	.002	.001	.138	.142	.153	.159	.179	.193	.170	.203	.219	.189	.339	.344
	28	.002	.001	.001	.139	.144	.155	.163	.184	.200	.174	.213	.230	.196	.393	.398
	35	.001	.000	.000	.141	.145	.157	.165	.188	.207	.178	.219	.237	.202	.422	.434
	42	.001	.000	.000	.141	.146	.157	.167	.191	.210	.180	.222	.242	.206	.438	.459
	49	.001	.000	.001	.142	.147	.158	.169	.194	.212	.183	.226	.247	.210	.451	.475
10	7	.007	.002	.002	.127	.132	.138	.143	.151	.159	.149	.161	.167	.161	.186	.187
	14	.002	.001	.001	.132	.137	.147	.154	.168	.189	.161	.182	.210	.179	.268	.272
	21	.001	.000	.000	.135	.139	.154	.159	.176	.202	.168	.191	.235	.189	.331	.350
	28	.000	.000	.000	.137	.141	.157	.163	.180	.209	.173	.197	.247	.197	.367	.417
	35	.000	.000	.000	.138	.143	.157	.166	.183	.215	.176	.201	.254	.202	.389	.468
	42	.000	.000	.000	.139	.144	.158	.168	.185	.218	.179	.205	.257	.207	.404	.500
	49	.010	.003	.000	.140	.145	.158	.170	.187	.219	.181	.207	.260	.210	.415	.520

TABLE VI

AVERAGE VALUE OF x_j FOR 100 TRIALS FOR ESTIMATORS I, II, AND III WHEN $\alpha = .30$ AND $\Delta k = 0$

Sample size (k)	Trial number (j)	$\theta_1 = .008$			$\theta_2 = .275$			$\theta_3 = .387$			$\theta_4 = .474$			$\theta_5 = .740$		
		I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
1	7	.100	.084	.080	.287	.284	.311	.367	.348	.371	.396	.397	.435	.472	.520	.540
	14	.074	.050	.058	.277	.272	.336	.369	.354	.402	.409	.423	.476	.514	.635	.681
	21	.060	.035	.045	.279	.274	.336	.371	.361	.415	.415	.432	.481	.538	.666	.721
	28	.049	.028	.044	.278	.275	.326	.374	.364	.420	.421	.439	.500	.553	.683	.746
	35	.045	.024	.041	.279	.276	.328	.375	.365	.422	.424	.441	.513	.564	.695	.753
	42	.040	.019	.035	.279	.275	.329	.375	.370	.423	.425	.443	.522	.573	.703	.756
	49	.036	.019	.033	.279	.275	.328	.376	.374	.429	.426	.444	.526	.580	.712	.759
5	7	.081	.048	.044	.280	.279	.287	.352	.356	.356	.385	.396	.404	.470	.533	.536
	14	.046	.019	.018	.278	.276	.283	.360	.366	.376	.403	.424	.436	.515	.651	.662
	21	.033	.014	.015	.278	.276	.281	.364	.373	.381	.412	.435	.447	.538	.681	.703
	28	.026	.012	.013	.277	.276	.283	.366	.376	.382	.415	.442	.454	.554	.698	.714
	35	.020	.011	.012	.277	.275	.285	.368	.377	.383	.420	.446	.459	.565	.707	.721
	42	.018	.010	.012	.277	.275	.282	.370	.378	.384	.423	.448	.460	.574	.712	.728
	49	.016	.010	.012	.277	.274	.281	.370	.380	.385	.425	.452	.463	.582	.715	.731
10	7	.076	.044	.038	.281	.283	.280	.358	.362	.363	.391	.397	.410	.471	.533	.535
	14	.041	.016	.014	.280	.280	.282	.365	.370	.379	.408	.421	.445	.515	.661	.669
	21	.029	.012	.013	.279	.280	.276	.369	.374	.378	.415	.431	.457	.538	.692	.712
	28	.022	.010	.012	.278	.279	.277	.370	.377	.382	.420	.437	.461	.554	.704	.725
	35	.019	.009	.010	.278	.278	.275	.372	.378	.386	.424	.440	.465	.565	.710	.730
	42	.016	.009	.010	.278	.278	.275	.373	.379	.387	.427	.443	.468	.574	.713	.732
	49	.014	.009	.010	.277	.277	.276	.373	.380	.389	.429	.445	.467	.581	.716	.733

TABLE VII

AVERAGE VALUE OF x_j FOR 100 TRIALS FOR ESTIMATORS I, II, AND III WHEN $a = .50$ AND $\Delta k = 0$

Sample size (k)	Trial number (j)	$\theta_1 = .062$			$\theta_2 = .388$			$\theta_3 = .500$			$\theta_4 = .612$			$\theta_5 = .841$		
		I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
1	7	.300	.281	.279	.436	.435	.428	.511	.504	.510	.569	.569	.596	.721	.739	.728
	14	.261	.211	.187	.429	.422	.402	.505	.501	.497	.580	.576	.598	.753	.787	.790
	21	.241	.182	.153	.426	.413	.412	.504	.500	.494	.584	.586	.602	.766	.810	.811
	28	.226	.168	.126	.424	.404	.403	.501	.500	.493	.585	.589	.606	.776	.820	.818
	35	.216	.155	.106	.421	.402	.400	.500	.500	.498	.588	.592	.607	.782	.822	.820
	42	.207	.144	.099	.420	.400	.394	.500	.503	.499	.590	.593	.608	.787	.824	.827
	49	.202	.138	.098	.416	.398	.394	.501	.502	.497	.591	.593	.609	.791	.826	.829
5	7	.317	.244	.258	.425	.413	.428	.504	.504	.494	.573	.578	.586	.721	.769	.770
	14	.274	.154	.156	.416	.403	.410	.503	.504	.495	.584	.593	.599	.758	.813	.821
	21	.251	.124	.120	.411	.399	.404	.503	.503	.496	.590	.599	.603	.774	.825	.830
	28	.237	.110	.098	.408	.397	.398	.503	.501	.498	.593	.602	.605	.784	.829	.834
	35	.226	.100	.088	.406	.395	.396	.502	.500	.499	.594	.603	.607	.790	.832	.836
	42	.218	.093	.084	.404	.394	.394	.502	.500	.499	.596	.603	.607	.795	.834	.838
	49	.211	.089	.080	.404	.394	.393	.502	.500	.499	.597	.604	.609	.798	.834	.839
10	7	.316	.255	.266	.425	.424	.411	.499	.502	.499	.573	.577	.585	.722	.770	.770
	14	.273	.159	.135	.416	.412	.403	.499	.500	.499	.584	.588	.600	.759	.810	.827
	21	.250	.128	.096	.413	.407	.396	.500	.499	.501	.587	.593	.605	.775	.821	.831
	28	.236	.114	.079	.410	.403	.392	.499	.500	.503	.590	.597	.606	.786	.825	.834
	35	.224	.105	.073	.408	.401	.390	.499	.499	.502	.592	.599	.606	.793	.829	.835
	42	.216	.100	.070	.407	.400	.391	.500	.499	.502	.593	.600	.608	.798	.831	.838
	49	.209	.096	.068	.406	.399	.392	.500	.499	.500	.595	.602	.608	.801	.832	.839

TABLE VIII

AVERAGE x_j FOR 100 TRIALS FOR ESTIMATORS I, II, AND III, k=1, Δk=5

α	Trial Number (j)	θ_1 = .00006			θ_2 = .112			θ_3 = .158			θ_4 = .194			θ_5 = .473		
		I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
.05	7	.095	.105	.081	.073	.077	.087	.077	.087	.092	.078	.086	.090	.081	.094	.094
	14	.091	.080	.056	.078	.078	.120	.085	.116	.129	.084	.115	.130	.090	.137	.137
	21	.088	.057	.034	.081	.069	.137	.089	.129	.158	.089	.140	.166	.095	.181	.181
	28	.085	.042	.012	.082	.067	.141	.092	.126	.176	.091	.159	.193	.099	.225	.224
	35	.083	.026	.004	.083	.069	.135	.094	.118	.179	.093	.154	.208	.101	.267	.268
	42	.082	.010	.001	.084	.077	.128	.095	.119	.176	.095	.143	.212	.104	.307	.311
	49	.080	.007	.001	.085	.083	.125	.096	.122	.172	.096	.143	.213	.106	.334	.352
		θ_1 = .0001			θ_2 = .158			θ_3 = .224			θ_4 = .274			θ_5 = .562		
.10	7	.044	.040	.026	.134	.137	.147	.149	.147	.177	.147	.169	.180	.161	.187	.187
	14	.037	.012	.004	.136	.134	.144	.159	.166	.228	.158	.209	.247	.180	.272	.273
	21	.031	.002	.000	.139	.131	.167	.164	.164	.248	.165	.217	.277	.190	.349	.356
	28	.028	.000	.000	.139	.137	.166	.167	.177	.243	.170	.224	.287	.197	.412	.433
	35	.025	.000	.000	.140	.142	.163	.169	.186	.237	.173	.234	.286	.203	.461	.495
	42	.022	.000	.000	.141	.145	.162	.171	.193	.233	.176	.238	.282	.207	.470	.530
	49	.020	.000	.000	.141	.147	.161	.173	.197	.231	.178	.242	.280	.211	.484	.541
		θ_1 = .008			θ_2 = .275			θ_3 = .387			θ_4 = .474			θ_5 = .740		
.30	7	.082	.082	.063	.272	.280	.290	.363	.351	.391	.389	.380	.425	.475	.522	.544
	14	.051	.025	.018	.273	.274	.286	.366	.362	.391	.403	.412	.448	.519	.619	.672
	21	.038	.012	.010	.273	.274	.282	.369	.369	.388	.411	.430	.455	.542	.666	.703
	28	.031	.009	.009	.273	.274	.279	.370	.373	.387	.416	.440	.460	.558	.696	.717
	35	.027	.009	.009	.273	.275	.278	.371	.375	.386	.420	.447	.463	.569	.711	.725
	42	.024	.008	.008	.274	.275	.277	.372	.378	.387	.422	.452	.465	.578	.719	.729
	49	.021	.008	.008	.274	.275	.276	.373	.379	.387	.425	.455	.466	.585	.724	.731
		θ_1 = .062			θ_2 = .388			θ_3 = .500			θ_4 = .612			θ_5 = .841		
.50	7	.301	.278	.263	.433	.435	.419	.497	.488	.497	.565	.570	.579	.713	.729	.772
	14	.259	.204	.168	.422	.419	.399	.499	.492	.500	.576	.586	.596	.749	.797	.821
	21	.240	.159	.115	.417	.413	.394	.499	.494	.501	.581	.595	.603	.766	.818	.832
	28	.225	.131	.093	.415	.408	.392	.499	.496	.502	.584	.598	.606	.777	.826	.835
	35	.215	.115	.082	.412	.405	.390	.499	.496	.501	.586	.601	.607	.784	.830	.838
	42	.207	.104	.077	.410	.403	.389	.499	.497	.501	.588	.603	.609	.789	.833	.838
	49	.201	.097	.072	.409	.401	.389	.499	.498	.500	.590	.604	.609	.793	.834	.839

CHAPTER V

PROPERTIES OF THE SEQUENCE $[c_j]$ ASSOCIATED WITH ESTIMATOR III

Let $H(y \mid x)$ be a family of distribution functions depending on the real parameter x, and let

$$M(x) = \int_{-\infty}^{+\infty} y\,dH(y \mid x) \tag{10}$$

be the corresponding regression function. It is assumed that $M(x)$ is unknown to the experimenter, who is, however, allowed to take observations on $H(y \mid x)$ for any value of x.

The recursive formula

$$x_{j+1} = x_j + c_j(a - y_j) \tag{11}$$

defines a sequence $[x_j]$ which in the limit would be desirable to converge with probability one to θ, which is a root of the equation

$$M(x) = a \tag{12}$$

The value c_j is an element of a sequence defined by the following rule:

$$c_1 = a_1 \tag{13}$$

$$c_2 = a_2$$

If $c_{j-1} = a_k$ for $k \geq 2$, then

$$c_j = \begin{cases} a_k & \text{when } \alpha \notin (y_j, y_{j-1}) \\ (x_{j-1} - x_j) \,/\, (y_{j-1} - y_j) & \text{when } \alpha \in (y_j, y_{j-1}) \end{cases}$$

$$c_{j+1} = \begin{cases} a_k & \text{when } c_j = a_k \text{ and } \alpha \notin (y_{j+1}, y_j) \\ (x_{j+1} - x_j)/(y_{j+1} - y_j) & \text{when } \alpha \in (y_{j+1}, y_j) \\ a_{k+1} & \text{when } \alpha \notin (y_{j+1}, y_j) \text{ and} \\ & c_j = (x_j - x_{j-1})/(y_j - y_{j-1}) \end{cases}$$

Where a_k is an element of a sequence, $[a_k]$, having the following properties:

(a) $a_k > 0$ for $k = 1, 2, 3, \ldots$

(b) $a_k > a_{k+1}$ for $k = 1, 2, 3, \ldots$

(c) $\sum_1^{\infty} a_j = \infty$

(d) $\sum_1^{\infty} a_j^2 < \infty$

The symbol ε in the quantity $\alpha \varepsilon (y_{j+1}, y_j)$ denotes between, that is, α is contained in the interval $[y_{j+1}, y_j]$ if $y_{j+1} < y_j$, or in $[y_j, y_{j+1}]$ if $y_j < y_{j+1}$.

It is assumed that M(x) is a non-decreasing continuous function and H(y | x) is such that

$$\Pr[Y > \alpha \mid x < \theta] < \Pr[Y > \alpha \mid x = \theta]$$

and $$\Pr[Y > \alpha \mid x > \theta] < \Pr[Y > \alpha \mid x = \theta]$$

These conditions and the restrictions listed below are the only restrictions placed on M(x) and H(y | x).

(a) $M(x) \leq c + |d| x$ c and d are real constants

(b) $\int_{-\infty}^{\infty} |y - M(x)|^2 dH(y \mid x) \leq \sigma^2 < \infty$

(c) $M(x) < \alpha$ for $x < \theta$, $M(x) > \alpha$ for $x > \theta$

(d) $\inf_{\delta_1 \leq |x-\theta| \leq \delta_2} |M(x) - \alpha| > 0$ for every pair of numbers (δ_1, δ_2) with $0 < \delta_1 < \delta_2 < \infty$

The properties of the sequence $[c_j]$ will be presented in the form of seven lemmas and a single theorem.

Lemma 5.1. If the elements c_k and c_{k-1} of the sequence $[c_j]$ are such that $c_{k-1} \in [a_j]$ and $c_k = (x_k - x_{k-1}) / (y_k - y_{k-1})$, then $0 < c_k \leq c_{k-1}$.

Proof: Since $c_{k-1} \in [a_j]$, $c_{k-1} > 0$. If $y_k < \alpha < y_{k-1}$, then $x_k < x_{k-1}$. Similarly, if $y_{k-1} < \alpha < y_k$, then $x_{k-1} < x_k$. It follows immediately that $c_k = (x_k - x_{k-1}) / (y_k - y_{k-1}) > 0$.

It remains to be proved that $c_k < c_{k-1}$. Since $x_k = x_{k-1} + c_{k-1}(\alpha - y_{k-1})$, we can write $c_k = c_{k-1}(\alpha - y_{k-1}) / (y_k - y_{k-1})$. Noting that both $y_k < \alpha < y_{k-1}$ and $y_{k-1} < \alpha < y_k$ imply that $0 < (\alpha - y_{k-1}) / (y_k - y_{k-1}) < 1$, it can be concluded that $c_k < c_{k-1}$. It should be noted that if $x_k < x_{k-1}$, then $y_{k-1} < y_k$ cannot be true. This follows immediately from the recursive formula, equation (11)

Lemma 5.2. For every k such that $c_k = (x_k - x_{k-1}) / (y_k - y_{k-1})$ and $c_{k+1} = (x_{k+1} - x_k) / (y_{k+1} - y_k)$, then $c_{k+1} < c_k$.

Proof: From the proof of Lemma 5.1, we know that

$c_{k+1} = c_k(a - y_k) / (y_{k+1} - y_k)$, and $c_k > 0$ and $0 < (a - y_k) / (y_{k+1} - y_k) < 1$, it follows that $c_{k+1} < c_k$.

It should be noted that in general c_{j+1} is not less than c_j for all $j = 1, 2, \ldots$.

Lemma 5.3. For each k and J the probability that $c_j = a_k$ for all $j \geq J$ is zero.

Proof: Let $c_j = a_k$ for all $j \geq J$, where $j = 1, 2, \ldots$. The sequence $[x_j]$ is monotone and converges to a finite value, say A, if the sequence is bounded, and diverges to either $-\infty$ or $+\infty$ if unbounded.

Let $[x_j]$ be non-increasing and bounded below by its limit A. Then for each $j > J$ there exists an $e_j > 0$ such that $x_j = A + e_j a_k$. The sequence $[e_j]$ is a non-increasing sequence of positive elements such that $\lim_{j\to\infty} e_j = 0$.

Then,

$$0 \leq x_{j+1} - A \leq e_j a_k \tag{14}$$

Simplifying,

$$0 \leq x_j - A + a_k(a - y_j) \leq e_j a_k$$

$$0 \leq e_j a_k + a_k(a - y_j) \leq e_j a_k$$

$$0 \leq e_j + (a - y_j) \leq e_j$$

$$a \leq y_j \leq a + e_j$$

Let us now consider the probability of such an event, that is, $\Pr[a \leq Y_j \leq a + e_j]$. If $H(y \mid x)$ is continuous, then, as $j\to\infty$ and $e_j \to 0$, $\Pr[a \leq Y_j \leq a + e_j] \to 0$. However, if $H(y \mid x)$ is discrete, $\Pr[a \leq Y_j \leq a + e_j]$ may not necessarily converge to zero as $e_j \to 0$. But $[a \leq Y_j < a + e_j]$ must hold for all j greater

than that one for which the inequality (14) holds. As e_j tends to zero, the probability of such an event is

$$\prod_{1}^{\infty} \Pr[Y_j = a] \leq \prod_{1}^{\infty} \max_{A \leq x \leq x_1} (\Pr[Y_j = a]) = 0$$

A similar argument holds when the sequence $[x_j]$ is non-decreasing and bounded.

Suppose the sequence $[x_j]$ is unbounded, then either $\lim_{j\to\infty} x_j = \infty$ or $\lim_{j\to\infty} x_j = -\infty$. In order for these events to occur, $y_j < a$ or $y_j > a$ for all $j > J$, respectively. Let us investigate the probability of such events, that is, $\Pr[Y_j > a, Y_{j+1} > a, \ldots]$ $= \Pr[\lim_{j\to\infty} x_j = -\infty]$ and $\Pr[Y_j < a, Y_{j+1} < a, \ldots] = \Pr[\lim_{j\to\infty} x_j = +\infty]$ Consider the latter of the two cases.

$$\Pr[Y_j < a, Y_{j+1} < a, \ldots] = \Pr[Y_j < a] \Pr[Y_{j+1} < a \mid Y_j < a] \ldots,$$
$$\Pr[Y_{j+L} < a \mid Y_j < a, \ldots, Y_{j+L-1} < a]$$
$$= \prod_{L}^{\infty} \Pr[Y_{j+L} < a]$$

There exists only a finite number of L such that $x < \theta$. It follows then that

$$\Pr[Y_j < a, Y_{j+1} < a, \ldots] \leq \prod_{L}^{\infty} \Pr[Y_j < a \mid x_j > \theta]$$
$$\leq \prod_{L}^{\infty} \Pr[Y < a \mid x = \theta]$$
$$= 0$$

A similar argument holds when $\lim_{j\to\infty} x_j = -\infty$, and the lemma is proved.

Lemma 5.4. If $c_j = (x_j - x_{j-1}) / (y_j - y_{j-1})$ for all $j > J$, then $\lim_{j\to\infty} (x_j - x_{j-1}) = 0$ almost surely is true for all c_j.

Proof: Suppose $x_{2j} > x_{2j-1}$ and $a < y_{2j}$. In order that c_j have the form restricted by the hypothesis of the lemma, $y_{2j-1} < a < y_{2j}$

for all $j \geq J$. The sequences $[x_{2j-1}]$ and $[x_{2j}]$ are monotone; the first is increasing, the second is decreasing. Since x_{j+1} is obtained by a linear interpolation between x_j and x_{j-1}, both sequences are bounded above and below. Let $\lim_{j\to\infty} x_{2j-1} = A$ and $\lim_{j\to\infty} x_{2j} = B$. Let $B - A = \Delta$, where $\Delta \geq 0$. Then for every $j > J$, there exists an $e_{2j-1} > 0$ such that $x_{2j-1} = A - e_{2j-1}$. The sequence $[e_{2j-1}]$ is monotonically decreasing and converges to zero. With each j there exists an e_{2j} such that $x_{2j} = B + e_{2j}$. The sequence $[e_{2j}]$ is monotonically decreasing and converges to zero as j increases without bound. Consider

$$(15) \qquad x_{2j+1} = x_{2j} + [(x_{2j} - x_{2j-1}) / (y_{2j} - y_{2j-1})](\alpha - y_{2j})$$

$$= B + e_{2j} + (B + e_{2j} - A + e_{2j-1}) \cdot [(\alpha - y_{2j}) / (y_{2j} - y_{2j-1})]$$

$$= B + \Delta[(\alpha - y_{2j}) / (y_{2j} - y_{2j-1})] + e_{2j} + (e_{2j} + e_{2j-1})[(\alpha - y_{2j}) / (y_{2j} - y_{2j-1})]$$

Taking the limit of both sides,

$$\lim_{j\to\infty} x_{2j+1} = B - \Delta[\lim_{j\to\infty}(y_{2j} - \alpha) / (y_{2j} - y_{2j-1})]$$

it is clear that

$$\Pr[\lim_{j\to\infty} (Y_{2j} - \alpha) / (Y_{2j} - Y_{2j-1}) = D] = 0 \text{ for any } D$$

Since the left side of equation (15) converges and $[\lim_{j\to\infty} (Y_{2j} - \alpha) / (Y_{2j} - Y_{2j-1})]$ almost surely does not exist, $\Delta = 0$, that is $A = B$. It follows immediately, then, that for almost all c_j $\lim_{j\to\infty} (x_j - x_{j-1}) = 0$, the desired result.

Lemma 5.5. Let the sequence $[z_k]$ be the union of all subsequences of $[Z_k]$ such that $\lim_{k\to\infty} z_k = \infty$, where Z_k is the number of times that

the kth element of $[a_k]$ appears in the sequence $[c_j]$. Then,

$$\Pr[\lim_{k\to\infty} z_k = \infty] = 0.$$

Proof: From Lemma 5.3, we know that for each k, Z_k is almost always finite. Since the sum of a denumerable number of sets of measure zero is also of measure zero, we can conclude that the probability of at least one element of the sequence of infinite terms in $[Z_k]$ being infinite is also zero. This still does not assure us that the sequence $[z_k]$ is almost always bounded.

Let $\lim_{k\to\infty} z_k = \infty$. Then for each $L > 0$, there must exist a k such that $z_k > L$. Consider the probability of such an event, that is

$$\Pr[z_k > L] = \Pr[Y_1 > a, \ldots, Y_L > a]$$

or $$\Pr[z_k > L] = \Pr[Y_1 < a, \ldots, Y_L < a]$$

But, from the proof of Lemma 5.3, we know

$$\lim_{L\to\infty} \Pr[Y_1 > a, \ldots, Y_L > a] = 0$$

or $$\lim_{L\to\infty} \Pr[Y_1 < a, \ldots, Y_L < a] = 0$$

Hence, we can conclude

$$\Pr[\lim_{L\to\infty} z_k = \infty] = 0$$

That is, the sequence $[z_k]$ is almost surely a bounded sequence.

Lemma 5.6. For every J, the probability that $c_j = (x_j - x_{j-1}) / (y_j - y_{j-1})$ for all $j \geq J$ is zero.

Proof: Suppose each element of the sequence $[c_j]$ takes on the form defined by the hypothesis of the lemma. Then the sequences $[x_{2j}]$ and $[x_{2j-1}]$ are monotonically decreasing and increasing sequences, respectively, when $y_{2j-1} < a < y_{2j}$ for all $j \geq J$. Similarly, the sequences are monotonically increasing and decreasing, respectively, if

$y_{2j} < \alpha < y_{2j-1}$ for all $j \geq J$.

By Lemma 5.4, we know that both these sequences converge to a common limit, A. Consider a neighborhood of A, say $\nu(A)$, such that at least one of the following probabilities is less than unity for all $x \varepsilon \nu(A)$: $\Pr[Y_{2j} > \alpha \mid x \varepsilon \nu(A)]$ and $\Pr[x_{2j-1} < \alpha \mid x \varepsilon \nu(A)]$. The existence of $\nu(A)$ is assured by the continuity of M(x). Suppose that at least one of the probabilities above is identically equal to unity, or at least in the limit equal to unity as $j \to \infty$ and $x \to A$. It is assumed that the variance of the random variable Y is finite for all values of x and that M(x) is continuous. Then if

$$\lim_{\substack{x \to A \\ j \to \infty}} \Pr[Y_{2j} > \alpha \mid x_{2j} \varepsilon \nu(A)] = 1$$

this must imply

$$\lim_{\substack{x \to A \\ j \to \infty}} \Pr[Y_{2j-1} < \alpha \mid x_{2j-1} \varepsilon \nu(A)] = 0$$

and vice versa.

Let there be a J such that $c_j = (x_j - x_{j-1}) / (y_j - y_{j-1})$ for all $j \geq J$. Consider the probability of such an event, that is,

$$\Pr[Y_1 < \alpha, \ldots, Y_{2j-1} < \alpha, \ldots] \Pr[Y_2 > \alpha, \ldots, Y_{2j} > \alpha, \ldots]$$

$$\leq \lim_{j \to \infty} \left(\max_{x_j \varepsilon \nu} \Pr[Y_{2j-1} < \alpha] \right)^j \lim_{j \to \infty} \left(\max_{x_j \varepsilon \nu} \Pr[Y_{2j} > \alpha] \right)^j$$

This is true since in $\nu(A)$ either $\max \Pr[Y_{2j-1} < \alpha]$ or $\max \Pr[Y_{2j} > \alpha]$ must be less than unity. Therefore, at least one of the limits will be identically zero.

Lemma 5.7. Let the sequence $[z_k]$ be the union of all subsequences of $[Z_k]$ such that $\lim_{k \to \infty} z_k = \infty$ where Z_k is the number of elements of the sequence $[c_j]$ having the form $(x_j - x_{j-1}) / (y_j - y_{j-1})$ which lie

between any two successive members of the sequence $[a_j]$. Then

$$\Pr[\lim_{k\to\infty} z_k = \infty] = 0.$$

Proof: Let $\lim_{j\to\infty} z_j = \infty$, then for each $2L > 0$ there exists a j such that $z_j > 2L$. Let us now consider the probability of such an event, that is, $\Pr[Y_1 < a, Y_2 > a, \ldots, Y_{2k-1} < a, Y_{2k} > a, \ldots, Y_{2L-1} < a, Y_{2L} > a]$. But, from the proof of Lemma 5.3, we know that $\lim_{L\to\infty} \Pr[Y_1 < a, \ldots, Y_{2L-1}] \Pr[Y_2 > a, \ldots, Y_{2L}] = 0$. It follows then that $\Pr[\lim_{j\to\infty} z_j = \infty] = 0$.

Theorem 5.1. Any given sequence $[c_j]$ is almost surely a member of the class of sequences $[b_j]$ where $[b_j]$ is defined by the following properties:

(a) $b_j > 0$ for all j

(b) $\sum_1^\infty b_j = \infty$

(c) $\sum_1^\infty b_j^2 < \infty$

Proof: Consider any sequence $[c_j]$ as defined in rule 12. By Lemma 5.1, each element of the sequence is necessarily positive. Condition (a) is satisfied.

Lemma 5.3, Lemma 5.6, and Lemma 5.7 assure us that every element of the sequence $[a_j]$ is almost surely contained in $[c_j]$. Therefore, since $c_j > 0$ for all j,

$$\sum_1^\infty c_j \geq \sum_1^\infty a_j, \text{ but } \sum_1^\infty a_j = \infty, \text{ then } \sum_1^\infty c_j = \infty$$

Hence, condition (b) is satisfied.

In order to show that the sequence $[c_j]$ satisfied condition (c),

consider the following infinite sum:

$$\sum_1^\infty c_j^2 = a_1^2 + a_2^2 + \ldots + a_2^2 + c_{11}^2 + c_{12}^2 + \ldots + c_{1M_1} + a_3^2$$

$$+ \ldots + a_3^2 + c_{21}^2 + \ldots + c_{2M_2}^2, \text{ etc.}$$

Where a_1 occurs once, a_2 occurs k_2 times, a_3 occurs k_3 times, etc. By Lemma 5.5, the sequence $[k_j]$ is almost surely bounded. By Lemma 5.7, the sequence $[M_j]$ is almost surely bounded. Let

$k = \max_j k_j$ and $M = \max_j M_j$.

If the sum $\sum_1^\infty c_j^2$ is convergent, it is absolutely convergent. The rearrangement of terms will not affect the convergence or the sum. Hence,

$$\sum_1^\infty c_j^2 \leq k \sum_1^\infty a_j^2 + M \sum_1^\infty a_j^2 = (k + M) \sum_1^\infty a_j^2 < \infty$$

which is the desired result, condition (c).

What is unusual about the theorem is that the conditions

(a) $c_j^2 > 0$

(b) $\sum_1^\infty c_j^2 < \infty$

(c) $\sum_1^\infty c_j = \infty$

are identical to those required by Blum (4) in his theorem which proves that the limit point of the sequence $[x_j]$ is θ with probability one for estimator I. The theorem can be stated as follows: Let $M(x)$ be the regression function corresponding to the family $H(y \mid x)$. Assume that $M(x)$ is a Legesque-measurable function satisfying

(a) $M(x) \leq c + |d| \; x$

(b) $\int_{-\infty}^{\infty} |y - M(x)|^2 \, dH(y \mid x) \leq \sigma^2 < \infty$

(c) $M(x) < a$ for $x < \theta$, $M(x) > a$ for $x > \theta$

(d) $$\inf_{\delta_1 \leq |x-\theta| \leq \delta_2} |M(x) - a| > 0$$

for every pair of numbers

(δ_1, δ_2) with $0 < \delta_1 < \delta_2 < \infty$

Let $[b_j]$ be a sequence of positive numbers such that

(e) $$\sum_1^\infty b_j = \infty$$

(f) $$\sum_1^\infty b_j^2 < \infty$$

Let x_1 be an arbitrary number. Define a sequence of random variables recursively by

(g) $$x_{j+1} = x_j + b_j (a - y_j)$$

where Y_j is a random variable distributed according to $H(y \mid x)$. Then x_j converges to θ with probability one.

CHAPTER VI

RESULTS OF THE EMPIRICAL STUDY

The most significant result of the empirical study is perhaps the apparent slowness with which estimator I converges to θ especially when $|x_j - \theta|$ is relatively large. For a test which involves less than fifty trials, estimator I when compared with II and III appears the least desirable in terms of bias. Figures 5 and 6 illustrate and emphasize the slowness of its convergence. A good rule is that unless the experimenter is certain that the initial value, x_1, is close to θ, he should avoid using estimator I (the Robbins-Monro stochastic approximation method).

On comparing estimators II and III, it is apparent that there are cases in which II appears better in terms of average bias than III, and vice versa. When $\alpha = 0.50$, the data from Table VIII indicate that III is slightly better for all sample sizes. Also, it should be noted that increasing the sample size had little effect in increasing the rate of convergence for all the estimators, I, II, and III. This is not true for other values of α. However, with sample size 10, estimator III gives a close approximation such that $|x_{49} - \theta| < 0.006$ for all θ_i for $i = 1, 2, 3, 4, 5$. The experimenter can be assured that estimator III will on the average give results with little bias when estimating θ for $\alpha = 0.50$.

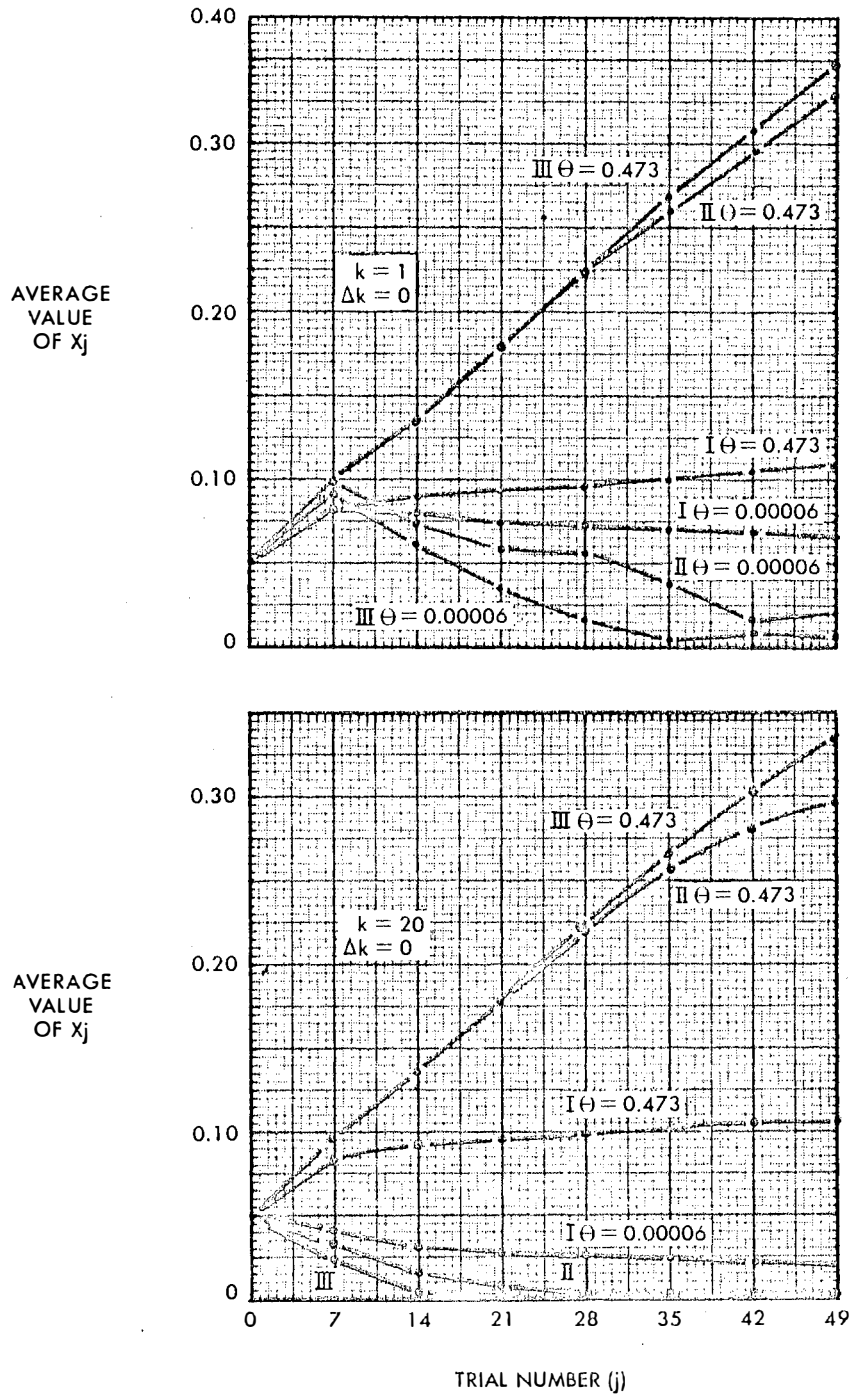

Fig. 5. Comparisons of the Rates of Convergence of Estimators I, II, and III.

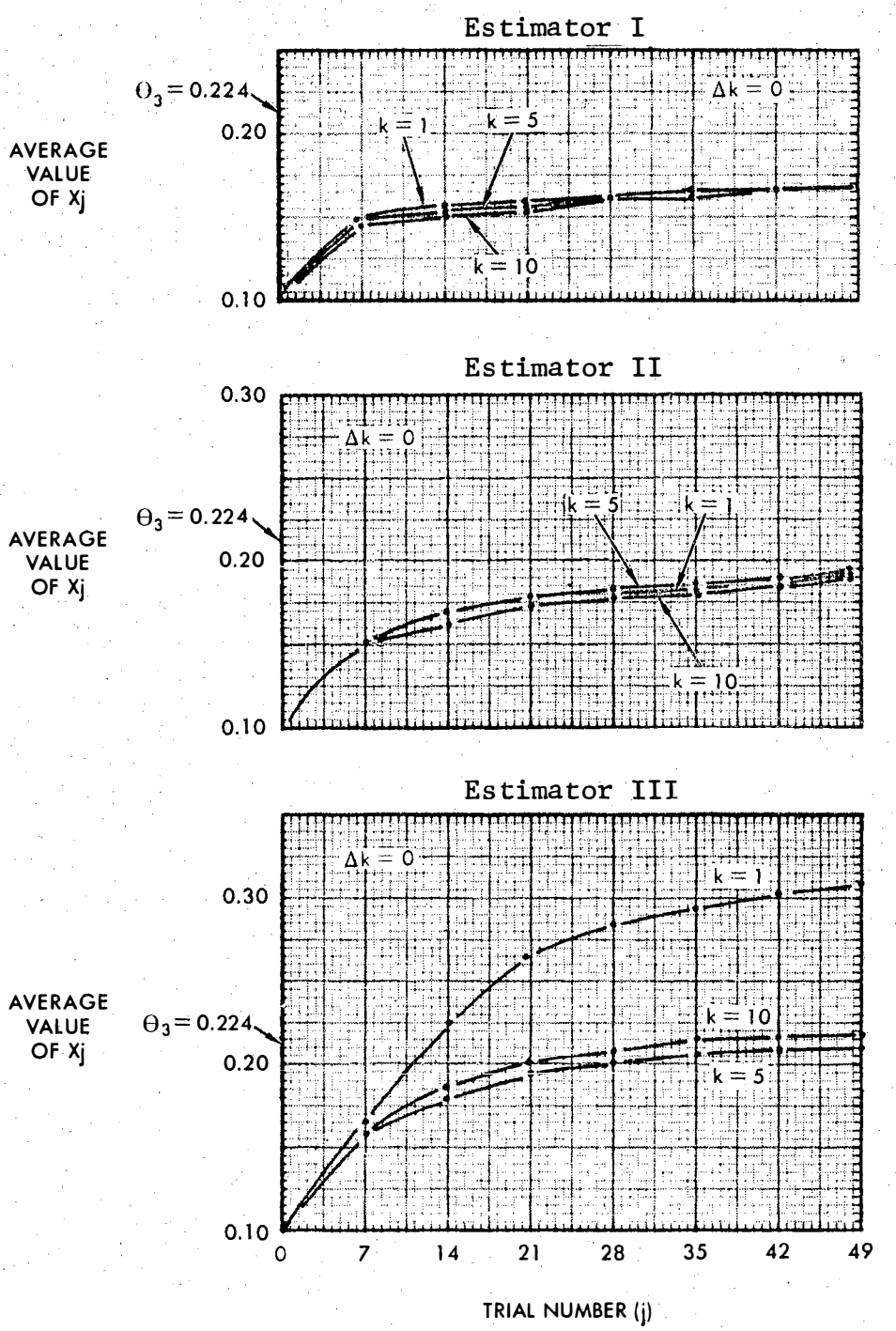

Fig. 6. The Effect of Sample Size on the Rate of Convergence when $\alpha = 0.10$.

On comparing estimators II and III for values of α other than $\alpha = 0.50$, it is seen that the bias in estimator II for sample size one is the smaller, but bias in estimators III becomes the smaller with increasing sample size. The data indicate that, for small sample sizes (1 and 5) and $\alpha = 0.50$, III is biased upward (see Fig. 3). In order to explain this, consider the following rationale.

Recalling that for sample size one

$$Y_j = \begin{cases} 0 & \text{if no response occurs} \\ 1 & \text{if a response occurs} \end{cases}$$

then if $\alpha \in (y_j, y_{j-1})$, a linear interpolation restricts x_{j+1} such that $x_j < x_{j+1} < x_{j-1}$ or $x_{j-1} < x_{j+1} < x_j$. Suppose that $\alpha = 0.05$, then one would expect in the neighborhood of θ that only one out of twenty trials would result in a response. Hence, there would occur on the average twenty steps to the right for one to the left. But when the response does occur, $x_{j+1} \in (x_{j-1}, x_j)$ or $x_{j+1} \in (x_j, x_{j-1})$, which offsets the large step back to the right which occurs in using I and II. Hence, one would expect estimator III to overestimate toward the right in the limit for $\alpha < 0.50$ and sample size one. It is assumed that α is always less than or equal to 0.50. But when $\alpha = 0.50$, the linear interpolation is meaningful and apparently there is little or no bias (see Table VII).

As the sample size increases, the bias associated with estimator III becomes smaller, indicating that either the symmetry of the density $dH(\bar{y} \mid x) / d\bar{y}$ or the decrease in the size of the variance of Y affects the convergence properties of III to θ.

Consider the function $p_k(x)$, which defines the probability that the

direction of the next step from x will not be in the direction of θ (see lower part of Fig. 1).

That is

$$p_k(x) = \begin{cases} s(x) & x \neq \theta \\ 1 & x = \theta \end{cases}$$

where

$$0 \leq s(x) \leq \max \left[\int_a^1 dG(\bar{y} \mid x), \int_0^a dG(\bar{y} \mid x) \right]$$

which in the limit as k increases without bound becomes $p_k(x) = 0$. This is sufficient for the estimator $x_{j+1} = x_j + c_j(a - \bar{y}_j)$ to converge in the limit to θ as k tends to ∞, and j tends to ∞.

The results support the following rules: For small sample sizes and a large number of trials, avoid using estimator III. For sample sizes larger than five and a small number of trials, estimator III gives smaller bias.

The direct relationship between smaller bias and large sample sizes poses a problem of efficiency of estimators, that is, the resolving of the problem of whether larger samples with a small number of trials is more desirable than unit sample sizes with a large number of trials. The solution depends on the nature of the test and must be solved for the specific test, hence, will not be considered here.

Increasing the sample size sequentially by increments of five (see Table VIII) does not, in the cases studied, decrease significantly the bias of the estimators, especially when the comparisons are based on sample sizes larger than one. This method can be used when sample sizes are not restrictive and relatively little bias is

important. However, it was observed that sample sizes will in some cases exceed one hundred experimental units at the forty-ninth trial. The absolute value of the bias in estimator III is decreased perhaps the most from such a scheme. It is important to note that increasing the sample size has little or no effect on the rate of convergence of estimators I and II.

CHAPTER VII

AN APPLICATION OF FINITE MARKOV CHAINS TO UP-AND-DOWN TESTING

Up-and-down testing is testing in which an increasing percentage of experimental units will respond as the severity of the test increases. The term response may denote a failure, an explosion, death, etc., depending upon the nature of the test. In such testing the precise severity of the test (that is, the precise magnitude of the variable concerned) which would result in a response cannot be measured. However, it can be observed that the applied severity results in a response or does not result in a response. In true up-and-down testing it is not possible to make more than one observation on a given experimental unit. Once a test has been made the experimental unit is altered, so that bona fide results cannot be obtained from a second test on the same experimental unit.

Let M(x) represent a monotonic increasing function which describes on the average the manner in which percentage of items responding changes with severity applied, say x. Let Y(x) be a random variable defined as the percentage of items responding in a single test at a fixed level x. Then $E[Y(x)] = M(x)$, where the form of M(x) is unknown to the experimenter. Since M(x) is monotonically increasing function, then the regression equation (1) has a single root $x = \theta$, which is to be estimated for $\alpha = 1/2$ in (1).

In this chapter a procedure to set confidence limits on the parameter θ will be formulated without making any assumptions concerning $M(x)$ other than continuity, monotonicity, linearity in the neighborhood of θ, and that an a priori estimate of a lower bound for the slope of $M(x)$ in the neighborhood of θ is available to the experimenter.

The formulation of the confidence interval will be based on three basic theorems applicable to regular and ergodic stochastic matrices. The theorems are stated in Appendix A as Theorems A. 1, A. 2, and A. 3.

Preliminary Notions and Notation

Let the random variable take on only two values, unity with probability $M(x)$ and zero with the probability $1 - M(x)$. Let there be two numbers a and b, where a is less than b, which are known to the experimenter and such that

(16)
$$\begin{aligned} Y(x) &= 0 && x \leq a \\ &= 1 && x \geq b \end{aligned}$$

The regression function $M(x)$ will have the following properties:

$$\begin{aligned} M(x) &= 0 && x \leq a \\ &= E[Y(x)] && a < x < b \\ &= 1 && x \geq b \end{aligned}$$

and $E[Y(x)]$ is a monotone increasing function unknown to the experimenter. The test plan will require that at each level, x_j, a sample of k values of Y be taken. Since the sample mean

$$\bar{Y}(x) = \Sigma Y_i / k$$

where

$$Y_i = \begin{cases} 0 & \text{if no response occurs} \\ 1 & \text{if a response occurs} \end{cases}$$

has the same expected value as the random variable $Y(x_j)$, $\overline{Y}(x_j)$ can be computed and compared with α in order to estimate whether x_j is less than, equal to, or greater than θ. Since

$$E[\overline{Y}(x)] = E[Y(x)] = M(x) \quad ,$$

the value of x such that $E[\overline{Y}(x)] = \alpha$ is also θ, the desired result, can be obtained by solving the latter equation.

Our purpose is to determine an interval estimate of θ, the severity at which, on the average, one half the experimental units will respond, that is, $\alpha = 1/2$ in (1). We have selected $\alpha = 1/2$ because in many up and down tests this is the value which is of interest. The value θ when $\alpha = 1/2$ may be called the mean lethal dose, mean critical point, mean breaking strength, etc., depending on the experiment being performed.

It is assumed that the experimenter knows the values a and b such that (16) is true. Partitioning the interval (a, b) into n-1 equally spaced abutting intervals, we let the n states (locations) be defined as follows:

$$S_1 = a$$

$$S_2 = a + (b - a) / (n - 1)$$

$$S_3 = a + 2(b - a) / (n - 1)$$

$$\vdots$$

$$S_j = a + (j - 1)(b - a) / (n - 1)$$

$$\vdots$$

$$S_n = a + (n - 1)(b - a) / (n - 1) = b$$

It is at these values of $x = S_j$ for $j = 1, 2, \ldots, n$ that the trials will be performed. A trial is defined as a test of sample size k at a specified level S_j. For each trial a decision is made in selecting the level at which the next trial will be performed by using the following rules:

Rule 1. If $\overline{Y}(x_j)$ is greater than α, then the next trial will be performed at S_{j-1}.

Rule 2. If $\overline{Y}(x_j)$ is equal to α, then the next trial will be performed again at S_j.

Rule 3. If $\overline{Y}(x_j)$ is less than α, then the next trial will be performed at S_{j+1}.

This procedure defines a random walk with reflecting barriers. We will make use of this fact in developing the theory for determining the desired confidence limits.

The Theory.

The random walk with reflecting barriers may be thought of as a finite Markov process. A finite Markov process is a stochastic process which moves through a finite number of states, and for which probability of entering a certain state depends only on the last state occupied. The transition matrix P of such a process is displayed on the following page. In the transition matrix P we use the notation so that the symbol p_j will denote the probability that the next trial will be performed at a state which is closer in distance to the value of θ than

	S_1	S_2	S_3	S_4	S_{i-2}	S_{i-1}	S_i	S_{i+1}	S_{i+2}	S_{i+3}	S_{n-2}	S_{n-1}	S_n
S_1	0	1	0	0									
S_2	q_1	r_1	p_1	0									
S_3	0	q_2	r_2	p_2									
S_4	0	0	q_3	r_3									
S_{i+2}					r_{i-3}	p_{i-3}	0	0	0	0			
S_{i-1}					q_{i-2}	r_{i-2}	p_{i-2}	0	0	0			
S_i					0	p_{-1}	p_o	p_{+1}	0	0			
S_{i+1}					0	0	p'_{-1}	p'_o	p'_{+1}	0			
S_{i+2}					0	0	0	p_{i+2}	r_{i+2}	q_{i+2}			
S_{i+3}					0	0	0	0	p_{i+3}	r_{i+3}			
S_{n-2}											r_{n-2}	q_{n-2}	0
S_{n-1}											p_{n-1}	r_{n-1}	q_{n-1}
S_n											0	1	0

The above array of elements is the transition matrix of a random walk with reflecting barriers and the associated states along upper left hand periphery of the matrix. We denote this matrix by the letter P.

the preceeding one. Similarly, the symbol q_j is used for the value of the probability that the next trial will be performed at a state farther in distance from θ than the previous trial. The symbol r_j is used for the probabilities that the next trial will take place at the same state as the preceeding one. The symbols p_{-1}, p_0, p_{+1}, p'_{-1}, p'_0, and p'_{+1} are used for those states in which no meaning can be attached to as to which states S_i, S_{i+1} is the closer location to the value of θ, since we assume that the value of θ lies somewhere between or on S_i and S_{i+1}. The probability is zero that S_i will be identically the value of θ.

Let us consider the upper and lower tolerance equations $L_1(x)$ and $L_2(x)$. These equations are loci of the points (x, y) such that

$$\Pr[\overline{Y}(x) \leq L_1(x)] \geq 1 - q$$

$$\Pr[\overline{Y}(x) \geq L_2(x)] \geq 1 - q$$

for all x belonging to the interval (a, b). The symbol q denotes a preassigned probability value whose range we will determine later.

Let $I(\theta)$ be the interval $[x(L_1), x(L_2)]$ as we have previously defined in Chapter 4, Fig. 1. By defining $p + q = 1$, it can be seen by Fig. 1 that the following inequalities are true:

$$\begin{aligned}
0 &\leq \Pr[\overline{Y} > \alpha \mid x \notin I(\theta) \text{ and } x < \theta] \leq q \\
0 &\leq \Pr[\overline{Y} < \alpha \mid x \notin I(\theta) \text{ and } x > \theta] \leq q \\
q &\leq \Pr[\overline{Y} > \alpha \mid x \in I(\theta) \text{ and } x < \theta] \leq 1/2 \\
q &\leq \Pr[\overline{Y} < \alpha \mid x \in I(\theta) \text{ and } x > \theta] \leq 1/2 \\
1/2 &\leq \Pr[\overline{Y} < \alpha \mid x \in I(\theta) \text{ and } x < \theta] \leq p \\
1/2 &\leq \Pr[\overline{Y} > \alpha \mid x \in I(\theta) \text{ and } x > \theta] \leq p \\
p &\leq \Pr[\overline{Y} < \alpha \mid x \notin I(\theta) \text{ and } x < \theta] \leq 1 \\
p &\leq \Pr[\overline{Y} > \alpha \mid x \notin I(\theta) \text{ and } x > \theta] \leq 1
\end{aligned} \tag{17}$$

From these inequalities the following bounds for the elements of the matrix P are defined:

$$0 \leq q_j \leq q \qquad \text{for all } j = 1, 2, 3, \ldots, i-1, i+2, \ldots, n-1$$

$$p \leq p_j \leq 1 \qquad \text{for all } j = 1, 2, 3, \ldots, i-1, i+2, \ldots, n-1$$

$$q \leq p_{-1} \leq 1/2$$

$$q \leq p'_{+1} \leq 1/2$$

$$1/2 \leq p_{+1} \leq p$$

$$1/2 \leq p'_{-1} \leq p$$

Since the sum of the elements of each row of P must be unity, we have established bounds for each element of the transition matrix.

From Theorem A. 3, it follows that

$$\lim_{n\to\infty} \Pr[\, \theta \in \{S_{i-1}, S_i, S_{i+1}, S_{i+2}\}\,] = \sum_{i-1}^{i+2} \beta_j$$

where β_j is the jth term of the β vector as defined in Appendix A, and n is the number of trials performed before terminating a test. Let us determine what the range of q must be in order that

$$\sum_{i-1}^{i+2} \beta_j \geq 1 - q \tag{18}$$

where q is the same q necessary to define the tolerance equations $L_1(x)$ and $L_2(x)$.

Solving $\beta P = \beta$ for the vector β in terms of β_1 where $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$. The results are:

$$\beta_2 = \beta_1/q_1$$

$$\beta_3 = \beta_1 p_1/q_1 q_2$$

$$\beta_4 = \beta_1 p_1 p_2/q_1 q_2 q_3$$

.

$$\beta_{i-1} = \beta_1 p_1 p_2 \cdots p_{i-3}/q_1 q_2 \cdots q_{i-2}$$

$$\beta_i = \beta_1 p_1 p_2 \cdots p_{i-2}/q_1 q_2 \cdots q_{i-2} p_{-1}$$

$$\beta_{i+1} = \beta_1 p_1 p_2 \cdots p_{i-2} p_{+1}/q_1 q_2 \cdots q_{i-2} p_{-1} p'_{-1}$$

$$\beta_{i+2} = \beta_1 p_1 p_2 \cdots p_{i-2} p_{+1} p'_{+1}/q_1 q_2 \cdots q_{i-2} p_{-1} p'_{-1} p_{i+2}$$

For simplicity and ease of manipulation, the remaining elements of the vector β are written in terms of β_{i+2}.

$$\beta_{i+3} = \beta_{i+2} q_{i+2}/p_{i+3}$$

$$\beta_{i+4} = \beta_{i+2} q_{i+2} q_{i+3}/p_{i+3} p_{i+4}$$

$$\beta_{n-1} = \beta_{i+2} q_{i+2} q_{i+3} \cdots q_{n-2}/p_{i+3} p_{i+4} \cdots p_{n-1}$$

$$\beta_n = \beta_{i+2} q_{i+2} q_{i+3} \cdots q_{n-1}/p_{i+3} p_{i+4} \cdots p_{n-1}$$

Considering the ratio

$$r = \frac{\sum_{i-1}^{i+2} \beta_j}{1 - \sum_{i-1}^{i+2} \beta_j} = \frac{\sum_{i-1}^{i+2} \beta_j}{\sum_{j \neq i-1, i, i+1, i+2} \beta_j}$$

we note that the ratio is void of β_1 and is a function of the elements of the transition matrix P. It follows that

$$\sum_{i-1}^{i+2} \beta_j = 1/(1 + 1/r)$$

Hence, if r is minimized, a lower bound for the left side of the above equality can be obtained.

Consider first

$$\sum_{i-1}^{i+2} \beta_j = \beta_1 p_1 p_2 \cdots p_{i-3}/q_1 q_2 \cdots q_{i-2} + \beta_1 p_1 p_2 \cdots p_{i-2}/q_1 q_2 \cdots q_{i-2} p_{-1} +$$

$$\beta_1 p_1 p_2 \cdots p_{i-2} p_{+1}/q_1 q_2 \cdots q_{i-2} p_{-1} p'_{-1} + \beta_1 p_1 p_2 \cdots p_{i-2} p_{+1} p'_{+1}/q_1 q_2 \cdots q_{i-2} p_{-1} p'_{-1} p_{i+2}$$

$$= (\beta_1 p_1 p_2 \cdots p_{i-2}/q_1 q_2 \cdots q_{i-2})(1/p_{i-2} + 1/p_{-1} + p_{+1}/p_{-1} p'_{-1} + p_{+1} p'_{+1}/p_{-1} p'_{-1} p_{i+2})$$

and

$$1 - \sum_{i-1}^{i+2} \beta_j = \sum_{j=i-1}^{i} \beta_j = \beta_1 + \beta_1/q_1 + \beta_1 p_1/q_1 q_2 + \cdots + \beta_1 p_1 p_2 \cdots p_{i-4}/q_1 q_2 \cdots q_{i-3} + \beta_{i+2} q_{i+2}/p_{i+3} + \cdots +$$

$$\beta_{i+2} q_{i+2} q_{i+3} \cdots q_{n-2}/p_{i+3} p_{i+4} \cdots p_{n-1} + \beta_{i+2} q_{i+2} q_{i+3} \cdots q_{n-2} q_{n-1}/p_{i+3} p_{i+4} \cdots p_{n-1}$$

$$= (\beta_1 p_1 p_2 \cdots p_{i-2}/q_1 q_2 \cdots q_{i-2})(q_1 q_2 \cdots q_{i-2}/p_1 p_2 \cdots p_{i-2}[1 + (1/q_1)(1 + p_1/q_2 + p_1 p_2/q_2 q_3 +$$

$$\cdots + p_1 p_2 \cdots p_{i-4}/q_1 q_2 \cdots q_{i-3}] + [p_{+1} p'_{+1}/p_{-1} p'_{-1} p_{i+2}][q_{i+2}/p_{i+3}(1 + q_{i+3}/p_{i+4} + \cdots +$$

$$q_{i+3}/p_{i+4} + \cdots + q_{i+3} q_{i+4} \cdots q_{n-2}/p_{i+4} q_{i+5} \cdots p_{n-1}] + q_{i+2} q_{i+3} \cdots q_{n-1}/p_{i+3} \cdots p_{n-1}])$$

On forming the ratio r and simplifying, we find

$$(19) \qquad r = \frac{1/p_{i-2} + 1/p_{-1} + p'_{-1} + p_{+1} p'_{+1}/p_{-1} p'_{-1} p_{i+2}}{(q_1 q_2 \cdots q_{i-2}/p_1 p_2 \cdots p_{i-2})[1 + 1/q_1(1 + p_1/q_2 + p_1 p_2/q_2 q_3 + \cdots +$$

$$p_1 p_2 \cdots p_{i-4}/q_2 q_3 \cdots q_{i-3})] + [p_{+1} p'_{+1}/p_{-1} p'_{-1} p_{i+2}][q_{i+2}/p_{i+3}(1 + q_{i+3}/p_{i+4} + \cdots +$$

$$q_{i+3} q_{i+4} \cdots q_{n-2}/p_{i+4} p_{i+5} \cdots p_{n-1}) + q_{i+2} q_{i+3} \cdots q_{n-1}/p_{i+3} p_{i+4} \cdots p_{n-1}]$$

In order to minimize r, we first note that if q_{i-1}, q_{i+2} both are equal to q, r_{i-1} and r_{i+2} are both equal to zero, p_{-1} and p'_{+1} are both equal to one-half and finally, p_o and p'_o both are equal to zero, then the probability of leaving the set of states $[S_{i-1}, S_i, S_{i+1}, S_{i+2}]$ is maximized within the constraints imposed by the inequalities (17). Using this fact and the remaining inequalities of (17), the ratio, r, can be bounded below by minimizing the numerator and maximizing the denominator of (19). Substituting these values and simplifying, it follows that

$$r > \frac{1 + 2 + 2 + 1}{(q/p)^{i-2}(1 + [1/q][(p/q)^{i-3} - 1]/[p/q - 1] + (1/p)([q/p][1 - (q/p)^{n-3-i}]/[1 - q/p] + [q/p]^{n-3-i}q)}$$

which on simplifying further becomes

$$r > \frac{3(1 - q)(1 - 2q)}{q} \tag{20}$$

Therefore,

$$\sum_{i-1}^{i+2} \beta_j > \frac{r^*}{1 + r^*} \tag{21}$$

where r^* is the right hand side of the inequality (20). To find the values of q such that

$$\sum_{i-1}^{i+2} \beta_j > 1 - q ,$$

consider the inequality

$$\frac{r^*}{1 + r^*} > 1 - q$$

which is equivalent to

$$\frac{3(1 - q)(1 - 2q)}{q + 3(1 - q)(1 - 2q)} > 1 - q$$

On simplifying,

$$(1 - q)(2q)(1 - 3q) > 0 \tag{22}$$

Solving the inequality (22), we find that if $0 < q < 1/3$, then the inequality (18) is true.

It is important to note that the lower bound for the quantity $\sum_{i-I}^{i+2} \beta_j$ as formulated by the inequality (21) is independent of the size of the transition matrix (that is, the number of states) and the location of the ith state (that is, the location of the solution $x = \theta$). Therefore, if the value of q is such that q is greater than zero, yet less than one-third, Theorem A. 3 of Appendix A assures that there exists an N' such that

$$\sum_{j=i-1}^{i+2} \beta_{kj}^{N'} > 1 - q \quad \text{for all } k \text{ and } N \geq N' \tag{23}$$

where the symbol

$$\beta_{kj}^{N'}$$

is the element of the kth row and the jth column of the transition matrix $P^{N'}$. That is, the fraction of times in the first N' steps that the process moves to at least one of the states, S_{i-1}, S_i, S_{i+1}, S_{i+2} will be at least $1 - q$.

Let the desired length of the confidence interval be a preassigned value 4δ and let $\tan C$ be an a priori estimate of the greatest lower bound of the slope of $M(x)$ at $x = \theta$. It is also required that 4δ be sufficiently small so that $M(x)$ is essentially linear in the neighborhood of $x = \theta$.

The size of the sample, say k, for each trial is selected such that

$$\Pr[L_2(\theta) \leq \overline{Y}(\theta) \leq L_1(\theta)] > 1 - 2q \tag{24}$$

$$L_2(\theta) = 1/2 + \delta \tan C$$

$$L_1(\theta) = 1/2 + \delta \tan C$$

and q is the preassigned confidence coefficient desired. It is also required that the inequalities

$$\Pr[L_2(\theta) \leq \overline{Y}(\theta)] \geq 1 - q$$

$$\Pr[L_1(\theta) \geq \overline{Y}(\theta)] \geq 1 - q$$

be true. Fig. 1 gives the graphical sketch of these conditions.

The value of x at which the variance of $\overline{Y}$ is maximum is the solution of (1), that is $x = \theta$. Therefore, the value $x = \theta$ is the one at which the tolerance equations, $L_1(x)$ and $L_2(x)$, deviate greatest from the regression equation M(x). Recalling that M(x) is assumed linear in the neighborhood of θ, and referring to Fig. 7, the following inequalities are seen to be true:

$$M(x) - \delta' \tan C' \leq L_2(x) < M(x) < L_1(x) \leq M(x) + \delta' \tan C'$$

where $\tan C'$ is the true slope of M(x) at $x = \theta$. The quantity δ' is defined by the ratio

$$\delta' = (\delta \tan C) / \tan C' \ .$$

By fixing $\overline{Y}$ to be equal to one-half, we can solve the following equations:

$$M(x) - \delta' \tan C' = 1/2$$

$$L_2(x) = 1/2$$

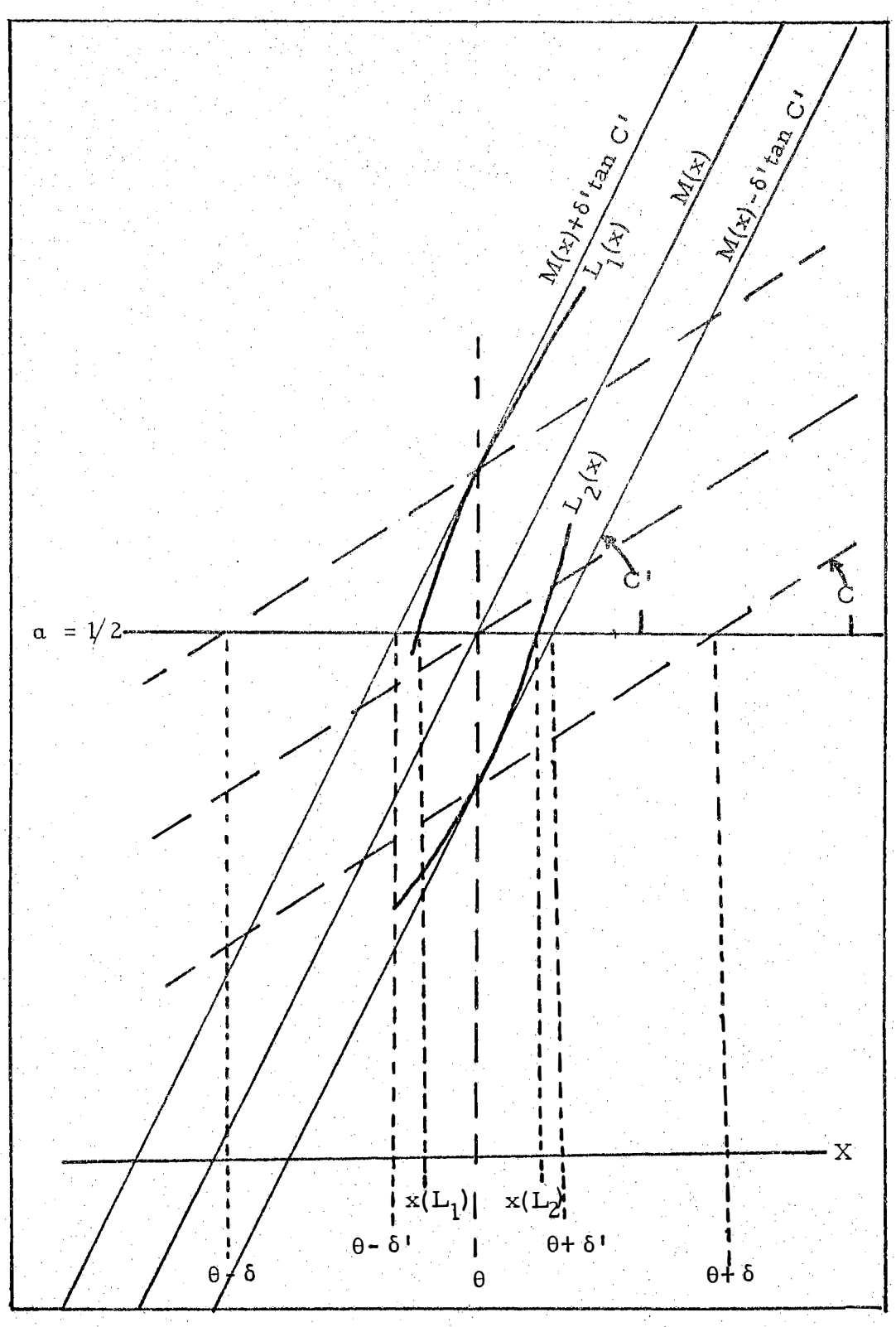

Fig. 7. Sketch of M(x) in the Neighborhood of x = θ

$$M(x) = 1/2$$

$$L_1(x) = 1/2$$

$$M(x) + \delta' \tan C' = 1/2$$

The solutions to the equations are $\theta - \delta'$, $x(L_1)$, θ, $x(L_2)$, $\theta + \delta'$, respectively. Due to the monotonicity of each function appearing in the inequality it follows that

$$\theta - \delta' \leq x(L_1) < \theta < x(L_2) \leq \theta + \delta'.$$

Previously, we defined the interval $I(\theta)$ to be $[x(L_1), x(L_2)]$. In order to be assured that the probability bounds imposed upon the elements of the matrix P to be valid, we must show that the length of $I(\theta)$ is such that it can contain at most only two states, S_i and S_j. Since $\tan C$ is known and is a lower bound for the slope of $M(x)$ at $x = \theta$, that is,

$$\tan C \leq \tan C'$$

it follows immediately that

$$\delta' \leq \delta .$$

Recalling that $|\theta + \delta - (\theta - \delta)| = |S_j - S_{j+1}| = 2\delta$, for all j, we can conclude that the interval $I(\theta)$ can cover at most two states. Since $I(\theta) \leq 2\delta$, the only two states that $I(\theta)$ can cover then are S_i and S_{i+1}.

The number of states necessary to perform the test is fixed by the value of δ. Let n be the number of states (the size of the transition matrix P), then n is the smallest integer larger or equal to

(25)
$$\frac{b - a}{\delta} + 1$$

Noting that the desired sample size is determined from (24), the number of trials necessary to complete a test is the only quantity that remains to be determined prior to the test. In order to assure ourselves that the transition matrix is regular, the sample size k must be the smallest even integer which equals or exceeds that which satisfies (24).

In order to determine the number of trials necessary so that (23) is true, consider the transition matrix that minimizes the probability that the next trial will occur at a state nearer to θ. That is, when i implies that $S_i < \theta < S_{i+1}$, we consider the matrix

$$\underline{P}_i = \begin{array}{c} \\ S_i \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array} \begin{bmatrix} 0 & 1 & 0 & 0 & & & & & & & & \\ q & 0 & p & 0 & & & & & & & & \\ 0 & q & 0 & p & & & & & & & & \\ & & & & q & 0 & p & 0 & 0 & 0 & & \\ & & & & 0 & 1/2 & 0 & 1/2 & 0 & 0 & & \\ & & & & 0 & 0 & 1/2 & 0 & 1/2 & 0 & & \\ & & & & 0 & 0 & 0 & p & 0 & q & & \\ & & & & & & & & & & p & 0 & q \\ & & & & & & & & & & 0 & 1 & 0 \end{bmatrix}$$

Let n_k be the power of the matrix $\underline{P}_i$ such that

$$\sum_{j=i-1}^{i+2} \beta_{mj}^{n_i} \geq 1 - q \quad \text{for all } m = 1, 2, \ldots, n-1$$

where m denotes the row. Repeating this procedure until i exhausts its range, the following set of values for n_i is determined

$$[n_k] = [n_1, n_2, \ldots, n_{n-1}]$$

The desired number of trials follows immediately and is equal to

(26) $$\max_i \; n_i$$

We will denote this number of trials as N. It is important to note that no restriction has been made concerning the location at which the initial trial is to be made. Therefore, if the first location must be selected by a random process, equation (26) gives the desired number of trials.

Upon performing N trials we are assured that

(27) $$\sum_{j=i-1}^{i+2} \beta_{mj}^{N} \quad 1 - q \geq \quad \text{for all } m.$$

From (27) it follows then that

$$Pr_m(S_N \in [S_{i-1}, S_i, S_{i+1}, S_{i+2}]) \geq 1 - q$$

regardless of the state at which the initial trial was performed. Noting that $S_i \leq \theta \leq S_{i+1}$, it follows that if S_N is the state at which the N + 1 trial would be performed that

$$Pr(S_{i-1} \leq S_N \leq S_{i+2}) \geq 1 - q$$

or equivalently,

(28) $$Pr(S_{N-2} \leq \theta \leq S_{N+2}) \geq 1 - q$$

which is the desired non-parametric confidence limit on the parameter θ with the coefficient $1 - q$.

The number N given by (26) is the smallest number such that the probability statement (28) will be true independent of the location of the state at which the initial trial will be performed. By selecting the initial trial to be S_m, a substantial decrease in the number of

trials necessary to assure that (28) will hold may be realized.

Consider the set M_i where M_i is the smallest power such that $\underline{P}_i^{M_i}$ assures that

$$(29) \qquad \sum_{j=i-1}^{i+2} \beta_{mj}^{M_i} \geq 1 - q$$

and m is that state at which the initial trial will take place. Forming the set

$$[M_i] = [M_1, M_2, \ldots, M_{n-1}]$$

the desired number of trials is

$$\max_i M_i$$

which is the least number of trials necessary given that the initial trial was performed at the m th state.

A natural choice of m would be that value of m such that the state at which the first trial would be performed would be near the mid-range of the interval (a, b).

Numerical Example.

Suppose that an interval estimate is desired for the mean lethal dosage of gamma radiation for a specified species of dogs. Let x be the unit of measure of radiation. Let the value of a = 0 and b = 1, $0 \leq x \leq 1$. It is desired that the length of the 90% confidence interval should be no longer than 0.20. From previous data it is known that the slope of M(x), the regression equation, is never less than 2 in the neighborhood of $\alpha = 1/2$. The procedure to obtain such a confidence limit is as follows:

1. Determination of the number of states and the size of the transition matrix P. Since the length of the confidence interval is

preassigned to be 0.20, it follows that the distance between states will be $\delta = I/4$, where I is the length of the confidence interval. Hence $\delta = 0.20/4 = 0.05$. Therefore, from (25) the number of states is $n = 1 + 1/0.05 = 21$.

2. Determination of the sample size at each trial. First we determine $L_1(\theta)$ and $L_2(\theta)$ from (24). That is,

$$L_1(\theta) = 0.50 + (0.05)(2) = 0.60$$

$$L_2(\theta) = 0.50 - (0.05)(2) = 0.40 .$$

The sample size k is such that

$$\Pr(0.40 \leq \bar{Y}(x = \theta) \leq 0.60) \geq 1 - 2q = 0.80 .$$

Using tables of the binomial probability density, we find that the value for $k = 40$.

3. Determination of the number of trials necessary in order to assure the validity of the confidence statement. In order to take advantage of the smaller sample size, one first defines the location at which the initial trial will take place, say S_{11}. Note that this will fix S_m as being S_{11}, that is $m = 11$ in the inequality (29). The selection of S_m is arbitrary but generally can be selected as that state nearest to the midrange of the interval $[a, b]$.

Letting $p = 0.90$ and $q = 0.10$ the matrix $\underline{P}_i$ is formed; where the subscript i implies the state such that

$$S_i < \theta < S_{i+1} .$$

Since it is unknown to the experimenter the value for i such that the above inequality holds, twenty matrices $\underline{P}_i$ must be built with i varying from 1 through 20. The $[\max_i M_i]$ is the number of trials that will

validate the confidence statement where M_i is the power of the matrix $\underline{P}_i$ such that the inequality (29) is true with m = 11.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
M_i	13	13	11	10	8	7	6	5	4	2	2	4	5	6	7	8	10	11	13	13

Hence, the number of trials will be 13.

4. Performing the test. Letting S_{11} be the state at which the initial trial is to be performed, one starts the test. After 13 trials, using Rules 1, 2, or 3, to determine the location at which each succeeding trial will be performed, the test ends, say, at S_{13}. The desired confidence interval follows from (28), that is we can say with 90% confidence that the closed interval [0.50, 0.70] will contain the solution of the regression equation θ.

BIBLIOGRAPHY

(1) Crow, E. L., U. S. Naval Ordnance Test Station, Statistics Manual, China Lake, California, NOTS, 1955. (NAVORD Report 3369, NOTS 948.) P. 93.

(2) Dixon, W. J. and Mood, A. M. "A Method for Obtaining and Analyzing Sensitivity Data," American Statistical Association Journal, Vol. 43 (1948), p. 109.

(3) Robbins, H. and Monro, S. "A Stochastic Approximation Method," Annals of Mathematical Statistics, Vol. 22 (1951), pp. 400-407

(4) Blum, J. R. "Approximation Methods Which Converge With Probability One," Annals of Mathematical Statistics, Vol. 2 (1951), pp. 382-386

(5) Keston, H "Accelerated Stochastic Approximation," Annals of Mathematical Statistics, Vol 29 (1938), pp. 41-59

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

(7) Kiefer, J. and Wolfowitz, J. "Stochastic Estimation of the Maximum of a Regression Function," Annals of Mathematical Statistics, Vol. 23 (1952), pp. 462-466.

(8) Wolfowitz, J. "On Stochastic Approximation Methods," Annals of Mathematical Statistics, Vol. 27 (1956), pp. 1151-1156

(9) Teichrow, D. "An Empirical Investigation of the Stochastic Approximation Method of Robbin and Monro." Unpublished Report.

(10) Guttman, L. and Guttman, R. "An Illustration of the Use of Stochastic Approximation," Biometrics, Vol. 15 (1959), pp. 551-559.

(11) Kemeny, J. G. and Snell, J. L. "Finite Markov Chains, New York, D. Van Nostrand Company, Inc. 1960

APPENDIX A

DEFINITIONS AND THEOREMS APPLICABLE TO FINITE MARKOV CHAINS

The theorems and definitions listed here are essentially the same as those given by Kemeny and Snell [11].

Fundamental Definitions

Definition A. 1. A finite Markov chain is a stochastic process which moves through a finite number of states, and for which the probability of entering a certain state depends only on the last state occupied.

Definition A. 2. An ergodic set of states is a set in which every state can be reached from every other state, and which cannot be left once it is entered.

Definition A. 3. An ergodic chain is one whose states form a single ergodic set; or--equivalently--a chain in which it is possible to go from every state to every other state.

Definition A. 4. A cyclic chain is an ergodic chain in which each state can only be entered at certain periodic intervals.

Definition A. 5. A regular chain is an ergodic chain that is not

cyclic.

Basic Theorems.

Theorem A.1. Let π_n be the induced measure for the outcome function for a finite Markov chain with initial probability vector π_o and transition matrix P. Then :

$$\pi_n = \pi_o P^n$$

This theorem shows that the key to the study of the induced measures of the outcome function of a finite Markov chain is the study of the powers of the transition matrix. The entries of these powers have themselves an interesting probabilistic interpretation. To see this take as initial vector π_o the vector with 1 in the ith component and 0 otherwise. Then by Theorem A.1, $\pi_n = \pi_o P^n$. But $\pi_o P^n$ is the ith row of the matrix P^n. Thus the ith row of the nth power of the transition matrix gives the probability of being in each of the various states under the assumption that the process started in state S_i.

Theorem A.2. If P is a regular transition matrix then

(i) The powers P^n approach a probability matrix A.

(ii) Each row of A is the same probability vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$.

(iii) The components of α are positive.

(iv) For any probability vector π, πP^n approaches the vector α as n tends to infinity.

(v) The vector α is the unique vector such that $\alpha P = \alpha$.

We note from Theorem A.2, for a regular Markov chain there is a limiting probability α_j of being in state S_j for a large number of steps. The symbol $v_j^{(n)}$ is the fraction of times in the first n steps that the process moves to state S_j. The law of large numbers for regular Markov chains can be stated as follows:

Theorem A.3. Consider a regular Markov chain with limiting vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$. For any initial vector π

$$E_\pi[v_j^{(n)}] \to \alpha$$

and for any $\epsilon > 0$

$$\Pr[\ |v_j^{(n)} - \alpha_j| > \epsilon] \to 0$$

as n tends to infinity.

VITA

Patrick Lowry Odell

Candidate for the Degree of

Doctor of Philosophy

Thesis: STOCHASTIC APPROXIMATION AND NON-PARAMETRIC INTERVAL ESTIMATION IN SENSITIVITY TESTING WHICH INVOLVES QUANTAL RESPONSE DATA

Major Field: Mathematical Statistics

Biographical:

Personal Data: Born at Watonga, Oklahoma, November 29, 1930, the son of Max V. and Pamela A. Odell.

Education: Attended grade school at Watonga, Oklahoma; graduated from Watonga High School in 1948; in 1952 received the Bachelor of Science degree from the University of Texas, Austin, Texas, with a major in mathematics and a minor in physics, attended the University of California for one year with major study in meteorology; in 1958 received the Master of Science degree from Oklahoma State University, Stillwater, Oklahoma, with a major in mathematics; attended University of New Mexico for two years with major study in mathematics; completed requirements for the Doctor of Philosophy degree in May, 1962.

Professional experience: Mathematician at Flight Determination Laboratory, White Sands Proving Ground, Las Cruces, New Mexico, from June, 1952 to July, 1953; meteorologist in the United States Air Force from July, 1953 to June, 1957; assistant in the Department of Mathematics, Oklahoma State University, as teaching assistant from September, 1957 to August 1958; research scientist from Kaman Aircraft Nuclear Division, Albuquerque, New Mexico, August 1958 to February, 1959; mathematician at United States Naval Weapons Evaluation Facility, Albuquerque, New Mexico, February 1959 to May, 1962, graduate assistant in the Department of Mathematics and Statistics, Oklahoma State University, September, 1960 to May, 1962.