STOCHASTIC DECISION MODEL FOR ARITHMETIC PROGRAMMING

Mohammad Zia-Ul Haque

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

STOCHASTIC DECISION MODEL
for
ARITHMETIC PROGRAMMING

by

MOHAMMAD ZIA UL HAQUE

Thesis Advisor: J. K. Arima

Approved for public release; distribution unlimited.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) Stochastic Decision Model for Arithmetic Programming	5. TYPE OF REPORT & PERIOD COVERED Thesis (March 1976)
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Mohammad Zia Ul Haque	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940	12. REPORT DATE March 1976
	13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Markovian decision model; branching criteria; all-or-none learning theory; programmed learning; probabilistic decision model; model for arithmetic programming; stochastic model; arithmetic learning; advancement criteria.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Few if any validated guidelines exist for making decisions about the design, media, or format of new instructional products. This study examined strings of programmed learning responses to create general guidelines for making such decisions. Using a Markov model, tables were developed relating the expected proportion of students to be in a solution state at a given accuracy level and at a given level of confidence with respect to the length of response strings.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
(Page 1) S/N 0102-014-6601 |

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

STOCHASTIC DECISION MODEL FOR ARITHMETIC PROGRAMMING

BY

MOHAMMAD ZIA UL HAQUE
FLIGHT LIEUTENANT, PAKISTAN AIR FORCE
B.E. (AERONAUTICAL), UNIVESITY OF KARACHI, 1970

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN OPERATIONS RESEARCH

FROM THE

NAVAL POSTGRADUATE SCHOOL
MARCH 1976

ABSTRACT

Few if any validated guidelines exist for making decisions about the design, media, or format of new instructional products. This study examined strings of programmed learning responses to create general guidelines for making such decisions. Using a Markov model, tables were developed relating the expected proportion of students to be in a solution state at a given accuracy level and at a given level of confidence with respect to the length of response strings.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Looking at the past history of instructional material development it has been found that much initial effort was spent in generating material and selecting media. A number of decisions were required. Many problems such as the selection of format, mode of response, reinforcement etc. were to be solved during this initial period. If we look at some other fields having the same kind of problems, we see that these problems are being solved with the help of validated models and there are very few decisions left to be made. In the field of education, generally, and in instructional education, specially, we have not been able to find such models in existence, though much research should have been done in this area. If we look at the literature, there are indications that people feel the need of such a model (Smith & Murry 1975). Murril & Boutwell (1975) have commented that mathematical evidence and specific component justification of current instructional development methods lack in empirical verification. Baker (1973) has even suggested that much of the literature in instructional development prescribed procedures was based upon faith alone. A book edited by Mayer (1975) points out the importance of clearcut guiderules in the instructional design rules.

We can see very clearly that there is a fundamental problem in the field of instructional education. The absence of robust, active, validated models or set of guiderules to help the developer determine the best material and procedures for the student does and will continue to effect our standard of education.

Presently it would be unfair to say that our researchers have not paid any attention to this ever existing problem. Quite a few instructional programs have

been developed over the years, yet in each case the program developer had to create a unique model to answer the design questions for each program. Simple basic questions regarding the operations of the program had no ready answers available which were empirically based or validated. In the absence of readily available answers and since there was no method to conveniently simulate various outcomes to arrive at the answers, each program became an exercise in rediscovery through trial and error. As a result the model developed for a program became suitable only for that particular program and it was not possible to generalize it for other programs. This is the situation in which an instructional product developer usually finds himself.

If a model could be developed for instructional education, it would give the developer a system and a method for testing out and selecting various combinations of the product components in order to achieve desired target behavior. Components such as accuracy level, length of lesson, response rate, etc. could be arranged to result in the fastest learning at the least cost. A model like this should be specific to the outcomes rather than the content so that its basic alogrithms could be applied to many different programs. Each program can have a different arrangment of components depending upon the required outcome. If a model like this existed, it would have resulted in the early development of instructional programs and their speedy validation. The result would have been a tremendous saving of time and cost in the field of education.

In reviewing the general history of instructional development it can be seen that the absence of such models is one of the most overriding problems in the area of instructional education. The obvious problem then is that no model exists which has been tested and validated and is

generalizable to a variety of instructional products. The potential benefits to be derived from even a modest model are sufficiently great to place this problem in high priority category. The emphasis is being put upon the need for validated workable models or guiderules which can assist the instructional developer in the construction of teaching material and procedures.

At the Behavioral Sciences Institute, Carmel, California, considerable work is being done in this area. They have developed some models and are in the process of validating them. In an early study Madson (1972) attempted to form a model for language learning on the basis of a markov chain process. Oertel (1975) showed the nonexistence of any etiological factors. The author, in doing this work for arithmetic programming, is pursuing the same theory and is attempting to produce the guiderules which are so badly needed.

MODEL DEVELOPMENT

BACKGROUND

Before we go about developing our model it is necessary to review the events which started the development of such model. Since 1885 when some work was done by Ebbinghaus, experimental studies on learning have been recorded and reported in quantitative form. The first application of mathematics was seen for the purpose of describing empirical functions. A learning curve was the most common method of reporting results of a learning experiment. A graph representing the changes in the performance of a subject or group of subjects over successive practice trials for particular experimental conditions was the best bet. We have seen some of the analytic functions which were proposed to be the learning functions. Many arguments heard regarding these functions were that none of them was derived from fundamental considerations about the nature of learning. All of them were good with closest fit to the data usually obtained by the function that had more free parameters.

In 1919 Thurstone set up a system of axioms based on psychological considerations that led to the derivation of rational learning functions. A very specific set of psychological identifications was used as the parameters. Moreover Thurstone was the one to suggest a probabilistic approach. He took as his aim the derivation of the probability of a correct response as a function of trial numbers. The same theory was later extended to the analysis of discrimination learning and transposition by Gulliksen and Wolfle (1938). However, only mean response curves were considered and no attention was paid to the prediction of response distributions and sequential statistics. Moreover no proceedures were devised for parameter estimation and no experiments were done to find the validity of the parameters of the model. Another group of experimenters attempted to derive learning curves from simplified conceptual models of

the nervous system but their efforts did not have any significant impact on experimental investigation of learning.

The picneer of theoretical learning was Clark Hull. In his major work, Principle of Behavior (1943), a number of postulates were stated which dealt with a number of variables that had not been identified in the earlier experiments. The postulate in many cases was simply a generalization of empirical results. It was hoped that the aggregation of postulates would jointly imply much more than the specified experimental facts from which they are individually derived. Hull aimed for comprehensiveness in his theory partially due to its relative clearity and generality. The theory stimulated considerable experimental research. It has gone through a variety of modifications and still guides the research of many contemporary experimenters. The most important contribution by Hull was the statement of a rich collection of qualitative concepts and propositions, some of which have had a lasting influence on the thinking of psychologists.

Somewhat later many other researchers started formulating their stochastic models for learning. At the same time another group worked in developing what has come to be known as Linear Models for learning. The basic idea for linear models is very simple. In a two-choice learning experiment, the probability that the subject will make response 1 on trial n is p_n. On each trial the subject responds and some reinforcing event is provided. If reinforcement event j occurs on trial n the new value of response probability on trial n+1 is

$$p_{n+1} = a_j p_n + b_j$$

this equation expresses the new value of response probability as a linear function of its old value. The parameters a_j and b_j specify whether event j effects an

increase or decrease in p_{n+1}.

At the same time work was being done on markov chain models with fewer states and they represent an especially promising line of theoretical development. The basis of original development was a paper by Estes (1959). Basic to this formulation is the idea that a subject s response probability can take on only a fixed set of values and that reinforcing events produce transitions from one value of response probability to another.

It has been proposed that performance in the experimental situation can be represented by three discrete performance levels: o, p,and 1. In these terms learning consists of two all-or-none transitions from lower to higher levels of response probability. This notion was originated by Estes who also introduced the technique of representing learning by markov chain. It was because of Estes prior theoretical work that we were led to examine our data for evidence of an intermediate performance level. In truth we have been astonished by the consistency with which such evidence has apperared throughout the range of data examined.

It will be noted that the evidence comes from experimental situations in which initially the probability of a correct response is zero and asymtotically it is unity. Such zero to one situations possess an important advantage for our method of data analysis. The arrangement enables one to identify responses between the first success and last failure as occurring in the intermediate state. The importance of this identification can be understood if one imagines trying to test decisively the notion of a single intermediate state for a learning situation in which the initial response probability is greater than zero or the asymtote is less than unity, or both. In such cases the

evidence has to be of a more indirect nature like predicting quantitative details of a variety of statistics. We know that data showing an intermediate performance level can be interpreted within the framework of stimulus sampling theory. Facts about intermediate performance level can also be interpreted in terms of multistage models of Restle and Greeno (1970) In constructing and testing the three-stage model, we have suppressed the stimulus sampling rationale and have presented simply a descriptive model about learning.

The learning model exploits the notion of an intermediate state in an obvious way. Certain general markovian properties were imposed regarding transition probabilities among the states, and the resulting model provided a fairly adequate description of the data on which it was tested. The specific form of the model is not arbitrary entirely since we had been able to reject various plausible alternative three-stage models because one of the models we have tested permits a direct, one-trial transition from the starting state to the terminal absorbing state. This alternative is diagramed in Figure 1 . Here it is assumed that with probability (1 - d) the subject skips the intermediate p state going directly to state 1. The alternative classes of learning models which can be considered are the continuous or incremental theories such as the linear operator models. Although extensive comparisions have not been undertaken, it seems evident that all contiuous models will be rejected for this kind of data. In particular, from continuous models one would expect performance to improve monotonically over trials between the first success and last error. Such upward trends simply failed to materialize in any of the studies. Our test for such trends were the CHI Square and the rank order correlation between intermediate trials and response probabilities. In none of many cases considered was this

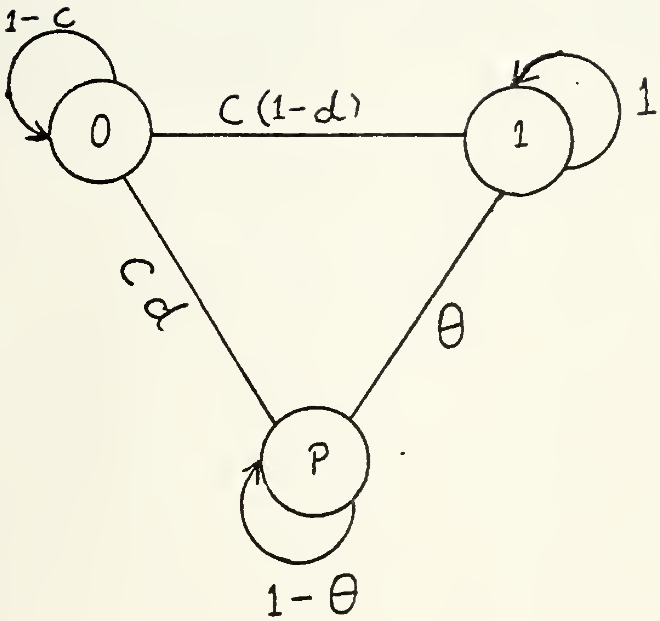


Fig 1. A Three Stage Model

correlation significantly different from zero, a result in line with the stationarity assumption. It might be objected that possible effects on the intermediate responses of individual differences in learning were not considered. To answer this objection experiments were conducted by Bush and Mosteller (1955). Two points were made from the results observed. One, that the argument of selection artifacts does not really rescue the continuous models from the stationarity data and two, that the statistical tests we routinely use to assess stationarity of intermediate responses have considerable power to reject the null hypothesis when it is false.

MODEL DEVELOPEMENT

A brief review of mathematical learning theory by Atkinson, Bower, and Crothers (1965) indicates that learning as probability models started in 1919. From 1919 to 1950 there were quite a few models proposed and tested. All of them were specific to certain learning situations. From 1950 onward there has been much work done in the area of stochastic learning. This resulted in two theories. the linear model and the markov model. The linear model basically depends upon the theory that the probability of success for a subject is given by the equation

$$P_n = 1 - (1-P_1)(1-\theta)^{n-1} \cdots\cdots\cdots (1)$$

where p_1 is initial probability of success and θ is his learning rate.

The markov model depends upon a different theory which states that if a subject is in an unlearned state (u) then the probability of a correct response is g(guess). If the subject is in the learned state(L), then the probability of correct response is 1. the probability of going from the unlearned state to the learned state on any presolution trial is c. The probability of a correct response on any trial n is given by

$$P_n = 1 - (1-g)(1-c)^{n-1} \cdots\cdots\cdots (2)$$

a comparision of equations (1) and (2) indicates that their forms are exactly the same. The difference in these equations lies in their theoretical background and the meaning of the parameters. Equation (1) states that a subject starts with a probability p_1 of making a correct response on the first trial. The probability of success on the second trial is greater due to incremental learning achieved on the first trial. The linear process continues

indefinitely and the subject s probability of success approaches 1 asymptotically. Equation (2) states that on each presolution trial a subject has a probability c of going into sclution. Once in solution the subject stays in solution and always responds correctly and this probability remains constant. The form of these two equations are compared by Restle and Greeno (1970). Based on their analysis it is stated "...the all-or-none theory is most interesting and we think it is the one most deserving of future work ".

Pilot research involving a computer simulation of the linear model suggested that it is inappropriate for mathmatical learning. The study of data from students showed that the markov principles of stationarity and independence are applicable to this program. Based on these results this work was done considering Markovian (all-or-none) principle.

ASSUMPTIONS

For the developement of the model, the following assumptions are necessary

1. The learning process is Markovian in nature

2. The subject can be correct on the first trial of any step by either (a) being in solution prior to the trial, (b) going into solution because of the information presented in the first stimulus or (c) guessing correctly in presolution. This assumption modifies equation (2) in that equation (2) contains the restriction that for the subject to be correct on the first response, he must guess correctly. therefore it does not allow the possibility of being in solution (the learned state) on the first trial. Allowing for the possibility that the subject is in solution on the trial (Atkinson, 1965) appears to be a more realistic approach and was used in this work.

3. The g factor in presolution is a function of step and the subject.

4. The c factor is a function of step and the subject.

5. g and c are constant over any step for a given subject.

6. The set of outcomes form a homogenous markov chain

$$\begin{array}{cc} & \begin{array}{cc} L_{n+1} & U_{n+1} \end{array} \\ \begin{array}{c} L_n \\ U_n \end{array} & \begin{bmatrix} 1 & 0 \\ c & 1-c \end{bmatrix} \end{array} \qquad \begin{array}{l} n = 0, 1, 2 \\ g \leq 1,\ c \leq 1 \end{array} \qquad \begin{array}{c} P(\text{Correct} \mid \text{Row State}) \\ \begin{bmatrix} 1 \\ g \end{bmatrix} \end{array}$$

MODEL

The equations developed in this work are based on the work done by Atkinson, Bower, and Crothers (1965), Coombs (1970), Restle (1970), Gray (1972), and Madson (1972). Since it is difficult to give credit to one source, only the equations are given with explanations. The first important thing is the probability of a correct response given that the subject is in an unlearned state (u). This state is assumed on the first trial and known to exist if an error occurs before reaching the advancement criterion. If no error occurs then there is no way to find out whether the subject was in learned state (L) or was in unlearned state and performed as follows

$$P(\text{CORRECT}) = C + g(1-c)c + g^2(1-c)^2c + \cdots\cdots \quad (3)$$

$$\lim_{n \to \infty} P(\text{CORRECT}) = \frac{c}{1-g(1-c)} = \rho \quad \cdots\cdots\cdots\cdots (4)$$

in the future whenever we refer to this probability we shall call it rho, the probability of errorless response given that the subject is in the unlearned state. The above equation says that either the subject goes into the learned state on the first trial, stays in the unlearned state and guesses correctly and then goes into learned state, or stays in the unlearned state twice, guesses correctly twice and then goes into the learned state, etc. The development indicates that the subject goes into the learned state eventually if errorless response is achieved after an error. The reader familiar with markov theory will note that the term relating to remaining in the unlearned state and having errorless responses was omitted in developing equation (4). The omission was committed since the term

$$g^n(1-c)^n$$

goes to zero in the limit as n approaches infinity.

The next development will be the expected number of

errors given g and c. The probability that the total number of errors is k is

$$P[E=K] = \sum_{i=0}^{\infty} \binom{K+i}{i} g^{i} (1-c)^{i} (1-g)^{K} (1-c)^{K} c$$

This would represent every feasible combination of events in which exactly k errors can occur. By using standard mathematical tables we can reduce the equation to the following

$$P[E=K] = \left[1 - \frac{c}{1-g(1-c)}\right]^{K} \left(\frac{c}{1-g(1-c)}\right) \cdots\cdots\cdots\cdots (5)$$

$$= (1-\rho)^{K} \rho$$

In words equation (5) gives the total number of response strings required untill the last error and after that the subject is in the learned state.

Since the probability of an errorless response string is rho, given that the subject is in an unlearned state, it follows that the error response is (1 - rho). This takes into account all possible numbers of correct responses before the error response which breaks the string. The occurence of an error demonstrates the unlearned state and also allows for another possible string of errorless responses which is independent of the length of previous strings and depends only on being in the unlearned state.

The next developement is the expected trial number of last error. The probability that the last error occurred on trial t equals

$$p(T=0) = \text{rho}$$

$$P[T=t] = (1-c)^{t} (1-g) \rho \cdots\cdots\cdots (6)$$

$$t=1,\ 2,\ 3, \ldots\ldots$$

In words equation (6) says that there were t trials in the unlearned state indicated by an error on trial t and then errorless response. The probability statement allows for any

sequence or number of correct and incorrect responses up to trial t. The only required knowledge is that an error occurred on trial t and then no more errors.

To find the expected value of t

$$E[T] = \sum_{t=0}^{\infty} t\, P[T=t] = e(1-g)(1-c) \sum_{t=1}^{\infty} t\,(1-c)^{t-1}$$

$$= \frac{(1-g)(1-c)}{[1-g(1-c)]\,c} \qquad \text{As} \quad e = \left[E[E]+1\right]^{-1}$$

$$\text{So} \quad c = e\,\frac{E[E]}{E[T]} \qquad \text{OR} \quad \hat{c} = \frac{E}{T[E+1]}$$

solving by using previous relations

$$\hat{c} \simeq \frac{1}{T}$$

so this equation says that c is approximately the inverse of the trial number of the last error. This is intuitively appealing as it states that the larger the factor c (probability of going into solution) the fewer the expected number of trials.

VERIFICATION OF MODEL

Subjects

All subjects from whom data were obtained for this analysis were public school students. They attended classes for the educationally handicapped in the state of Pennsylvania. All were going through the Monterey Arithmetic Program which was developed by Behavioral Sciences Institute in Carmel, California. The number of subjects used in this analysis was 48. There were 20 girls and 28 boys. The age range was between 5 and 11 years. Their IQ ranged from 60 to 80. The subjects were randomly selected for analysis by the supervisor in Pennsulvania. There was no effort to constrain subject selection by age, sex, etiology or any other parameter.

Data Source

The subjects were given problems to solve. Depending upon what subprogram they were in , they performed addition, subtraction, multiplication or division. When a subject completed a problem it was checked by a teacher for accuracy. Depending upon the outcome it was marked as a correct or incorrect response. Thus, for the purposes of this study, each problem which was worked was counted as one response and each lesson was comprised of a sequential string of responses.

The total number of responses was 3000. For any subject the sequence of responses generated in a single lesson consisted of two parts. First, a string consisting of correct and incorrect responses and second, a string of 10 continously correct responses. Some of the response strings were not used in the analysis. The string of continous correct responses indicates a solution state and since we were considering only the presolution state, the string of continous correct responses was not utilized. There were 480 responses in this category. The situations where the subject started with correct responses and did not make any error indicated that the subject was already in the solution state. The responses in situations like this were not used. The number of responses of this kind was 320. In situations where the subject did not complete the lesson, he gave us no indication of the number of responses necessary to go into solution state. We were also unable to use those responses. The number of responses of this type was 1196. After disregarding all those responses mentioned above we were left with a total of 1004 responses which comprised 48 strings of correct and incorrect responses (lessons). Thus each subject contributed one response string to the data pool.

Program

The arithmetic program consists of material and procedures which are specially designed for the purpose of achieving a high degree of skill and accuracy in the computation of arithmetic problems. It is divided into four subprograms of addition, subtraction, multiplication, and division. Each subprogram consists of 42 steps. These steps are in increasing order of difficulty. The first step is very basic and the last step is most difficult. A subject completing the last step is considered capable of performing all the calculations of that subprogram. This program is designed to be used in a classroom but it can be administered on an individual basis. It is useful for both kinds of students, those who did not have any arithmetic before and those who had had it but could not achieve the required accuracy level. This program is applicable to all students of all ages and takes into consideration all kinds of differences which occur among them. It uses a locator test which helps the teacher to place each student at the appropriate location in the program. It also uses an automatic branching proceedure which takes care of slow learners. This program is built in such a way that the teacher can respond equally to both remedial and developmental students.

ANALYSIS and RESULT

The raw data consisted of 48 strings of correct and inccrrect responses. For this analysis values of 0 and 1 were assigned to correct and incorrect responses, respectively. The data are shown in appendix A. As the basic characterstics in Markov chain process are independence and stationarity and since other aspects of performance are closely related to these properties, it was decided to test the data for these two characterstics. The proceedure for the tests was the same as proposed by Oertel (1975) for pooled data. Independence was tested by calculating for each subject the observed frequency of the four possible combinations (1-1, 1-0, 0-0, 0-1) and then computing the value of Chi Square by appropriate formula for a 2 x 2 contingency table (incorporating the correction for continuity). Whenever the subjects had cell entry less then 5, the data were combined with as many adjacent subjects as necessary to get a frequency of at least 5. The Chi Square values were then summed . The results are shown in Table 1 and the observed values in appendix B. The table shows that the data has the property of independence.

For testing stationarity the proportion of correct responses in the first and second halves were compared. The difference in proportions for each subject was tested by a direct difference t test. The results are in Table 2 and it establishes the property of stationarity.

Once the properties of independence and stationarity were confirmed, the next step was to find the distribution of L (number of responses). To find the distribution a histogram was plotted (appendix C). The distribution appeared to be exponential. A Chi square goodness-of-fit test was used to test the null hypothesis that the distribution was exponential. The test did not reject the

null hypothesis. Calculations are shown in appendix D. Since the data was discrete, it was decided to test the data for having a negative binomial or a geometric distribution. A Kolmogorov-Smirnoff goodness-of-fit test was done to find the distribution. The result of the test are shown in Table 3, and the linear relationship between the observed and generated data is shown in appendix E. From the table we can see that the data best fits the Geometric distribution with q = 0.96. This gives c the maximum absolute difference in comulative distribution function = 0.12 and the probability of occurance is 0.7167. The value of alpha for the test was 0.1. Once the distribution was confirmed we were able to predict the percentage of students in the solution state for any given number of responses using the cumulative distribution function table shown in appendix F. The values of L (number of responses) for different percentages are given in table 4.

The next step was to find the estimated value of the parameter c. From our theoretical background we know that c is approximately the inverse of the expected number of incorrect responses T. To find the expected value of T for any given number of responses a regression analysis was carried out between T and L. The result was a linear equation with a value of r = 0.8673

$$L = 4.8T + 3.3$$

The expected values of L for any given T are shown in table 5. Similarly, expected values of T for different L are shown in the same table. Hence for any L we were able to find the value of T and so the value of C. The values of L, T and C for different accuracy levels (Q) are given in table 6.

The next step was to find some kind of representation

or trend from the number of incorrect responses within the first 10, 15 or 20 responses. This was attempted to enable us to predict the expected number of responses from a subject to reach the solution state and to find a branching criterion. The relationship of the density, sequence, and patterning of incorrect responses to the total number of responses was examined. Unfortunately we were unable to find any significant trends or relationships.

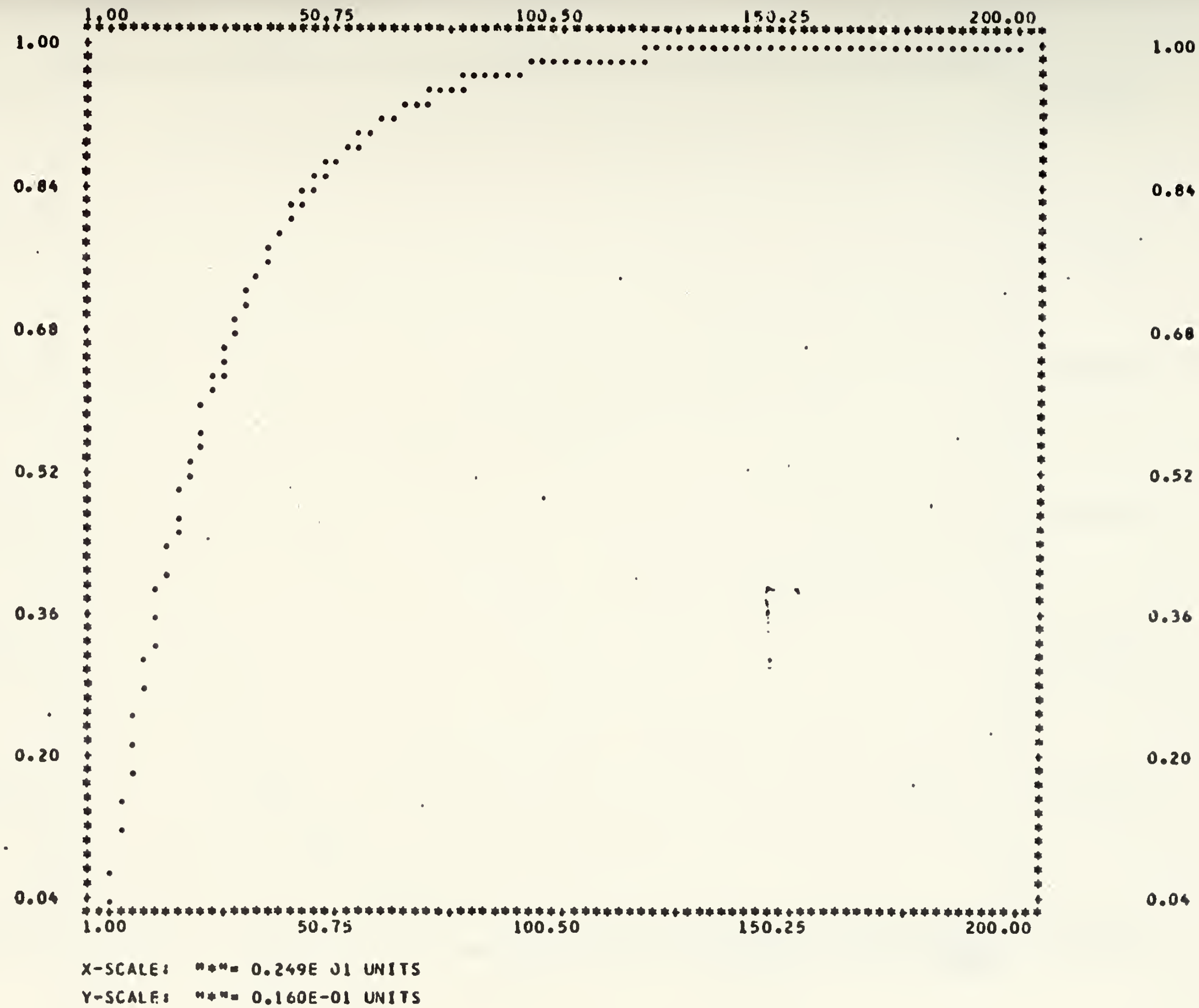


FIGURE II. Graph of Cumulative Distribution Function

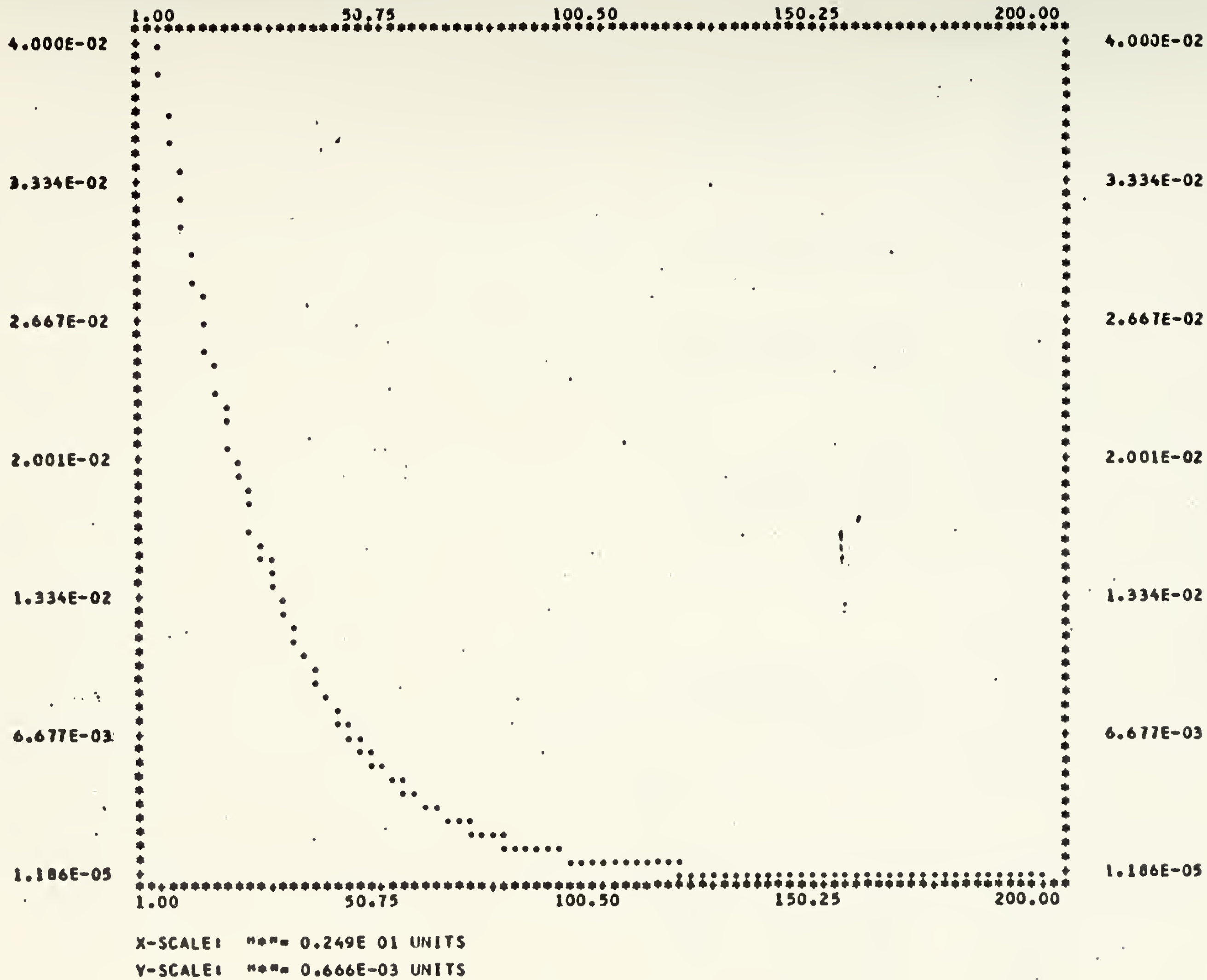


FIGURE III. Graph of Probability Distribution Function

Table 1

Chi-square values for independence of transition probabilities

subject	A	B	C	D	Chi square value
1-9	5	25	33	30	.0027
10	10	8	9	30	.0010
11-24	12	45	57	165	.000044
25-40	8	43	53	176	.00011
41-48	5	23	29	116	.000034
					total .00388

Table 2

Tabulated values of the proportion of correct responses in first and second half and the values of a direct difference t test

subject	1st half	2nd half	diff
1	3/3	3/3	0
2	5/6	5/6	0
3	6/7	6/7	0
4	14/18	15/18	1
5	16/23	15/23	1
6	5/5	4/5	1
7	3/4	3/4	0
8	3/3	2/3	1
9	4/5	4/5	0
10	3/5	4/5	1
11	10/11	9/11	1
12	24/27	24/27	0
13	3/8	3/8	0

14	6/10	9/10	3
15	5/5	5/5	0
16	12/13	11/13	1
17	20/20	20/20	0
18	2/2	2/2	0
19	2/3	2/3	0
20	3/7	5/7	2
21	7/8	7/8	0
22	2/2	1/2	1
23	13/16	14/16	1
24	2/2	1/2	1
25	4/5	4/5	0
26	6/7	5/7	1
27	19/22	13/22	6
28	2/4	3/4	1
29	2/2	1/2	1
30	7/7	5/7	2
31	15/18	15/18	0

32	5/8	4/8	1
33	21/29	18/29	3
34	7/10	8/10	1
35	14/20	14/20	0
36	13/17	12/17	1
37	16/19	13/19	3
38	5/5	3/5	2
39	12/15	10/15	2
40	5/7	5/7	0
41	5/5	4/5	1
42	5/6	5/6	0
43	5/6	5/6	0
44	4/4	3/4	1
45	30/34	30/34	0
46	1/1	1/1	0
47	16/24	17/24	1
48	6/7	6/7	0
total	392/488	372/488	20

$t(\text{observed}) = 1.45$

$t(\text{critical}) = 2.01$

Result: The data had the property of stationarity

Table 3

Kolmogorov-Smirnoff goodness-of-fit test for the number of responses (L) to the Negative binomial and Geometric distributions

distribution	parameter	c	p
negative binomial	aipha=27.36 K = 0.91	0.98	0.00000
geometric	q = 0.85	0.54	0.00000
	q = 0.95	0.14	0.5487
	q = 0.96	0.12	0.7167 *
	q = 0.97	0.20	0.1786
	q = 0.99	0.52	0.0000

c = absolute difference in c.d.f.   p = prob. of occur.

Table 4

Tabled values of the number of responses (L) required for a given percentage of students to be in the solution state at a specific level of confidence

in solution (percent)	confidence level (percent) 80	90	95	99
50	5	6	7	9
60	10	11	12	14
75	23	24	26	29
80	28	29	31	36
85	35	37	40	47
90	47	50	55	70
95	63	69	82	>200
96	69	76	96	>200

Table 5

Tabled values of the expected number of errors (T) and the total number of responses (L) given T or L

T to L		L to T	
T	L	L	T
1	8	10	1
2	13	20	3
3	18	30	5
4	22	40	7
5	27	50	9
6	32	60	11
7	37	70	13
8	42	80	15
9	46	90	18
10	51	100	20

Table 6

Tabled values of T, C, and L for a given percentage of students in solution and a given accuracy level

	percentage in solution											
	50		75		80		85		90		95	
q	t	c	t	c	t	c	t	c	t	c	t	c
.5	3	333	12	083	14	071	18	055	25	040	34	029
.4	2	416	10	104	12	086	15	067	20	050	28	036
.3	2	555	7	139	9	115	11	090	15	067	21	048
.25	1	999	6	167	7	138	9	067	12	080	17	058
.2	1	999	5	208	6	172	7	135	10	100	14	072
.15	0	999	4	277	4	230	6	180	7	133	10	097
.1	0	-	2	416	3	345	4	270	5	200	7	145
.05	0	-	1	999	1	690	2	540	2	400	3	290
1		5		22		27		34		45		60

Q = (1-p), probability of incorrect response

DISCUSSION and SUMMARY

The basic idea behind this work was to develop some guidelines to help the designer of the learning program in deciding, before the program is run, the required amount of work to be performed by the students and the teacher. The ability to make this decision validly would be helpful in speeding learning and cutting down the costs. For these reasons model verification was required. First of all the data was observed to see the kind of process that would be useful. As we know there are two kinds of models in existence, the linear model and the stochastic model. It was especially necessary to see whether the data agreed with the stochastic model, since there are certain parameters--namely L, T, C--which, if determined correctly, would enable us to predict values which are very close to observed values. The work done by Oertel had shown that this was possible. So our main emphasis was to establish first that the data is a product of Markov process and then to find these parameters.

As shown in the analysis, we were able to describe the learning process to be a Markov process by testing for stationarity and independence. Once these properties were established, we were able to use all the assumptions mentioned earlier. The distribution, once found, enabled us to predict the expected number of responses required for any given percentage of students to be in the learned state. This would help the designer of the program to determine his requirement for the number of problems, depending upon his target of achievement.

The next step was to determine the values of the parameters t and c. The linear regression equation helped us in predicting the expected number of incorrect responses when the total number of responses was known. If the designer of the program can determine the number of

responses required to be in the solution state, he could determine a branching criterion easily. The rule could be made that if a subject made more than a specified number of incorrect responses, he should be branched. Once the value of T was found, it was an easy step to find the value of C. These values can be used to calculate different probabilities as shown in the theory.

In the next step we tried to find some kind of representation of incorrect responses. This was done in order to be able to predict the students to be branched by observing the first 10 or 15 responses. This was done by different methods such as density, pattern, and frequency. Unfortunately we were unable to find any significant trends. The reason for not finding the trend could be that there is none, but it could also be that we did not have a sufficient number of response strings.

It is suggested that if further work is done in the future then the data to be collected should beat least four- or fivefold of the present data. If with that data trends are still not visible, it will suggest that they donot exist, however if a trend is observed, it would be a great help to the designer of program for determining the branching rule right after the few initial responses. As stated this would save much effort and time of both students and teachers and would be a major factor in reducing the cost of running the program.

Appendix --A

Raw Data

0 0 0 1

1 0 0 0 0 0 0 1 0 0 1

0 0 1 0 0 0 0 0 1 0 0 0 0 1

0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1
0 1 1 1 0 1 0 1 0 1 0 0 1 0 1

0 1 1 0 0 0 0 1

0 0 0 1

0 0 0 0 0 0 0 0 0 1 0 0 0 1

0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0
0 0 1 0 0 0 1

0 0 0 0 1 1 1 0 0 1 0 0 1 1 0 1

0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 1 1 1 1 1 1
1 1 1 0 0 1 1 0 0 0 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1

1 0 1 1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 1

0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 1 1 0 0 0
1 0 1 0 0 0 0 0 1 0 1

0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1 0 1 1 1 0 1 0 1 0
1 0 0 1 0 1

0 0 0 0 0 0 0 1 0 0 1 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0
0 0 0 1 1 0 1

0 0 0 0 0 0 0 1 0 1

0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 1 0 1 0 0 1 0 0 1 0 0 0 1

0 0 1 1 0 0 0 0 0 0 1 0 1 0 1

0 0 0 0 0 0 0 0 1

0 0 1 0 0 0 0 0 0 0 0 1

0 0 0 0 1 0 0 0 0 0 0 0 1

0 0 0 0 0 0 1

0 0 0 0 0 1 0 0 0 1 0 0 1

0 0 0 1 0 0 0 0 0 0 0 0 0 1

0 0 0 0 1 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0
0 0 1 0 0 0 1

1 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 1 1 0 0 0 1 1 0 0 0
0 0 0 1 1 1 0 1 0 0 0 1 0 1 0 0 1

0 0 0 0 0 0 0 0 0 1

0 0 0 1 0 0 0 1

0 0 0 0 0 1

1 0 0 0 0 0 0 0 0 1

1 0 0 0 1 0 0 0 0 1

0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 1

0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0
0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 1

1 0 0 0 0 0 0 1

0 0 1 1 0 1 0 1 0 0 0 0 0 0 0 0 1 0 0 0 1

0 0 0 1 0 0 0 0 0 0 1

0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1

0 0 0 0 1

0 1 0 0 0 1

1 0 1 0 1 1 0 0 1 0 0 0 0 1

1 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 1

0 0 0 1

0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0

0 1

0 0 0 0 0 0 0 1

0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0
0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0
0 0 0 0 0 0 0 1

1

0 0 1

0 0 1 0 0 0 0 1 0 0 1 1 0 1 0 0 0 0 1 0 0 1 1 0 0 0 0 0 1 1
0 1 1 0 0 0 0 0 0 0 0 1 0 0 0 0 1 1

Appendix --B

Frequencies of (1-1, 1-0, 0-1, 0-0) sequences

frequencies of sequences

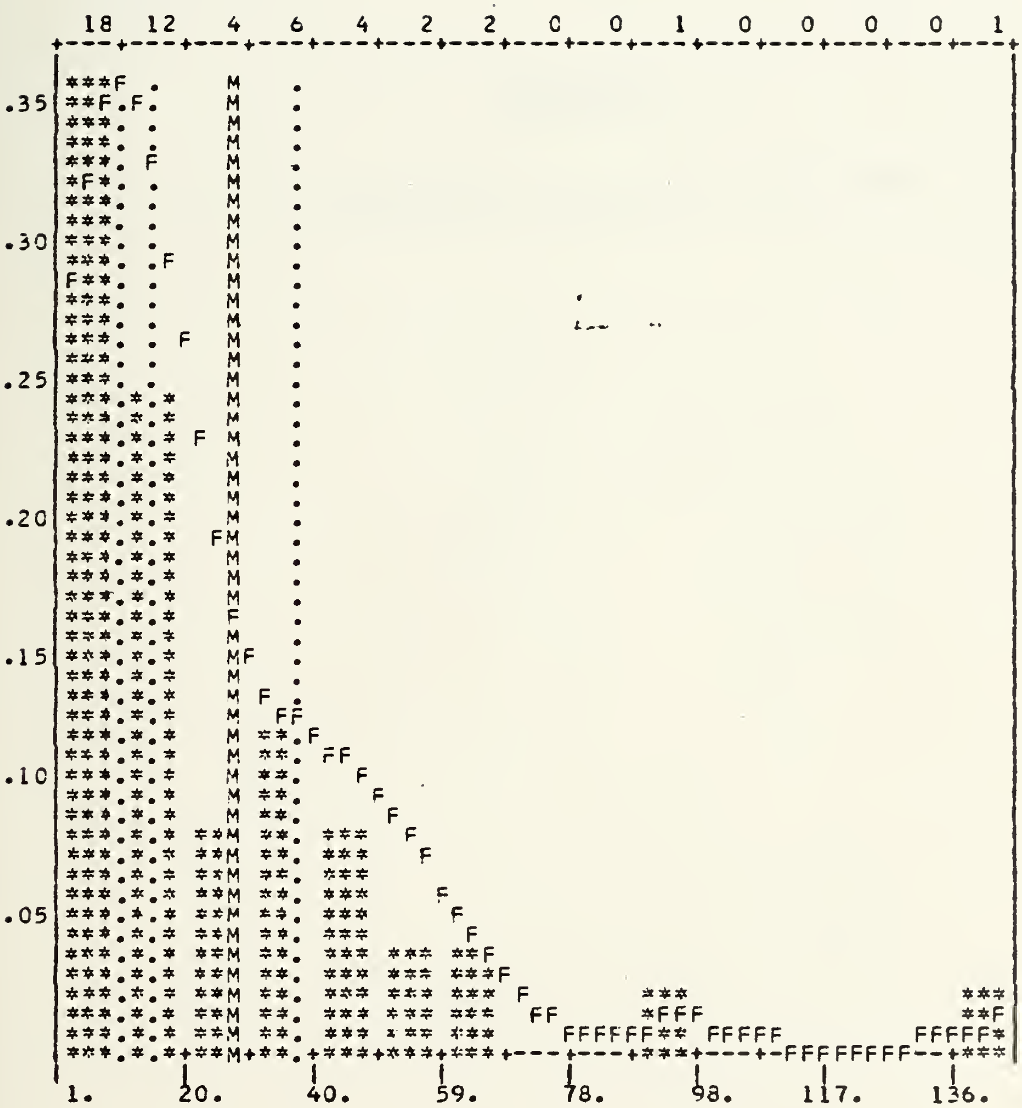
1	0	0	1	2
2	0	2	2	6
3	0	2	3	8
4	2	10	11	21
5	1	1	2	3
6	0	0	1	2
7	0	1	2	10
8	0	6	7	23
9	2	3	4	5
10	10	8	9	30
11	1	4	4	11
12	6	6	7	21
13	1	8	9	16
14	2	5	6	23
15	0	1	2	6
16	0	8	8	12
17	1	3	4	6
18	0	0	1	7
19	0	1	2	8
20	0	1	2	9
21	0	0	1	5
22	0	2	3	7
23	0	1	2	10
24	0	5	6	24
25	4	11	11	20
26	0	0	1	8
27	0	1	2	4
28	0	0	1	4
29	0	1	1	7
30	0	2	2	5
31	0	2	3	16
32	0	6	7	41
33	0	1	1	5

34	1	4	5	10
35	0	1	2	7
36	0	2	3	20
37	2	12	12	30
38	0	0	1	3
39	0	1	2	2
40	1	4	4	4
41	0	3	3	10
42	0	0	1	2
43	0	4	5	22
44	0	0	1	6
45	0	7	8	52
46	0	0	0	0
47	0	0	1	1
48	5	9	10	23

Appendix --C

Histogram of the data

FREQUENCIES SAMPLE SIZE = 50



Appendix --D

Chi square goodness-of-fit test

Chi. Sqr. goodness of fit test

H_0= the distribution is exponential

H_1= the ditribution is not exponential

alpha = 0.1

interval	1-exp(-alpha x)	theo freq	obs freq	dif
10	0.33	16	18	2
20	0.55	10	9	1
30	0.70	8	7	1
40	0.80	4	5	1
50	0.865	4	4	0
60	0.91	2	1	1
70	0.94	1	1	0

Chi. Sqr. = 1.1125

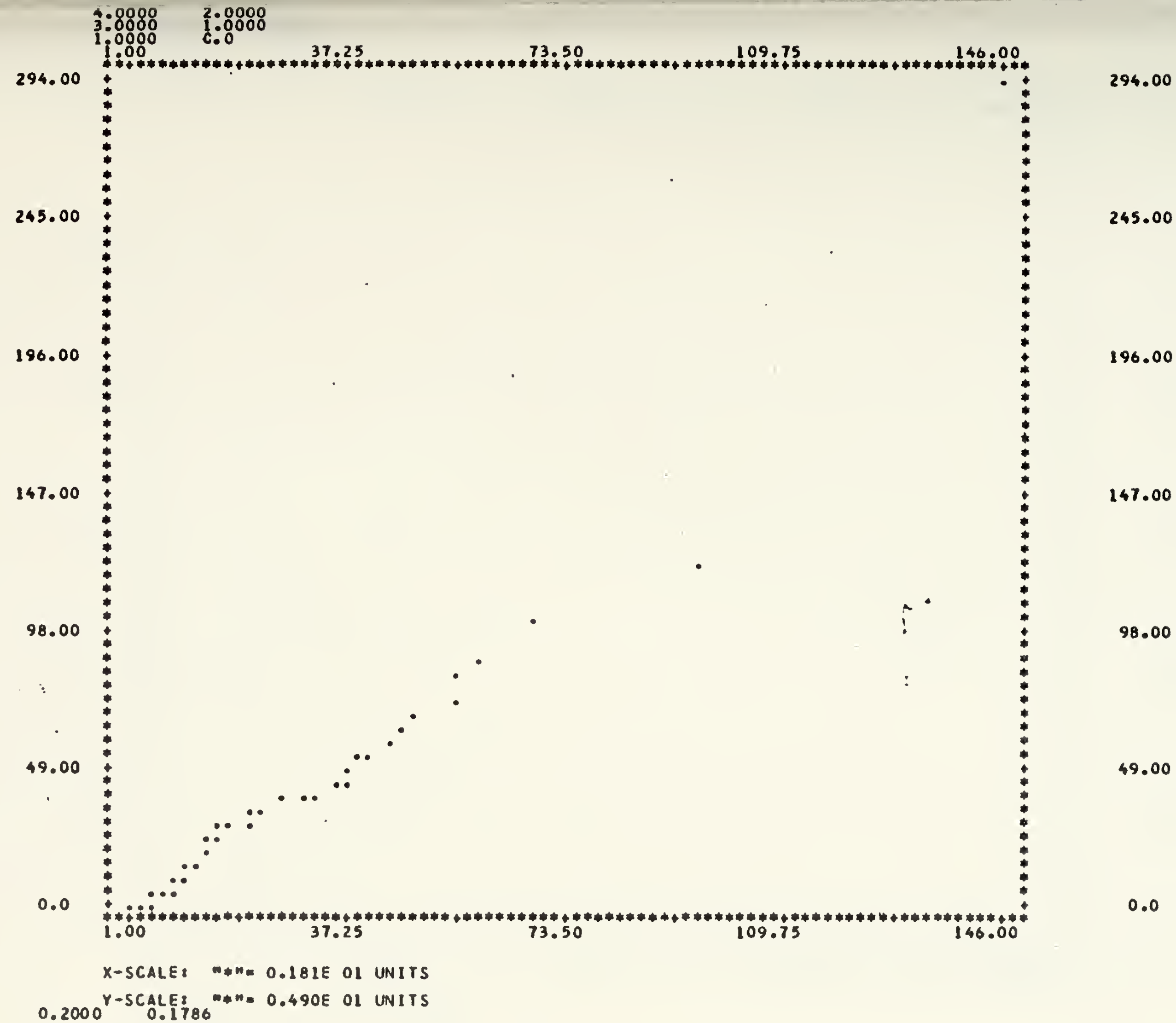
Chi. Sqr. (.05) = 1.64

df=6

Result: accept H_o

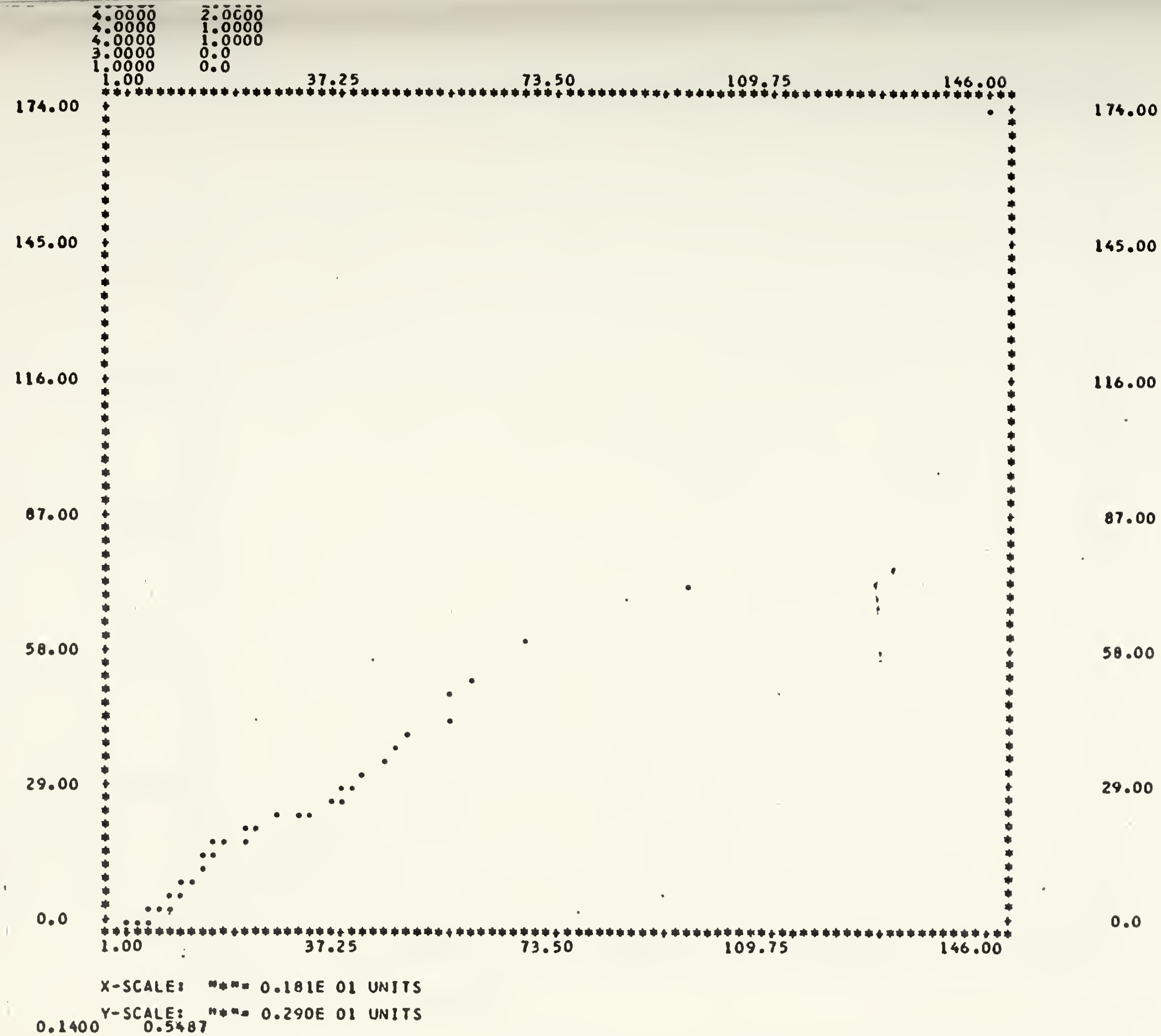
Appendix --E

Graphical representation of the linear relationship between observed and generated data


4.0000 2.0000
3.0000 1.0000
1.0000 C.0
1.00
37.25
73.50
109.75
146.00
294.00
245.00
196.00
147.00
98.00
49.00
0.0
X-SCALE: "*"= 0.181E 01 UNITS
Y-SCALE: "*"= 0.490E 01 UNITS
0.2000 0.1786

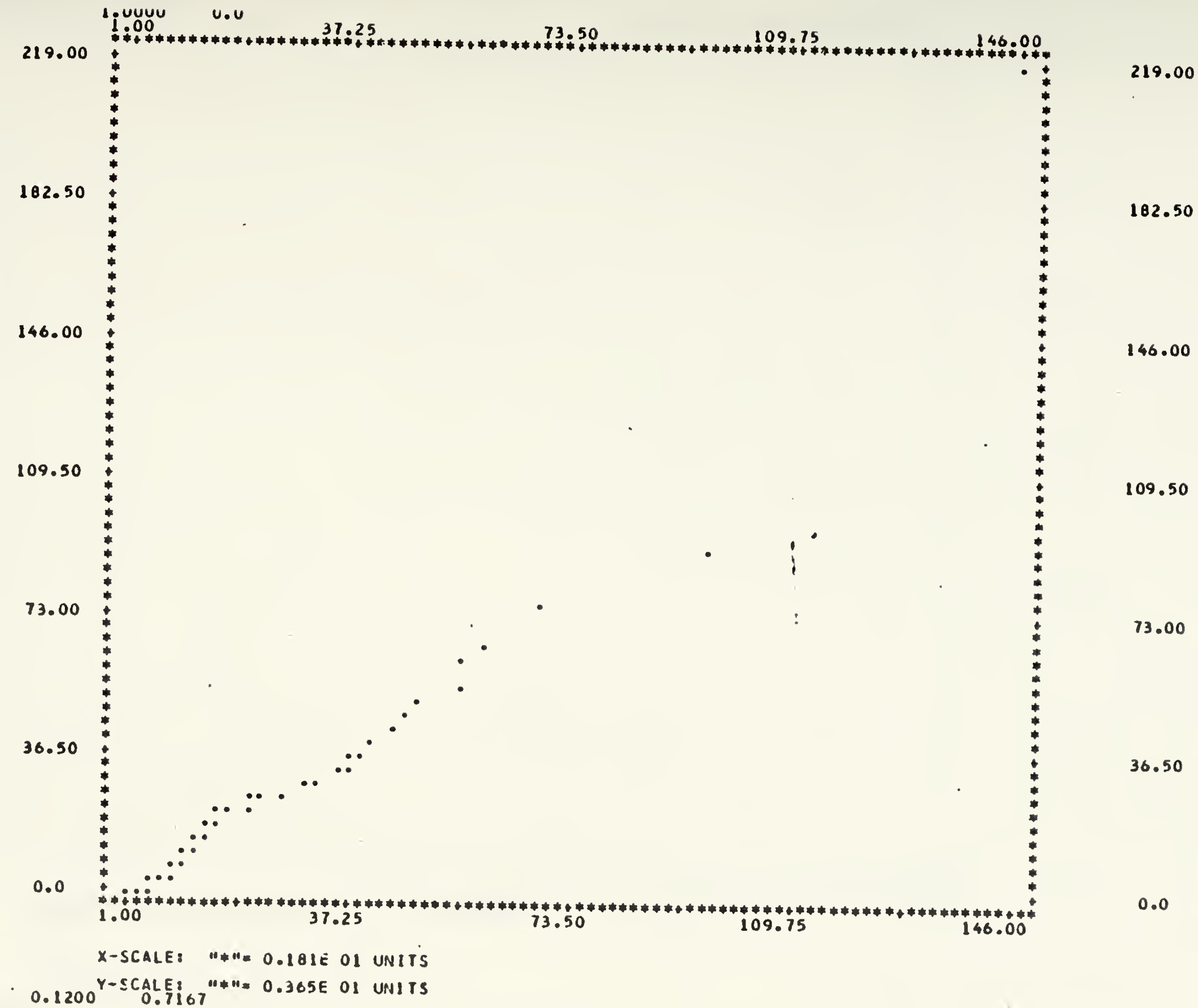
Q=0.85

(i)


4.0000 2.0000
4.0000 1.0000
4.0000 1.0000
3.0000 0.0
1.0000 0.0
1.00
37.25
73.50
109.75
146.00
174.00
145.00
116.00
87.00
58.00
29.00
0.0
X-SCALE: "*"= 0.181E 01 UNITS
Y-SCALE: "*"= 0.290E 01 UNITS
0.1400 0.5487

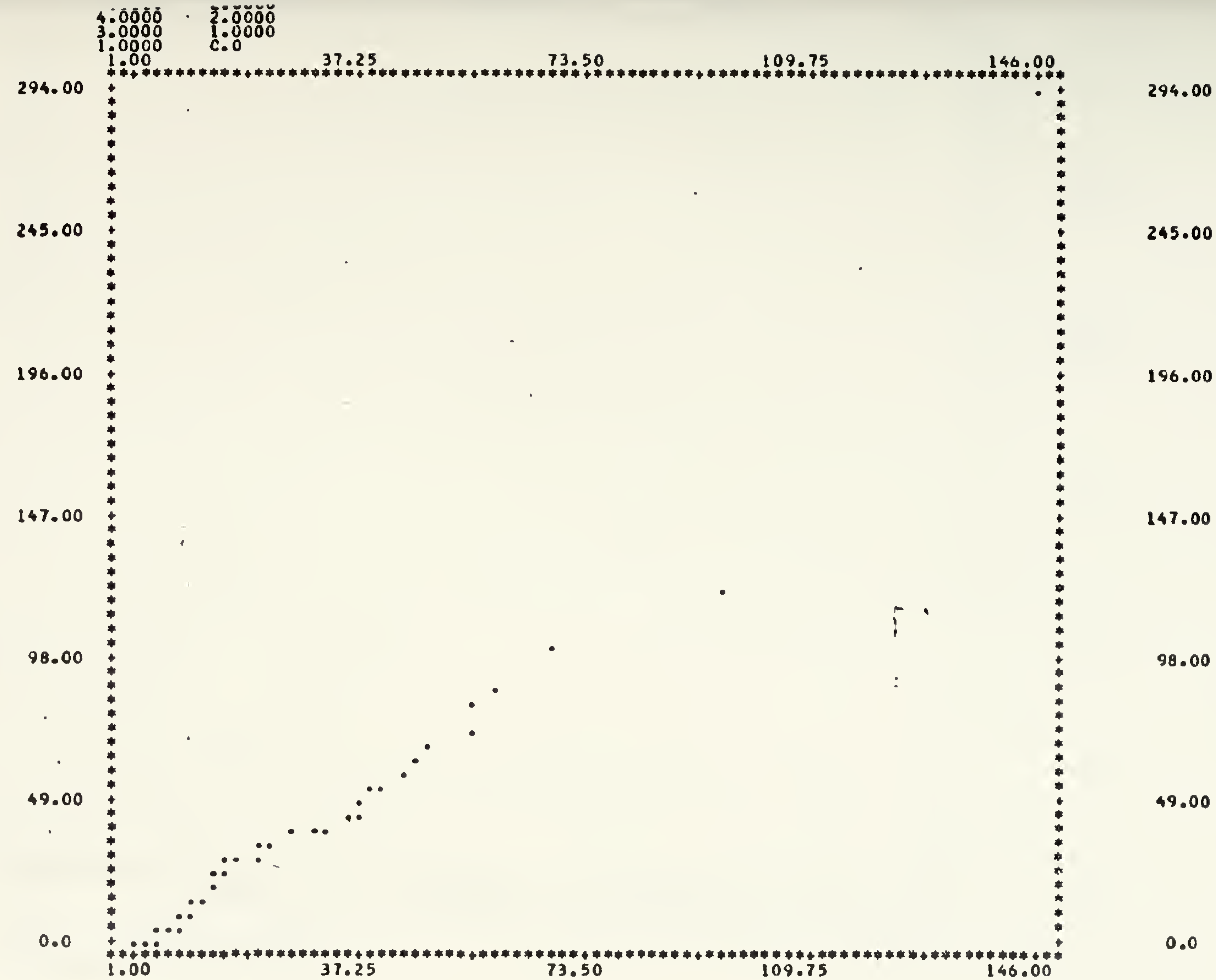
Q=0.95

(ii)


1.0000
0.0
1.00
37.25
73.50
109.75
146.00
219.00
182.50
146.00
109.50
73.00
36.50
0.0
X-SCALE: "*"= 0.181E 01 UNITS
Y-SCALE: "*"= 0.365E 01 UNITS
0.1200
0.7167

Q=0.96

(iii)



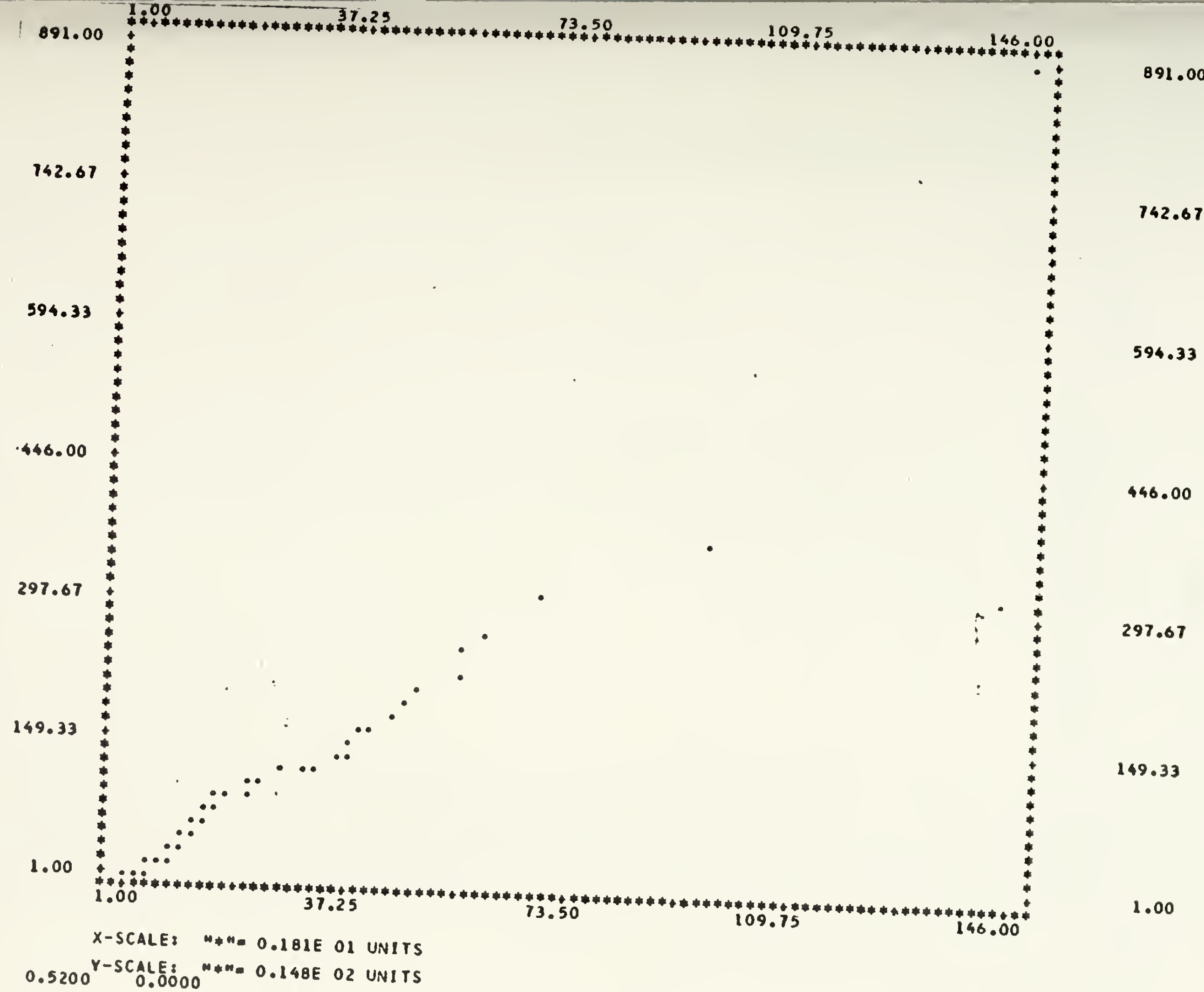
X-SCALE: "*"= 0.181E 01 UNITS

Y-SCALE: "*"= 0.490E 01 UNITS

0.2000 0.1786

Q=0.97

(iv)


1.00
37.25
73.50
109.75
146.00
891.00
742.67
594.33
446.00
297.67
149.33
1.00
1.00
37.25
73.50
109.75
146.00
891.00
742.67
594.33
446.00
297.67
149.33
1.00
X-SCALE: "*"= 0.181E 01 UNITS
Y-SCALE: "*"= 0.148E 02 UNITS
0.5200
0.0000

Q=0.99

(v)

Appendix --F

Cumulative distribution function and probability distribution function values

Number	Pdf	Cdf
1	0.0400	0.0400
2	0.0384	0.0784
3	0.0369	0.1153
4	0.0354	0.1507
5	0.0340	0.1846
6	0.0326	0.2172
7	0.0313	0.2486
8	0.0301	0.2786
9	0.0289	0.3075
10	0.0277	0.3352
11	0.0266	0.3618
12	0.0255	0.3873
13	0.0245	0.4118
14	0.0235	0.4353
15	0.0226	0.4579
16	0.0217	0.4796
17	0.0208	0.5004
18	0.0200	0.5204
19	0.0192	0.5396
20	0.0184	0.5580
21	0.0177	0.5757
22	0.0170	0.5926
23	0.0163	0.6089
24	0.0156	0.6246
25	0.0150	0.6396
26	0.0144	0.6540
27	0.0138	0.6679
28	0.0133	0.6811
29	0.0128	0.6939
30	0.0122	0.7061
31	0.0118	0.7179
32	0.0113	0.7292
33	0.0108	0.7400
34	0.0104	0.7504
35	0.0100	0.7604
36	0.0096	0.7700
37	0.0092	0.7792
38	0.0088	0.7880
39	0.0085	0.7965
40	0.0081	0.8046
41	0.0078	0.8124
42	0.0075	0.8199
43	0.0072	0.8272
44	0.0069	0.8341
45	0.0066	0.8407
46	0.0064	0.8471
47	0.0061	0.8532
48	0.0059	0.8591
49	0.0056	0.8647
50	0.0054	0.8701
51	0.0052	0.8753
52	0.0050	0.8803
53	0.0048	0.8851
54	0.0046	0.8897

55	0.0044	0.8941
56	0.0042	0.8983
57	0.0041	0.9024
58	0.0039	0.9063
59	0.0037	0.9100
60	0.0036	0.9136
61	0.0035	0.9171
62	0.0033	0.9204
63	0.0032	0.9236
64	0.0031	0.9267
65	0.0029	0.9296
66	0.0028	0.9324
67	0.0027	0.9351
68	0.0026	0.9377
69	0.0025	0.9402
70	0.0024	0.9426
71	0.0023	0.9449
72	0.0022	0.9471
73	0.0021	0.9492
74	0.0020	0.9512
75	0.0020	0.9532
76	0.0019	0.9551
77	0.0018	0.9569
78	0.0017	0.9586
79	0.0017	0.9602
80	0.0016	0.9618
81	0.0015	0.9634
82	0.0015	0.9648
83	0.0014	0.9662
84	0.0014	0.9676
85	0.0013	0.9689
86	0.0012	0.9701
87	0.0012	0.9713
88	0.0011	0.9725
89	0.0011	0.9736
90	0.0011	0.9746
91	0.0010	0.9756
92	0.0010	0.9766
93	0.0009	0.9775
94	0.0009	0.9784
95	0.0009	0.9793
96	0.0008	0.9801
97	0.0008	0.9809
98	0.0008	0.9817
99	0.0007	0.9824
100	0.0007	0.9831

101	0.0007	0.9838
102	0.0006	0.9844
103	0.0006	0.9851
104	0.0006	0.9857
105	0.0006	0.9862
106	0.0006	0.9868
107	0.0005	0.9873
108	0.0005	0.9878
109	0.0005	0.9883
110	0.0005	0.9888
111	0.0004	0.9892
112	0.0004	0.9897
113	0.0004	0.9901
114	0.0004	0.9905
115	0.0004	0.9908
116	0.0004	0.9912
117	0.0004	0.9916
118	0.0003	0.9919
119	0.0003	0.9922
120	0.0003	0.9925
121	0.0003	0.9928
122	0.0003	0.9931
123	0.0003	0.9934
124	0.0003	0.9937
125	0.0003	0.9939
126	0.0002	0.9942
127	0.0002	0.9944
128	0.0002	0.9946
129	0.0002	0.9948
130	0.0002	0.9950
131	0.0002	0.9952
132	0.0002	0.9954
133	0.0002	0.9956
134	0.0002	0.9958
135	0.0002	0.9960
136	0.0002	0.9961
137	0.0002	0.9963
138	0.0001	0.9964
139	0.0001	0.9966
140	0.0001	0.9967
141	0.0001	0.9968
142	0.0001	0.9970

143	0.0001	0.9971
144	0.0001	0.9972
145	0.0001	0.9973
146	0.0001	0.9974
147	0.0001	0.9975
148	0.0001	0.9976
149	0.0001	0.9977
150	0.0001	0.9978
151	0.0001	0.9979
152	0.0001	0.9980
153	0.0001	0.9981
154	0.0001	0.9981
155	0.0001	0.9982
156	0.0001	0.9983
157	0.0001	0.9983
158	0.0001	0.9984
159	0.0001	0.9985
160	0.0001	0.9985
161	0.0001	0.9986
162	0.0001	0.9987
163	0.0001	0.9987
164	0.0001	0.9988
165	0.0000	0.9988
166	0.0000	0.9989
167	0.0000	0.9989
168	0.0000	0.9989
169	0.0000	0.9990
170	0.0000	0.9990
171	0.0000	0.9991
172	0.0000	0.9991
173	0.0000	0.9991
174	0.0000	0.9992
175	0.0000	0.9992
176	0.0000	0.9992
177	0.0000	0.9993
178	0.0000	0.9993
179	0.0000	0.9993
180	0.0000	0.9993
181	0.0000	0.9994
182	0.0000	0.9994
183	0.0000	0.9994
184	0.0000	0.9994
185	0.0000	0.9995
186	0.0000	0.9995
187	0.0000	0.9995
188	0.0000	0.9995
189	0.0000	0.9995
190	0.0000	0.9996
191	0.0000	0.9996
192	0.0000	0.9996
193	0.0000	0.9996
194	0.0000	0.9996
195	0.0000	0.9996
196	0.0000	0.9997
197	0.0000	0.9997
198	0.0000	0.9997
199	0.0000	0.9997
200	0.0000	0.9997

LIST OF REFERENCES

1. Atkinson, R. C.; Bower, G. H.; and Crothers, E. J. An Introduction to Mathematical Learning Theory New York: Wiley, 1965

2. Coombs, C. H.; Dawes, R. M.; and Tversky, A. Mathematical Psychology An Elementary Introduction Englewood Cliffs, N. J.: Prentice-Hall, 1970

3. Restle, F. and Green, J. G. Introduction to Mathematical Psychology Reading, Mass: Addison-Wesley, 1970

4. Gray, B. B. and Ryan, B. P. Language Training Through Programmed Conditioning Champaign, Ill.: Research Press (In Press), 1973

5. Madson, G. G. Establishing Criteria in Programmed Learning M. S. Thesis, Naval Postgraduate School, Monterey, Ca.: 1972

6. Oertel, W. W. An Analysis of Response Distribution, M. S. Thesis, Naval Postgraduate School, Monterey, Ca.: 1975

INITIAL DISTRIBUTION LIST

1.	DEFENCE DOCUMENTATION CENTRE CAMERON STATION ALEXANDRIA, VIRGINIA 22314	2
2.	LIBRARY CODE 0212 NAVAL POSTGRADUATE SCHOOL MONTEREY, CA. 93940	2
3.	ASSOC. PROFESSOR JAMES K. ARIMA CODE 55AA, DEPARTMENT OF OPERATION RESEARCH NAVAL POSTGRADUATE SCHOOL MONTEREY, CA. 93940	1
4.	FLIGHT LIEUTENANT M. ZIA-UL-HAQUE P.A.F. BASE KORANGI CREEK, KARACHI, PAKISTAN	1
5.	DEPARTMENT OF OPERATION RESEARCH AND ADMINISTRATIVE SCIENCES NAVAL POSTGRADUATE SCHOOL MONTEREY, CA. 93940	1
6.	BEHAVIORAL SCIENCES INSTITUTE 72, FERN CANYON WAY CARMEL, CA. 93921	2
7.	P.A.F. COLLEGE OF AERONAUTICAL ENGINEERING P.A.F. BASE KORANGI CREEK, KARACHI, PAKISTAN	1
8.	DIRECTORATE OF OPERATIONS RESEARCH AIR HEADQUARTERS PESHAWAR, PAKISTAN	1