A METHODOLOGY OF MATHEMATICAL MODELS WITH AN APPLICATION

A Thesis

by

RICHARD BRIAN WOOD

Submitted to the Graduate College of
Texas A&M University in
partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE

December 1972

Major Subject: Mathematics

A METHODOLOGY OF MATHEMATICAL MODELS

WITH AN APPLICATION

A Thesis

by

RICHARD BRIAN WOOD

Approved as to style and content by:

(Chairman of Committee)

(Head of Department)

(Member)

(Member)

December 1972

ABSTRACT

A Methodology of Mathematical Models
with an Application. (December 1972)
Richard Brian Wood, B.S., University
of Texas at Arlington
Directed by: Dr. Darald J. Hartfiel

The objective of this thesis is to develop a methodology for the construction of mathematical models. The principle results are a statement of a methodology of mathematical models and an example of a Markov chain model constructed using that method.

The methodology is based on the traditional axiomatic method. It is stated in very general terms for universality of application.

The Markov chain model is discussed and developed as well as the necessary results from Markov chain theory. The model is developed strictly from the stated methodology.

ACKNOWLEDGMENTS

I wish to express my gratitude and thanks to my committee chairman, Dr. D. J. Hartfiel, for his guidance and inspiration in the preparation of this thesis. I also wish to thank Dr. J. L. Tollefson and Dr. Guy Curry for serving on my committee.

I especially want to thank my wife, Debbie, for her encouragement and her patience.

DEDICATION

I wish to dedicate this thesis to my parents who made everything possible.

TABLE OF CONTENTS

CHAPTER I

A METHODOLOGY OF MATHEMATICAL MODELS

We begin our discussion of mathematical models with some definitions.

An *axiom* is an assumption in a logical discourse. It is expressed in undefined technical terms called *primitives*. A collection of several axioms is called a *system of axioms*.

A *theorem* is any statement which may be logically deduced from a system of axioms.

A system of axioms is said to be *consistent* if there is no statement S so that S and not S may both be deduced from the system.

An *interpretation* of a system of axioms is an assignment of meanings to the primitives in the system in such a way that the axioms become simultaneously "true." If an interpretation exists for a system of axioms, then the system is said to be *satisfiable*. A *mathematical model* is a set of mathematical expressions resulting from an interpretation of a system of axioms. If these expressions involve probabilities, then the model is said to be *stochastic*. Otherwise, the model is said to be *deterministic*.

We now wish to examine why mathematical models are of value and consider the extent of their usefulness.

The format of this paper is patterned after the Proceedings of the American Mathematical Society.

A mathematical model is a means by which some concrete phenomenon may be studied through abstract expressions using the tools of mathematics and logic. At best, the model is a partial description of certain aspects of reality. Its value depends upon how closely it approximates the characteristics of the phenomenon under study.

In any mathematical model certain assumptions are made. These assumptions are based on a study of the phenomenon to be modeled. This should result in a consistent system of axioms since we tacitly assume nature to be consistent. Any other statements made in the course of developing the model should be logically deduced from the system of axioms, even if they are intuitively clear.

Mathematical models allow concise analysis of complicated physical systems. Very often a study of the mathematics involved will lead to previously unknown information about the physical system. Thus mathematical models may be predictive as well as descriptive.

From the preceding discussion it is clear that the method by which mathematical models are constructed should be explicitly stated. We now wish to develop that method.

The approach to the construction of a mathematical model is essentially the same as that of solving any scientific problem. It is necessary to produce a clear statement of the problem. More precisely, a statement of what we wish to model and what information we wish to obtain from our model should be stated at the outset. This description should include a discussion of terminology, definitions, and any notation to be introduced.

After a thorough description of the problem, the basic assumptions upon which the model will be based should be clearly stated. These assumptions form the system of axioms from which all other conclusions must be logically deduced.

Based upon these axioms, mathematical expressions are derived. These expressions form the basis of the mathematical model. Further, these expressions, under suitable interpretation, describe the phenomenon under study and may be used to predict novel occurrences.

When the mathematical development of the model has been completed, the results should be translated into the nonmathematical terms of the situation we are trying to model. All conclusions drawn from the mathematical model should be stated in terms of what they infer in the "real" world.

With the finished mathematical model we now need to evaluate to what degree it is valid. Many statistical tools are available to determine this. The test known as chi square is probably the best known. It gives a measure of the difference between expected and observed values. Hence it is a natural choice to use to test the validity of mathematical models.

The following chapters develop an example of how mathematical models are constructed using the method we have just set forth.

CHAPTER II

FINITE MARKOV CHAINS

This chapter develops those tools of Markov chain theory which are necessary to the development of our model. For this we assume the results of matrix theory, probability theory, and statistics. The language and notation of matrix theory is consistent with [4]. The language and notation of probability and statistics is consistent with [6].

Consider a sequence of trials whose outcomes $x_1, x_2, \ldots$ satisfy the following conditions:

1) Each outcome belongs to a finite set of outcomes $\{S_1, S_2, \ldots, S_m\}$; if the outcome on the nth trial is S_i, then we say the system is in state S_i at the nth step.
2) The outcome of any trial depends at most upon the outcome of the immediately preceding trial; with each pair of states S_i and S_j there is a probability p_{ij} that S_j occurs immediately after S_i occurs.

Such a process is called a *finite Markov chain*. The numbers p_{ij}, called the *transitional probabilities* can be arranged in a matrix called the *transitional probability matrix* as follows:

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ p_{m1} & p_{m2} & & p_{mm} \end{pmatrix} .$$

The matrix P is nonnegative and row stochastic, i.e., all row sums equal one.

THEOREM 2.1. The probability of moving from state S_i to state S_j in precisely n steps, denoted by $p_{ij}^{(n)}$, is the i,jth entry of P^n.

PROOF. Clearly $p_{ij}^{(1)} = p_{ij}$, the i,jth entry of P. The probability of moving from state S_i to state S_j in two steps is the sum of the probabilities of moving from state S_i to each of the possible states on the first step and then moving to state S_j on the second step. Therefore

$$p_{ij}^{(2)} = \sum_{k=1}^{m} p_{ik}\, p_{kj}.$$

Matrix multiplication shows $p_{ij}^{(2)}$ is the i,jth entry of P^2. In general, the probability of moving from state S_i to state S_j in precisely n steps is the sum of the probabilities of moving from state S_i to each of the possible states in n-1 steps and then moving to state S_j on the nth step. Therefore

$$p_{ij}^{(n)} = \sum_{(k)} p_{ik_1}\, p_{k_1k_2}\, p_{k_2k_3} \cdots p_{k_{n-1}j}$$

where the summation is over all (k) where $(k) = (k_1, k_2, \ldots, k_{n-1})$ is any n-1 selection of $\{1,2,3,\ldots,n\}$. Again matrix multiplication shows $p_{ij}^{(n)}$ is the i,jth entry of P^n.

THEOREM 2.2. If π is a row vector giving the probabilities of being in each state at the present step, πP^n gives these probabilities after n additional steps.

PROOF. Let $\pi = (x_1, x_2, \ldots, x_m)$, where x_i is the probability of being in state S_i at the present step. From the previous theorem, $p_{ki}^{(n)}$ is the probability of moving from state S_k to state S_i in precisely n steps. Then $x_k p_{ki}^{(n)}$ is the probability of the chain being in state S_i after n additional steps from state S_k given that the probability of initially being in state S_k is x_k. Therefore the probability of the chain being in state S_i after n additional steps is

$$\sum_k x_k p_{ki}^{(n)} .$$

This is the ith entry of πP^n.

A *path* is said to exist from state S_i to state S_j if there is a sequence of positive transitional probabilities $p_{i,k_1}, p_{k_1,k_2}, \ldots, p_{k_{r-1}, k_r}, p_{k_r,j}$, i.e., it is possible to reach state S_j from state S_i.

A state S_i is said to be an *absorbing state* if and only if $p_{ii} = 1$, i.e., once the state is entered it cannot be left. A chain is said to be an *absorbing chain* if it has at least one absorbing state and if a path exists from every state to an absorbing state.

Given any transitional probability matrix P for an absorbing Markov chain there exist permutation matrices $\tilde{P}$ and $\tilde{P}^t$ such that $\tilde{P}\, P\, \tilde{P}^t = \overline{P}$ where $\overline{P}$ is of the form

$$\overline{P} = \begin{pmatrix} I_r & 0 \\ R & Q \end{pmatrix} , \qquad (2.1)$$

where I_r is an r x r identity matrix given r absorbing states; Q is s x s given s nonabsorbing states; $R \neq 0$ is s x r, and 0 is an r x s matrix of zero entries.

THEOREM 2.3. Let B be a square matrix. Then

$$\sum_{k=0}^{\infty} B^k \text{ converges if and only if } \lim_{k \to \infty} B^k = 0. \text{ Moreover if}$$

$$\sum_{k=0}^{\infty} B^k \text{ converges, it converges to } (I - B)^{-1}.$$

PROOF. Let $\{B_1, B_2, \ldots\}$ be an infinite sequence of square matrices. The series $\sum_{k=1}^{\infty} B_k$ converges if $\sum_{k=1}^{\infty} b_{ij}^{(k)}$ converges where $b_{ij}^{(k)}$ is the i,jth element of B_k. From calculus we know that if $\sum_{k=1}^{\infty} b_{ij}^{(k)}$ converges then $\lim_{k \to \infty} (b_{ij}^{(k)}) = 0$. Thus if $\sum_{k=1}^{\infty} B_k$ converges, $\lim_{k \to \infty} B = 0$. Therefore as a consequence we have that if $\sum_{k=0}^{\infty} B^k$ converges, $\lim_{k \to \infty} B^k = 0$. Conversely if $\lim_{k \to \infty} B^k = 0$, the eigenvalues of B are less than one in magnitude [3, p. 112] and hence the eigenvalues of I - B are all nonzero. Thus I - B is nonsingular, i.e., $(I - B)^{-1}$ exists. But $I + B + B^2 + B^3 + \ldots + B^k = (I - B)^{-1} (I - B^{k+1})$. Hence as B^k approaches 0, then $\sum_{k=0}^{\infty} B^k = (I - B)^{-1}$ [9].

From (2.1) we see that

$$\overline{P}^{\,n} = \begin{pmatrix} I_r & 0 \\ R' & Q^n \end{pmatrix} . \qquad (2.2)$$

A basic result of finite absorbing Markov chain theory is:

THEOREM 2.4. Suppose P is any transitional probability matrix for an absorbing Markov chain. Then $\lim_{n \to \infty} Q^n = 0$.

PROOF.

Case 1. Q irreducible.

Let S_i be the row sum of the ith row of Q. Let $S = \max_i S_i$ and $s = \min_i S_i$. Let $\sigma = \sum_{i=1}^{n} \frac{S_i}{n}$, $\alpha = \min_i q_{ii}$ and k be the least of the positive off-diagonal elements of Q. Let r be the Perron root of Q. Brauer found that

$$s + \varepsilon(\sigma - s) \leq r \leq S - \varepsilon(S - \sigma), \qquad (2.3)$$

where $\varepsilon = [\frac{k}{S-\alpha}]^{n-1}$ [10, p. 157]. Comparable results were also found by Hartfiel [5]. Now $0 < S \leq 1$ by hypothesis. Also we know that $0 < \sigma < 1$ since $0 < S_i < 1$ for at least one i by hypothesis. Further $\varepsilon > 0$ since k is positive. Hence $0 < S - \sigma < 1$ and $S - \varepsilon\ (S - \sigma) < S$ which implies $r < 1$. Therefore as r is the maximal eigenvalue of Q [10], all of the eigenvalues of Q are less than one, which implies

$$\lim_{n \to \infty} Q^n = 0 \qquad [3, p. 112].$$

Case 2. Q reducible.

If Q is reducible then there exists a permutation matrix P_1 so that

$$P_1 Q P_1^t = \begin{pmatrix} Q_1 & 0 \\ R_1 & Q_2 \end{pmatrix}, \qquad (2.4)$$

where Q_1 is $n_1 \times n_1$ and Q_2 is $(s - n_1) \times (s - n_1)$. If both Q_1 and Q_2 are irreducible, then by the argument of Case 1, all of the eigenvalues of Q are less than one and the conclusion follows.

Without loss of generality, suppose Q_1 is reducible. Then there exists a permutation matrix P_2 so that

$$P_2 Q_1 P_2^t = \begin{pmatrix} Q_{11} & 0 \\ R_{11} & Q_{12} \end{pmatrix} \tag{2.5}$$

where Q_{11} and Q_{12} are both square matrices. Hence

$$\begin{pmatrix} P_2 & 0 \\ 0 & I_{s-n_1} \end{pmatrix} P_1 Q P_1 \begin{pmatrix} P_2 & 0 \\ 0 & I_{s-n_1} \end{pmatrix} = \left(\begin{array}{cc|c} Q_{11} & 0 & 0 \\ R_{11} & Q_{12} & \\ \hline \multicolumn{2}{c|}{R_1} & Q_2 \end{array}\right) \tag{2.6}$$

By continuing this argument we will eventually obtain a permutation matrix $\hat{P}$ so that

$$\hat{P} Q \hat{P}^t = \begin{pmatrix} Q_1' & 0 & 0 & 0 & \cdots & & 0 \\ R_1' & Q_2' & 0 & 0 & \cdots & & 0 \\ R_2' & R_3' & Q_3' & 0 & \cdots & & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \\ R_j' & \cdot & \cdot & \cdot & \cdots & R_k' & Q_n' \end{pmatrix}$$

where each Q_i' $(i = 1, \cdots, n)$ is square and irreducible. (Each 1 x 1 matrix is irreducible by definition.) Again by arguing as in Case 1 the eigenvalues of Q are all less than one which implies

$$\lim_{n \to \infty} Q^n = 0.$$

As a consequence of Theorems 2.3 and 2.4 we have that the matrix series $N = I + Q + Q^2 + Q^3 \cdots$ always converges and $N = (I - Q)^{-1}$. The matrix N is called the *fundamental matrix* of the chain.

DEFINITION 2.1. Let f be any numerical function defined on a possibility space S. Then the mean of the function f is

$$M\,[f] = \sum_j j \cdot P_r\,[f = j] \tag{2.7}$$

Example 1. A coin is tossed twice. Then the possibility space S = {HH, HT, TH, TT} where H denotes heads, T denotes tails. Let f be the number of heads that turn up. Then

$$M[f] = 2 \cdot 1/4 + 1 \cdot 1/4 + 1 \cdot 1/4 + 0 \cdot 1/4 = 1.$$

Example 2. Let A = {1,2,7,9,9}. A number is drawn at random from this set. The possibility space is simply the set A itself. Let f be the outcome function. Then

$$M[f] = 1 \cdot 1/5 + 2 \cdot 1/5 + 7 \cdot 1/5 + 9 \cdot 1/5 + 9 \cdot 1/5 = 5\ 3/5.$$

THEOREM 2.5. Consider the following possibility space $S = \left\{x_0, x_1, x_2, \ldots, x_m, \cdots \text{ where } x_0 = S_i, \text{ and } x_i\ \varepsilon \{S_1, \ldots, S_n\} \text{ for } i = 0,1,2,\ldots\right\}$.

Let

$$X_{ij}^{(k)} = \begin{cases} 1 \text{ for all members of S where } x_k = S_j \\ 0 \text{ otherwise} \end{cases} \tag{2.8}$$

Set $f = \sum_{k=0}^{\infty} X_{ij}^{(k)}$. Then $n_{ij} = M[f]$, i.e., M[f] is the mean number of times the chain passes through a fixed nonabsorbing state

S_j starting at a fixed nonabsorbing state S_i.

PROOF. From Definition 2.1 we have

$$M[X_{ij}^{(k)}] = q_{ij}^{(k)} , \tag{2.9}$$

where $q_{ij}^{(k)}$ is the i,jth entry of Q^k. It can be shown that

$$M[\sum_{n=0}^{\infty} f_n] = \sum_{n=0}^{\infty} M[f_n]. \tag{2.10}$$

(For proof see [7].)

Using (2.10), we define $\hat{n}_{ij}$ as follows:

$$\hat{n}_{ij} = M[\sum_{k=0}^{\infty} X_{ij}^{(k)}] = \sum_{k=0}^{\infty} M[X_{ij}^{(k)}]. \tag{2.11}$$

Then from (2.9) and (2.11) we have that

$$\hat{n}_{ij} = q_{ij}^{(o)} + q_{ij}^{(1)} + q_{ij}^{(2)} + \cdots$$

Thus $\hat{n}_{ij} = n_{ij}$, the i,jth entry of the matrix $N = I + Q = Q^2 + Q^3 + \cdot\cdot$ which is the fundamental matrix.

THEOREM 2.6. Let b_{ij} be the probability that an absorbing chain will be absorbed in state S_j if it starts in nonabsorbing state S_i. Let B be the matrix with entries b_{ij}. Let R be the s x r matrix as defined in (2.1). Then B = NR where N is the fundamental matrix.

PROOF. From (2.1) we have

$$\overline{P} = \begin{pmatrix} I & 0 \\ R & Q \end{pmatrix}$$

Then

$$\overline{P}^2 = \begin{pmatrix} I & 0 \\ R+QR & Q^2 \end{pmatrix}, \text{ and}$$

$$\bar{P}^3 = \begin{pmatrix} I & 0 \\ R + QR + Q^2R & Q^3 \end{pmatrix} .$$

Continuing this multiplication we have

$$\bar{P}^n = \begin{pmatrix} I & 0 \\ R + QR + Q^2R + \cdots + Q^{n-1}R & Q^n \end{pmatrix} .$$

From Theorems 2.3 and 2.4 we have the following:

$$\bar{P}^\infty = \begin{pmatrix} I & 0 \\ (I + Q + Q^2 + \cdots)R & 0 \end{pmatrix}$$

$$= \begin{pmatrix} I & 0 \\ NR & 0 \end{pmatrix} .$$

Hence by definition B = NR, and B is row stochastic.

CHAPTER III

A MARKOV CHAIN MODEL IN SOCIOLOGY

3.1 *Introduction.*

Many sociological and psychological experiments deal with the behavior of small groups of individuals. This chapter will attempt to develop a mathematical model for one such type of experiment.

3.2 *Statement of the problem.*

The experiment to be modeled was first performed by Asch [1] with subsequent development by Cohen [2] and Kemeny and Snell [8].

A subject is seated with a group of confederates. They are shown two cards, card A and card B. Card A has three lines of different lengths while card B has only one line. Each person is asked to choose which one of the lines on card A they feel is the same length as the line on card B.

The group of confederates consistently and unanimously give wrong answers. The uninformed subject is then confronted with the following dilema: he may conform with the majority and answer incorrectly, or he may reject the pressure of the majority and answer correctly.

The data for this experiment is a sequence of responses. The letters a and b are used to denote correct and incorrect responses respectively.

The same number of trials are performed for each subject. A sufficiently large number of trials are performed so that each subject will eventually give a sequence of consistent responses. This sequence of consistent responses is called the *terminal segment*. The consistent response given in the terminal segment is called the *terminal response*. The others form the *initial segment*.

3.3 *The model.*

We begin a discussion of the model by stating several axioms.

AXIOM 3.1. On any given trial the subject must be in one of four states:

State S_1 - If the subject is in this state, he will answer correctly on this trial and all those which follow.

State S_2 - If the subject is in this state, he will answer correctly on this trial, but not necessarily on those which follow.

State S_3 - If the subject is in this state, he will answer incorrectly on this trial, but not necessarily on those which follow.

State S_4 - If the subject is in this state, he will answer incorrectly on this trial and all those which follow.

AXIOM 3.2. The probability of being in a particular state on trial n depends only on the state the subject was in on the preceding trial.

AXIOM 3.3. Each subject is initially in state S_2, and must at some time be in either state S_1 or state S_4.

Based on Axiom 3.3 we may consider a as the "0th" response. We then define n_{aa} as the total number of a to a transitions in all modified initial segments. Define n_{bb} similarly. We define n_{ab} and n_{ba} as the total number of a to b and b to a transitions in all segments. Define n_a to be the total number of a responses in all modified initial segments. Define n_b similarly.

The following relations follow from these definitions:

$$n_a = n_{aa} + n_{ab} \tag{3.1}$$

$$n_b = n_{bb} + n_{ba} \quad . \tag{3.2}$$

Define t_a and t_b to be the total number of subjects whose terminal responses are a and b respectively. Then it follows that

$$t_b = n_{ab} - n_{ba} \quad . \tag{3.3}$$

AXIOM 3.4. The following state-to-state transitions are not allowed:

State S_1 to States S_2, S_3, or S_4.

State S_4 to States S_1, S_2, or S_3.

State S_2 to State S_4.

State S_3 to State S_1.

Let S (k,n) denote the state person k is in on the nth trial. Then p_{ij} is the probability that $S(k,n+1) = S_j$ given $S(k,n) = S_i$.

It is possible to arrange the transitional probabilities into a transitional probability matrix as follows:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ p_{21} & p_{22} & p_{23} & 0 \\ 0 & p_{32} & p_{33} & p_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} . \qquad (3.4)$$

From Axiom 3.1, $p_{11} = p_{44} = 1$. From Axiom 3.4, $p_{12} = p_{13} = p_{14} = p_{24} = p_{31} = p_{41} = p_{42} = p_{43} = 0$. From Axiom 3.3, either p_{21} or p_{34} is strictly positive.

3.4 *Objectives of the model.*

1) To formulate a method of determining the unknown entries of the matrix P.
2) To predict the mean number of correct responses.
3) To predict the mean number of times a subject will change from one response to another.
4) To predict the proportion of subjects who will not be intimidated.
5) To find the mean number of subjects who will change responses exactly k times.

3.5 *Mathematical treatment*

Markov chain theory (see Chapter II) permits rewriting the transitional probability matrix P listing the absorbing states first:

$$\tilde{P} = \begin{pmatrix} I & 0 \\ R & Q \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ p_{21} & 0 & p_{22} & p_{23} \\ 0 & p_{34} & p_{32} & p_{33} \end{pmatrix}. \quad (3.5)$$

The fundamental matrix is then

$$N = (I - Q)^{-1} = \begin{pmatrix} 1 - p_{22} & - p_{23} \\ - p_{32} & 1 - p_{33} \end{pmatrix}^{-1}. \quad (3.6)$$

Computing this inverse by the adjoint method yields

$$N = \frac{\text{adj } (I - Q)}{\det (I - Q)} = \frac{1}{\delta} \begin{pmatrix} 1 - p_{33} & p_{23} \\ p_{32} & 1 - p_{22} \end{pmatrix} =$$

$$\begin{pmatrix} n_{22} & n_{23} \\ n_{32} & n_{33} \end{pmatrix}, \quad (3.7)$$

where $\delta = (1 - p_{22})\ (1 - p_{33}) - p_{23}p_{32}$. (3.8)

The matrix B of absorption probabilities is

$$B = NR = \frac{1}{\delta} \begin{pmatrix} (1 - p_{33})\ p_{21} & p_{23}p_{34} \\ p_{32}p_{21} & (1-p_{22})\ p_{34} \end{pmatrix} \quad (3.9)$$

$$= \begin{pmatrix} b_{21} & b_{24} \\ b_{31} & b_{34} \end{pmatrix}, \quad (3.10)$$

and B is row stochastic.

Since the n_{23} entry of the matrix N gives the mean of the total number of times a subject is in state S_3 during a sequence of trials, then

$$M[\frac{n_b}{n}] = n_{23} - f_{23} \ , \tag{3.11}$$

where f_{23} is the mean number of times a subject is in state S_3 in the terminal segment, n is the total number of subjects, and $\frac{n_b}{n}$ is the average number of b responses per subject in the initial segments.

Let q_m be the probability of exactly m occurrences of state S_3 in the terminal segment. If the terminal state is S_1 then there can be no occurrences of state S_3. Thus,

$$q_o = b_{21} \ . \tag{3.12}$$

Suppose we know q_m. We may then find q_{m+1} very easily. The probability of m+1 occurrences of state S_3 is, by elementary laws of probability, the probability of m occurrences of state S_3 multiplied by the probability of one more occurrence of state S_3. Therefore

$$\begin{aligned} q_{m+1} &= q_m p_{33} \\ &= q_1 p_{33}^m \ . \end{aligned} \tag{3.13}$$

Now

$$\begin{aligned} 1 = \sum_{m=0}^{\infty} q_m &= q_o + \sum_{m=1}^{\infty} q_m \\ &= b_{21} + \sum_{m=1}^{\infty} q_m \\ &= b_{21} + \frac{q_1}{1 - p_{33}} \ , \end{aligned}$$

since $\sum_{m=1}^{\infty} q_m$ is a geometric series by (3.13). From (3.14) we have

$$q_1 = (1 - p_{33})\ (1 - b_{21})$$
$$= (1 - p_{33})\ b_{24}\ .$$

Hence,

$$q_m = p_{33}^{m-1}\ (1 - p_{33})\ b_{24}.$$

Now

$$f_{23} = \sum_{m=1}^{\infty} mq_m$$
$$= \sum_{m=1}^{\infty} mp_{33}^{m-1}\ (1 - p_{33})\ b_{24}$$
$$= b_{24} \sum_{m-1}^{\infty} mp_{33}^{m-1}\ \ (1 - p_{33})$$
$$= b_{24} \sum_{m=0}^{\infty} p_{33}^{m}$$
$$= \frac{b_{24}}{1 - p_{33}}$$
$$= \frac{p_{23}\ \ p_{34}}{\delta(1 - p_{33})}\ .$$

Then from (3.11) we have

$$M\ [\frac{n_b}{n}] = \frac{p_{23}}{\delta} - \frac{p_{23}\ \ p_{34}}{\delta(1 - p_{33})}$$
$$= \frac{p_{23}\ (1 - p_{33} - p_{34})}{\delta(1 - p_{33})}$$
$$= \frac{p_{23}\ \ p_{32}}{\delta(1 - p_{33})}\ . \qquad (3.15)$$

Similarly we may write $M[\frac{n_a}{n}] = n_{22} - f_{22}$. Using a computational procedure analagous to the one above we find

$$M[\frac{n_a}{n}] = \frac{1 - p_{33}}{\delta} - \frac{p_{21}\ (1 - p_{33})}{\delta(1 - p_{22})}$$

$$= \frac{p_{23}\ (1 - p_{33})}{\delta(1 - p_{22})} \ . \qquad (3.16)$$

From (3.9) and (3.10) we find

$$M[\frac{t_b}{n}] = b_{24} = \frac{p_{23}\ p_{34}}{\delta} \qquad (3.17)$$

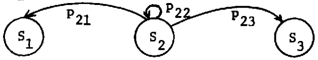
and

$$M[\frac{t_a}{n}] = b_{21} = \frac{(1 - p_{33})\ p_{21}}{\delta} \ . \qquad (3.18)$$

Suppose the chain is only observed when a change of state occurs, rather than a change of response. Then the transitional probability matrix for this chain is

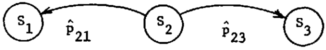
$$\hat{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{p_{21}}{1 - p_{22}} & 0 & \frac{p_{23}}{1 - p_{22}} & 0 \\ 0 & \frac{p_{32}}{1 - p_{33}} & 0 & \frac{p_{34}}{1 - p_{33}} \\ 0 & 0 & 0 & 1 \end{pmatrix} . \qquad (3.19)$$

To see this, consider the original chain. Suppose k subjects are in state S_2. Then the distribution after one step would be as follows:



$p_{21}k$ $\qquad$ $p_{22}k$ $\qquad$ $p_{23}k$

Now if we consider the chain in which only changes of state are considered, then the distribution after one step will be as follows:



$$\hat{p}_{21}(k - p_{22}k) \qquad k - p_{22}k \qquad \hat{p}_{23}\,(k - p_{22}k)$$

Hence,

$$\hat{p}_{21}\,(k - p_{22}k) = p_{21}k \quad ,$$

and

$$\hat{p}_{21} = \frac{p_{21}}{1 - p_{22}} \quad .$$

The other entries of $\hat{P}$ may be found in a similar manner.

The fundamental matrix $\hat{N}$ for this chain is

$$\hat{N} = \frac{1}{\delta}\begin{pmatrix} (1 - p_{22})\,(1 - p_{33}) & p_{23}\,(1 - p_{33}) \\ p_{32}\,(1 - p_{22}) & (1 - p_{22})\,(1 - p_{33}) \end{pmatrix} \quad ,$$

where δ is defined as in (3.8).

The $\hat{n}_{23}$ element of $\hat{N}$ is the mean number of changes from state S_2 to state S_3 before absorption. This is the same as the mean number of changes from response a to response b since changes from state S_2 to state S_4 are not permitted. Thus

$$M_{ab} = M[\frac{n_{ab}}{n}] = \hat{n}_{23} = \frac{p_{23}\,(1 - p_{33})}{\delta} \quad . \qquad (3.20)$$

The mean of $\frac{n_{aa}}{n}$ may be computed from (3.1), (3.16) and (3.20):

$$M_{aa} = M[\frac{n_{aa}}{n}] = M[\frac{n_a - n_{ab}}{n}]$$

$$= M[\frac{n_a}{n}] - M[\frac{n_{ab}}{n}]$$

$$= \frac{p_{23}\ (1 - p_{33})}{\delta\ (1 - p_{22})} - \frac{p_{23}\ (1 - p_{33})}{\delta}$$

$$= \frac{p_{22}\ p_{23}\ (1 - p_{33})}{\delta\ (1 - p_{22})}\ . \qquad (3.21)$$

Similarly from (3.3), (3.17) and (3.20),

$$M_{ba} = M[\frac{n_{ba}}{n}] = \frac{p_{23}p_{32}}{\delta}\ , \qquad (3.22)$$

and from (3.1), (3.11) and (3.22),

$$M_{bb} = M[\frac{n_{bb}}{n}] = \frac{p_{23}p_{32}p_{33}}{\delta\ (1 - p_{33})}\ . \qquad (3.23)$$

It is possible to determine the transitional probabilities from the means (3.20)-(3.23). Thus, in a sense, the transitional probabilities are group mean probabilities rather than individual subject probabilities. If we consider the elementary definition of probability as number of successes divided by the number of possible outcomes and use means, then a natural assumption would be as follows:

$$p_{22} = \frac{M_{aa}}{M_{aa} + M_{ab}}\ . \qquad (3.24)$$

This result may be easily verified from (3.20) and (3.21). Similarly from (3.22) and (3.23) we have that

$$p_{33} = \frac{M_{bb}}{M_{bb} + M_{ba}} \quad . \tag{3.25}$$

Using the row stochastic property of the matrix P, we can verify that

$$\begin{aligned}
p_{21} &= \frac{M_{ab}(1 - M_{ab} + M_{ba})}{(M_{aa} + M_{ab})(1 + M_{ba})} \\
p_{23} &= \frac{M_{ab}^{2}}{(M_{aa} + M_{ab})(1 + M_{ba})} \\
p_{32} &= \frac{M_{ba}^{2}}{M_{ab}(M_{ba} + M_{bb})} \\
p_{34} &= \frac{M_{ba}(M_{ab} - M_{ba})}{M_{ab}(M_{ba} + M_{bb})} \quad .
\end{aligned} \tag{3.26}$$

Further it is seen that these are the only solutions.

3.6 *Interpretation of results.*

The Law of Large Numbers says that if an experiment is repeated a large number of times then computed averages will approach their predicted means. In other words, given an arbitrarily small positive number ε, there is some positive integer N such that for all n greater than N, the difference between $\frac{n_{ab}}{n}$ and M_{ab} will be less than ε. For purposes of interpretation, we assume this difference will be zero, i.e., the averages exactly equal the predicted means.

In one of Cohen's experiments, it was found that

$n_{aa} = 196$, $n_{ab} = 117$, $n_{ba} = 106$, $n_{bb} = 102$, $n_a = 313$,

$n_b = 208$, $t_a = 22$, $t_b = 11$, $n = 33$.

The assumption then yields

$$M_{aa} = \frac{196}{33}\ ,\ M_{ab} = \frac{117}{33}\ ,\ M_{ba} = \frac{106}{33}\ ,\ M_{bb} = \frac{102}{33}\ .$$

From (3.24)-(3.26) we obtain the transitional probability matrix

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ .06 & .63 & .31 & 0 \\ 0 & .46 & .49 & .05 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \qquad (3.27)$$

We now examine this model to see if it meets the objectives stated previously.

1) The model clearly gives a method of finding the unknown entries of the transitional probability matrix.
2) Using (3.20), (3.21) and (3.27), we may predict the mean number of correct responses as $M_{aa} + M_{ab}$.
3) Using (3.20), (3.22) and (3.27), we may predict the mean number of times a subject will change from one response to another as $M_{ab} + M_{ba}$.
4) To predict the proportion of subjects who will not be intimidated, i.e., who will never answer incorrectly, we must find the proportion of subjects who change responses zero times. We note that those who answer incorrectly consistently, those who give b responses on every trial, are not considered to be among those who change responses zero times. This is because the "0th" response of each subject is a. Hence these individuals have changed responses once. Thus

those who change responses zero times are those who give response a consistently. Therefore the probability of changing response zero times is the same as the probability $\hat{p}_{21}$ from the transition matrix (3.19). Thus the model predicts that, for n total subjects, $\hat{p}_{21}n$ of those subjects will not be intimidated.

5) To find the mean number of subjects who will change responses exactly k times it is necessary to expand the idea presented in 4). It should be explained that two cases shall be considered, k even and k odd. This is because if k is even the terminal response is a, if k is odd the terminal response is b. It has already been shown that for $k = 0$, the probability of exactly k changes is $\hat{p}_{21}$. Now if $k = 1$, the terminal response is b, hence the terminal state is S_4. Since everyone is initially in state S_2, the probability of one change in response is given by $\hat{p}_{23}\hat{p}_{34}$. By continuing this argument, we see that the probability of 2k changes is $\hat{p}_{21}[\hat{p}_{23}\hat{p}_{32}]^k$. Similarly the probability of $2k + 1$ changes is $\hat{p}_{23}\hat{p}_{34}[\hat{p}_{23}\hat{p}_{32}]^k$.

We will use the predicted quantities in 4) and 5) above to test the model. The following table gives the observed values and the predicted values for the mean number of subjects who change responses exactly k times.

k	Number of subjects who switched responses exactly k times	Mean number predicted by the model
0	10	5.34
1	1	2.71
2 or 4	5	7.16
3 or 5	4	3.63
6,8, or 10	2	5.47
7, 9, or 11	1	2.77
Large even	5	4.02
Large odd	5	2.01
Total	33	33.11

The standard statistical test used to measure the difference between predicted values and observed value is the chi square test. If we apply this test to this data, we find the deviations to be within the expected margin of error 90 percent of the time. Depending upon the tolerance needed for an experiment of this type, this may or may not be judged a suitable model. Normally 95 per cent reliability is considered acceptable.

3.7 *Concluding remarks.*

If ninety per cent reliability is judged to be inadequate, two courses of action are possible: (1) reject the system of axioms, or (2) retain the system of axioms. Assuming case 1, it could be argued, for example, that it is erroneous to consider an experiment of this type to be Markov. This would require new axioms and result in a completely new model. Another possible source of error is the assumption that the predicted means are exactly equal to the computed average. Increases in the number of subjects or the number of responses per subject could give better results. If all of the axioms and assumptions are accepted, then the model as stated must be accepted.

Computational errors could also be a problem source. One could possibly calculate the transition matrix (3.27) with more significant digits to reduce the effects of roundoff error for more accurate results.

The method presented is one of many possible ways in which problems of this nature may be handled. Further experimentation would probably lead to a more sophisticated approach.

CHAPTER IV

CONCLUSION

The principle results of this thesis were a methodology for the construction of mathematical models, and a Markov chain model constructed using this method.

The model presented in this thesis was deliberately chosen from a nonmathematical, nonscientific field of study to demonstrate the fact that mathematical models are powerful and useful tools regardless of the area of application. It was developed strictly from the stated methodology. The presentation differs from that of Kemeny and Snell [8] in that proofs of important results are included rather than relying on intuitive explanations.

The study of Markov chain models is useful since so many stochastic processes exhibit Markov properties. These models also offer an opportunity to study certain properties of irreducible matrices.

The methodology stated in this thesis is an attempt to give an organized approach to the task of constructing mathematical models. It is felt that the more rigorous the methodology, the more useful the resulting model will be.

BIBLIOGRAPHY

1. S. E. Asch, Social psychology, Prentice Hall, Inc., New York, 1952.

2. B. P. Cohen, "A probability model for conformity," Sociometry, 21 (1958), 69-81.

3. J. N. Franklin, Matrix theory, Prentice Hall, Inc., Englewood Cliffs, N.J., 1968. MR 38 #5798.

4. F. R. Gantmacher, The theory of matrices, Vol. 2, Chelsea, New York, 1959. MR 16 #438.

5. D. J. Hartfiel, "Bounds for eigenvalues and eigenvectors of a nonnegative matrix which involve a measure of irreducibility," SIAM Journal on Applied Mathematics (to appear).

6. J. G. Kemeny, A. W. Knapp, and J. L. Snell, Denumerable markov chains, Van Nostrand, Princeton, N.J., 1966. MR 34 #6858.

7. J. G. Kemeny, H. Mirkil, G. L. Thompson, and J. L. Snell, Finite mathematical structures, Prentice Hall, Inc., Englewood Cliffs, N.J., 1959. MR 21 #4888.

8. J. G. Kemeny and J. L. Snell, Mathematical models in the social sciences, Ginn, Boston, 1962. MR 25 #3797.

9. H. A. Luther, unpublished class notes, 1972.

10. M. Marcus and H. Minc, A survey of matrix theory and matrix inequalities, Allyn and Bacon, Inc., Boston 1964. MR 29 #112.

VITA

Richard Brian Wood was born on the 11th of August, 1949, in Tyler, Texas. His parents are Richard Lee and Dorothy Wood.

He attended the University of Texas at Arlington from September 1967 to May 1971. He graduated with honors in 1971 with a Bachelor of Science degree in Mathematics.

He attended Texas A&M University from September 1971 to December 1972 and received a Master's degree in Mathematics.

He is married to the former Debra Hurry.

His permanent address is

2212 Rosewood Drive

Waco, Texas 76710

This thesis was typed by Mrs. Janie Leighman.