Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

RANDOM GENETIC DRIFT DIFFUSION MODEL

AND

DETERMINISTIC AND STOCHASTIC MODELS OF EPIDEMICS

NORMAN W. LONEY

A Thesis presented to the Faculty of the Graduate School at NJIT, in partial fulfillment of the requirements for the degree of Master of Science in Applied Mathematics.

1985

APPROVAL SHEET

Title of thesis: Random Genetic Drift Diffusion Model and Deterministic and Stochastic Models of Epidemics

Name of candidate: Norman W. Loney

Master of Science in Applied Mathematics

May, 1985

Thesis and abstract approved:

Rose Ann Dios

John Tavantzis

Roman Voronka - Thesis Supervisor
Professor
Department of Mathematics

VITA

Name: Norman W. Loney

Permanent address:

Degree and date to be conferred: Master of Science in Applied Mathematics, 1985

Date of birth:

Place of birth:

Secondary education: Excelsior, October 1970

Collegiate institutions attended	Dates	Degree	Date of Degree
NJIT	1974-77	B.S.	
NJIT	1984-85	M.S.	1985

Major: Mathematics

Positions held: Teaching Assistant in Mathematics Department
New Jersey Institute of Technology
Newark, New Jersey 07102

ABSTRACT

Title of Thesis: RANDOM GENETIC DRIFT DIFFUSION MODEL

AND

DETERMINISTIC AND STOCHASTIC MODELS

OF EPIDEMICS

NORMAN W. LONEY, Master of Science in Applied Math, 1985

Thesis directed by: Dr. Roman Voronka

In the Random Genetic Drift Diffusion model two approaches are taken. First we examined a discrete model that represent a relatively idealised version of the phenomena. We further make the assumption that the population reproduces itself and then dies, thus maintaining a finite population size at all times. If at a given locus there are two possible allels A and B and if X(t) is the number of A type in the genetic pool of size 2N, then 2N-X(t) is the number of B type. We then proceed to obtain a probability density function of X(t) by an Exact method and the Monte Carlo method.

Based on a χ^2 for each generation examined there are no significant difference between the results obtained from either method. However, for large N (N > 20) the Exact method is cumbersome. and as a result the Monte Carlo is more appropriate for such N.

As a second approach, we approximated the Discrete model for large N with a Diffusion model (a singular parabolic partial differential equation) where x and t

are assumed continuous. By separation of variables we obtained the Hypergeometric equation which has an infinite series solution. From this we obtained the probability density as a function of gene frequency and compare these results with those of the previous methods (Discrete model). We found that there is favourable comparison between all three methods and in particular between the Diffusion Approximation and the Monte Carlo.

The Monte Carlo method was also utilized in the Stochastic models of Epidemics. The models we examined are the Chain Binomial models of Reed-Frost and Greenwood. We confirmed that for a household of 3 and smaller, both models are indistinguishable, whereas a household of 5 produced different chains based on the inherent assumptions in each model.

Establishing the existence of a threshold population size, we used a continuous model(Deterministic Theory). This approach resulted in a system of nonlinear ordinary differential equations.The solution of which using the Runge-Kutta (order four) established a relative removal rate above which no epidemic seems to occur, as well as demonstrate the existence of a threshold population size.

TABLE OF CONTENTS

CHAPTER I

THE FISHER-WRIGHT MODEL OF RANDOM GENETIC DRIFT

In this stochastic model from mathematical genetics we consider a diploid population whose size is N individuals. Thus in the genetic pool at a given locus there will be exactly 2N genes. We will assume that the population reproduces itself and then dies, so that the population size is N at all times. If we further assume that at the given locus there are two possible alleles A and B and if X(t) is the number of A type in the genetic pool of size 2N, then 2N - X(t) is the number of B type.A model due to Wright assumes that X(t) is a random variable binomially distributed with parameter X(t)/2N ; thus if the value of X(t) = i,then the probability P_{ij} that X(t+1) = j is given by

$$P_{ij} = \binom{2N}{j} (i/2N)^{j} \; (1 - i/2N)^{2N-j} \qquad (1)$$

This model assumes that there is no mutation from A to B or B to A and that there are no selective pressures favoring one allele over another . (Ewens)[1]

There are two ways in which we shall obtain the probability density function (pdf) of X(t).

I) EXACT METHOD

The first will be an exact method in which time measured in generations is discrete. Since the process is Markovian with $P = \{p_{ij}\}$ the transition matrix, we have

$$X(t+1) = X(t)P$$

with X(t) a row vector giving the probability density function of the random variable X at time t: $X(t) = \{X_0(t), X_1(t), \ldots X_{2N}(t)\}$
where $X_j(t)$ is the probability that the frequency of A is j/2N at time t.

Thus

$X(t) = X(0)\ P^t$ where X(0) is the initial probability vector

For example if 2N = 4, then :

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ .3164 & .4219 & .2109 & .0469 & .0039 \\ .0625 & .25 & .375 & .25 & .0625 \\ .0039 & .0469 & .2109 & .4219 & .3164 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

If at time zero X = 3 then X(0) = (0, 0, 0, 1, 0) and X(1) =X(0)P= (.0039,.0469,.2109,.4219,.3164). Here .0469 is the probability that at time 1 (one generation later) the value of X = 1 and 0.3164 is the probability that at time 1 the value of X = 4 (gene A is fixed).

$$P^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ .4632 & .2329 & .1780 & .0923 & .0336 \\ .1660 & .2109 & .2461 & .2109 & .1660 \\ .0336 & .0923 & .1780 & .2329 & .4632 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P^3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ .5484 & .1471 & .1353 & .0943 & .0748 \\ .2490 & .1604 & .1813 & .1604 & .2490 \\ .0748 & .0943 & .1353 & .1471 & .5484 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

where the powers of P corresponds to the generation in question.

When 2N = 6, then :

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ .3347 & .4019 & .2009 & .0536 & .0080 & .0006 & .00002 \\ .0878 & .2634 & .3292 & .2195 & .0823 & .0165 & .0014 \\ .0156 & .0937 & .2344 & .3125 & .2344 & .0937 & .0156 \\ .0014 & .0165 & .0823 & .2195 & .3292 & .2634 & .0878 \\ .00002 & .0006 & .0080 & .0536 & .2009 & .4019 & .3349 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ .4880 & .2196 & .1601 & .0842 & .0351 & .0110 & .0021 \\ .2084 & .2145 & .2196 & .1739 & .1110 & .0544 & .0180 \\ .0728 & .1326 & .1893 & .2106 & .1893 & .1326 & .0728 \\ .0180 & .0544 & .1110 & .1739 & .2196 & .2145 & .2084 \\ .0021 & .0110 & .0351 & .0842 & .1601 & .2196 & .4880 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P^3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ .5769 & .1389 & .1195 & .0815 & .0484 & .0243 & .0103 \\ .3024 & .1622 & .1657 & .1413 & .1080 & .0712 & .0490 \\ .1374 & .1261 & .1550 & .1631 & .1550 & .1261 & .1374 \\ .0490 & .0712 & .1080 & .1413 & .1657 & .1622 & .3024 \\ .0103 & .0243 & .0484 & .0815 & .1195 & .1389 & .5769 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

We used this exact method to examine cases up to size 2N = 20 . The matrix representation for such population of size 10 is cumbersome (a 21x21 matrix) . Histograms at 4 generations (1,2,10 & 20) for the case 2N = 20 are included in figures 4-7.

II) MONTE CARLO METHOD

When N is larger than 20, the exact method generates matrices of size (2N+1x2N+1) which is clearly cumbersome to manipulate. Thus we use another approach to model (1),a Monte Carlo method[2]. In our scheme,given the population of size 2N which at a given time t is in state i, we calculate the transition probabilities (a row in the matrix P). Here the transition probabilities are given by

$$p_{ij} = P\{ X(t+1)=j | X(t)=i \}$$

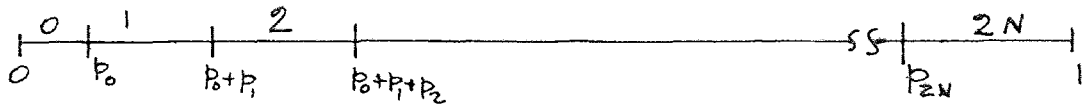
To decide in what state will the population be at time t+1 ,we consider the interval 0 <= y <= 1 and divide it into 2N sub-intervals with lengths p_0, $p_1, p_2, \ldots, p_{2N}$ the coordinates of the division points will be

$$y = p_0,\ y = p_0 + p_1,$$

$$y = p_0 + p_1 + p_2 + \ldots,$$

$$y = p_0 + p_1 + \ldots + p_{2N-1}.$$

We can further identify these subintervals with the numbers 0,1,...,2N as in the sketch below.



At this stage we generate a random number ,

$$0 <= \gamma <= 1.$$

If this number falls into the j-th subinterval of the partitioned line $0 <= y <= 1$ then we conclude that

$$X(t+1) = j.$$

In this method, the random variable γ is uniformly distributed in (0,1), the probability of γ lying within one of the sub-intervals is equal to the length of the sub-interval in question. Therefore:

$$P\{ 0 < \gamma < p_1 \} = p_1 ,$$

$$P\{ p_1 < \gamma < p_1 + p_2 \} = p_2 ,$$

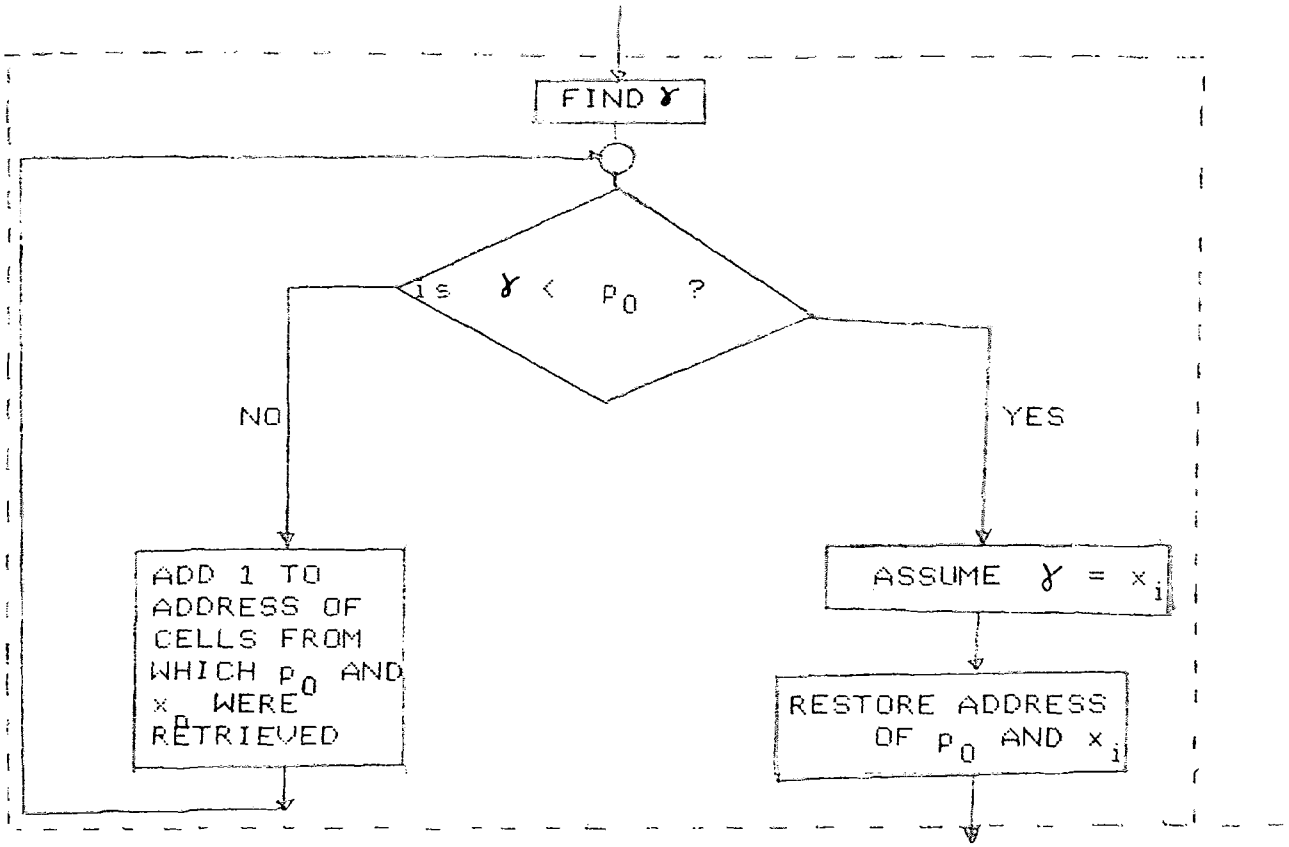
. . . . . ,

$$P\{ p_1 + p_2 + \ldots + p_{n-1} < \gamma < 1 \} = p_n$$

$$X(t+1) = j$$

if $p_1 + \ldots + p_{j-1} < \gamma < p_1 + \ldots + p_j$

Shown below is the subroutine used in our analysis: Suppose that the numbers 0, 1, ...,2N are placed in succession in storage cells and the probabilities p_0, p_0+p_1, $p_0+p_1+p_2$,...,1 also form a sequence in data storage. Then :



For example, to draw 10 values of the random variable with the distribution

$$P(\theta = 0) = .58 \ , \ P(\theta = 1) = .42$$

Select as values of γ ten pairs of numbers from a table of random numbers and multiply by .01. Thus suppose γ = 0.86,0.51,0.59,0.07,0.95,0.66 0.15, 0.56 0.64,0.34 (appendix A-1 table of random numbers). Then based on our scheme the value θ = 0 corresponds to the values of γ smaller than 0.58 and θ = 1, to the values of γ >= 0.58 i.e. θ = 1, 0, 1, 0, 1, 1, 0, 0, 1, 0. Note here that the order of enumerating

the numbers 0, 1,...,2N in the partition of 0 <= y <=1 can be arbitrary,but it must be fixed prior to drawing.

DIFFUSION APPROXIMATION

III)When the population size N is large, model (1) (a discrete model) can be approximated by a model where both x = X(t)/2N and t are continuous. We consider the derivation of the diffusion model along the lines given in Crow and Kimura[3],and Ludwig[4]. Let x diffuse on [0,1]. Assume that Δx has the conditional probability density $q(\Delta t,x,s)$ if X(t)=x. Thus

$$q(\Delta t,x,s)\Delta s = \text{Prob}[s <= \Delta x <= s + \Delta s \mid X(t) = x]$$

with

$$\int q(\Delta t,x,s)ds = 0$$

Let Q(x,t) be the pdf of X at time t. Then

$$Q(t + \Delta t,x) = \int Q(t, x-s)\, q(\Delta t, x-s,s)ds + o(\Delta t)$$

Since

$$Q(t,x) = Q(t,x)\int q(\Delta t,x,s)ds = \int Q(t,x)\, q(\Delta t,x,s)ds$$

we have

$$Q(t + \Delta t,x) - Q(t,x) = \int [Q(t,x-s)q(\Delta t,x-s,s) - Q(t,x)q(\Delta t,x,s)]ds$$

Expanding the integrand about x

$$Q(t+\Delta t,x)-Q(t,x) = \int [-s\, \partial/\partial x(Q(t,x)q(\Delta t,x,s)) + \tfrac{1}{2}S^2\partial^2/\partial x^2(Qq)+...]ds$$

$$= -\frac{\partial}{\partial x}\int s\,(Qq)|_{t,x,s}ds + 1/2\,\frac{\partial^2}{\partial x^2}\int s^2\,(Qq)|_{t,x,s}ds$$

$$= -\frac{\partial Q}{\partial x}\int sqds + 1/2\frac{\partial^2 Q}{\partial x^2}\int s^2 qds$$

We now make assumtions about the moments of q.

Let

$$E[\Delta x \mid X(t) = x] = \int sq(\Delta t, x, s)ds = b(x)\Delta t + O(\Delta t)$$

and

$$E[(\Delta x)^2 \mid X(t) = x] = \int s^2 q(\Delta t, x, s)ds = a(x)\Delta t + O(\Delta t).$$

Thus

$$Q(t + \Delta t, x) - Q(t, x) = \Delta t[-\partial/\partial x(Qb) + 1/2\,\partial^2/\partial x^2(Qa)] + O(\_t)$$

or letting $\Delta t \longrightarrow 0$

$$\partial Q/\partial t = 1/2\,\frac{\partial^2}{\partial x^2}[a(x)Q] - \partial/\partial x[b(x)Q] \quad 0 < x < 1$$

a singular parabolic partial differential equation.

Since X(t) was the number of alleles of type A in model (1), let

$$x = X(t)/2N.$$

According to assumptions in (1), X(t+1) is binomially distributed, and given X(t) = 2Nx then

$$E[X(t+1)] = 2Nx \quad \sigma^2[X(t+1)] = 2Nx(1-x)$$

let

$$\Delta X = X(t+1) - X(t) \text{ with } \Delta t = 1$$

then

$$E[\Delta X] = E[X(t+1)] - E[X(t)] = 2Nx - 2Nx = 0$$

$$E[(\Delta X)^2] = E[X(t+1) - 2Nx)^2\} = [\_X] = 2Nx(1-x)$$

$$E[(\Delta x)^2] = E[(\Delta X/2N)^2]$$

$$= E[(\Delta X)^2]/4N^2 = 2Nx(1-x)/4N^2 = x(1-x)/2N$$

Thus in the diffusion model

$$a(x) = x(1-x)/2N \text{ and } b(x) = 0.$$

If we rescale t by absorbing the factor 2N in t,

we obtain the following equation for the probability density function :

$$Q(x_0,x;t) = P\{X/2N = x \text{ at } t = t \mid X = 2Nx_0 \text{ at } t=t_0\}$$

$$\partial Q/\partial t = [x(1-x)Q]_{xx} \quad ; \quad 0 < x < 1 \qquad (1)$$

$$Q(x_0,x;0) = \delta(x-x_0) \qquad (1A)$$

$$df(0,t)/dt = 1/2\ Q(0,t)$$
$$df(1,t)/dt = 1/2\ Q(1,t) \qquad (1B)$$

These last two equations(1B) describe the rate at which fixation occurs at the boundaries $x = 0$ and $x = 1$.

To solve (1) we assume a separation of variables solution

$$Q = X(x)T(t) \qquad (2)$$

then

$XT' = [x(1-x)X]_{xx}T$ which can be reduced to :

$$T'/T = -\lambda_\lambda ; \quad [x(1-x)X]_{xx} = -\lambda_\lambda X$$

thus

$$x(1-x)X'' + 2(1-2x)X' - (2-\lambda_\lambda)X = 0 \qquad (3)$$

Equation (3) is the Hypergeometric equation :

$$x(1-x)y'' + [c-(a+b+1)x]y' - aby = 0$$

whose solution is

$$y = AF(a,b;c,x) + Bx^{1-c}F(a-c+1,b+1-c;2-c,x)$$

where

$$F(a,b;c,x) = 1+ab/c(x)+a(a+1)b(b+1)/c(c+1)2!(x^2)+...$$

from (3) it is evident that $c = 2$ and since

Q is to be bounded at $x = 0$ we have $B = 0$ further, comparing coefficients in (3) with those in the hypergeometric differetial equation we have

$$b = 3-a \text{ and } \qquad a = 1/2[3+\sqrt{1+4\lambda_x}].$$

for (3)

$$X(x) = 1 + ab/2(x) + a(a+1)b(b+1)/3!2!(x^2 + \ldots$$

at $x = 1$,

$$X(1) = 1 + ab/2 + a(a+1)b(b+1)/3!2! + \ldots$$

We note that if either a or b or both are negative integers or zero then X is a polynomial.To see for which values the series converges we use Raabe's test :

$$\lim_{n\to\infty} n(a_n/a_{n+1} - 1) = L; \text{ for } L < 1 \text{ diverge}, \quad L > 1 \text{ conv.}$$

Here

$$a_n/a_{n+1} = 1 + (3-a-b)/n + O(1/n_2).$$

This implies that the series converges if $3 - a - b > 1$ .But in our problem $3-a-(3-a) = 0$, and therefore the series diverges. We conclude that for solution to (3) to exist a ,b must be negative integers or zero.

By letting

$$a = -(i-1), \; i = 1, 2, 3, \ldots$$

and

$$b = 3-a = 2 + i$$

we have the eigenvalues of the problem.

$$\lambda_i = i(i+1)$$

and

$$Q_i(x,t) = F(1-i,2+1;2,x)e^{-i(i+1)t/4}$$

or

$$Q(x,t) = \sum_{i=1}^{\infty} C_i F(1-i,2+i;2,x)e^{-i(i+1)t/4} \qquad (5)$$

where F(1-i ,2+i ;2, x) is always a polynomial.

To determine C_i ,we apply the generalised Fourier series method together with equation (2) (normalization integral - orthogonal functions) giving :

$$C_i = x_o(1-x_o)\, i(i+1)(2i+1)F(1-i,2+i,2,x_o)$$

Thus the required solution that satisfies the singular diffusion equation (3) can be expressed in hypergeometric function as follows

$$Q(x_o,x;t) = \sum_{i=1}^{\infty} x_o(1-x_o)i(i+1)(2i+1)F(1-i,2+i,2,x_o) \times F(1-i,2+i,2,x)e^{-i(1+i)t/4} \qquad (6)$$

the probability of fixation at x = 0 and x = 1 are given by

$$f(0,t) = (1-x_o) + \sum_{i=1}^{\infty}(2i+1)(1-x_o)x_o F(i+2,1-i,2,1-x_o) \times (-1)^i e^{-i(i+1)t/4}$$

$$f(1,t) = x_o + \sum_{i=1}^{\infty}(2i+1)x_o(1-x_o)F(i+2,1-i,2,x_o) \times (-1)^i e^{-i(i+1)t/4}$$

we observe that at any time

$$f(0,t) + \int Q(x_o,x;t)\,dx + f(1,t) = 1$$

Based on equation (6) the process of change in

probability distribution of gene frequency when the population starts at x = 0.5, 0.1 and 0.3 is illustrated in figures 1, 2 and 3 respectively.

DISCUSSION of RESULTS

All of the models display (figures 1,2,3,4) the diffusion of genes through the population. Initially the graphs are very peaked but with increasing time , the graphs flatten out. After 2N generations the graph is almost linear (uniform distribution) which is confirmed by the solution given in (5). From this formula when we take the leading term we see that it dominates for large t, i.e. $Q(x, x_o; t) \sim Ce^{-t/2N}$ for $t \to \infty$. For all cases the $\int Q(x, x_o; t)\, dx$ decreases with time.This is due to the fixation occuring at x = 0 and x = 1.

Figure 1 shows a maximum probability density at the same gene frequency as the initial gene frequency (0.5), with the first generation (t = N/10) being the most pronounced. This is not unusual since $Q(x_o, x, t)$ approaches zero as t approaches infinity,more directly a small t produces a large Q(t is measured in generations). For the generations beyond 2N; the curves are flat and all frequencies seem equally probable. That is , fixation or loss of the allele in question proceeds at a constant

rate. For the smaller generations (t<=2N) the proportion of alleles lost is larger than the proportion of alleles fixed in a given population.

Figures 2 and 3 are more suitable for use as comparison to figure 1. Here the initial gene frequency is 0.1 (figure 2). Fixation occurs very rapidly at x = 0. Figure 3 shows characteristics similar to figure 1 (maximum Q occuring at initial gene frequency). However in some generations (t=N/2 , N , and 2N) the maximum seem to occur prior to 0.3. Also at least 4 generations are required before all gene frequencies become equally probable.

Discrete Model

Following the Monte Carlo simulation (appendix A2) of model (1) figures 4, 5, 6, and 7 were prepared. Included in these are the results of the transition matrix at the indicated observation periods. Also included for comparison is the solution of the continuous model (previously discussed) for certain generation. Further the matrix results provide a standard for direct comparison with the Monte Carlo simulation. As seen in figure 4, that the simulations compare very favorably with the matrix result. A χ^2 for each generation examined does show that there are no significant difference between the two sets of results

obtained and indeed the Monte Carlo scheme used is reliable. For the generation displayed in figure 4 (t = N/10) only 5 percent of the time this method will yield poor results.

Even though there is good comparison between the exact and Monte Carlo results , the matrix result is symmetric about the class mark 10 while there is some skewness in the Monte Carlo result. The absence of skewness is due in part to the underlying computations that produce the row vectors of the transition matrix (theoretical binomial density function).The continuous solution compares better with the matrix result (area under the curve and symmetry) than with the Monte Carlo result. In the case of figure 5 (t = N/5) the χ^2 indicates that 2 percent of the time the method may give poor result.However the standard deviation is larger here than in the previous case(σ^2=3.2 vs 2.3 for t = N/10 case). Also the skewness is more pronounced. This increase in standard deviation is due in part to the wide variation at both the class marks 7 and 13 . As t gets large both ends of the fixed classes should have large proportion of the gene pool. Even though the continuous model(figure 6) does not show this,the Monte Carlo and matrix results do. Again relatively large variations

occur more frequently,resulting in even larger standard deviation than before,but the overall method is stil good. In figure 7 except for the fixed classes the matrix and Monte Carlo result in a flat profile.That is, for this generation (t = 2N) the gene frequency of the unfixed classes are becoming equally probable; but since there were some losses and fixations prior to this generation, there is a cumulative effect for both the fixed classes at x = 0 and x = 1.

PROBABILITY DENSITY AS A FUNCTION
OF
GENE FREQUENCY (AT $t=0$ FREQUENCY IS 0·5)

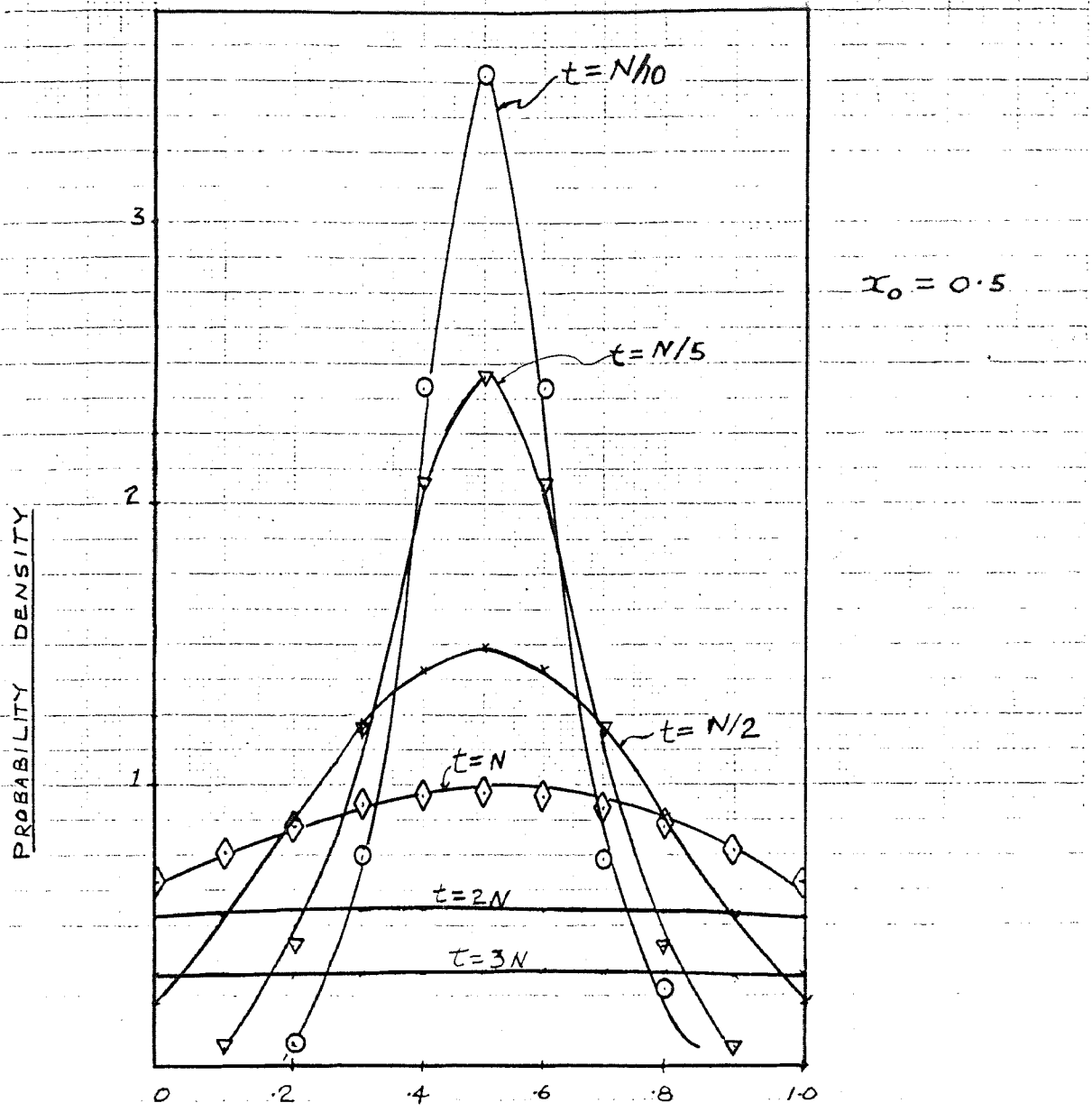
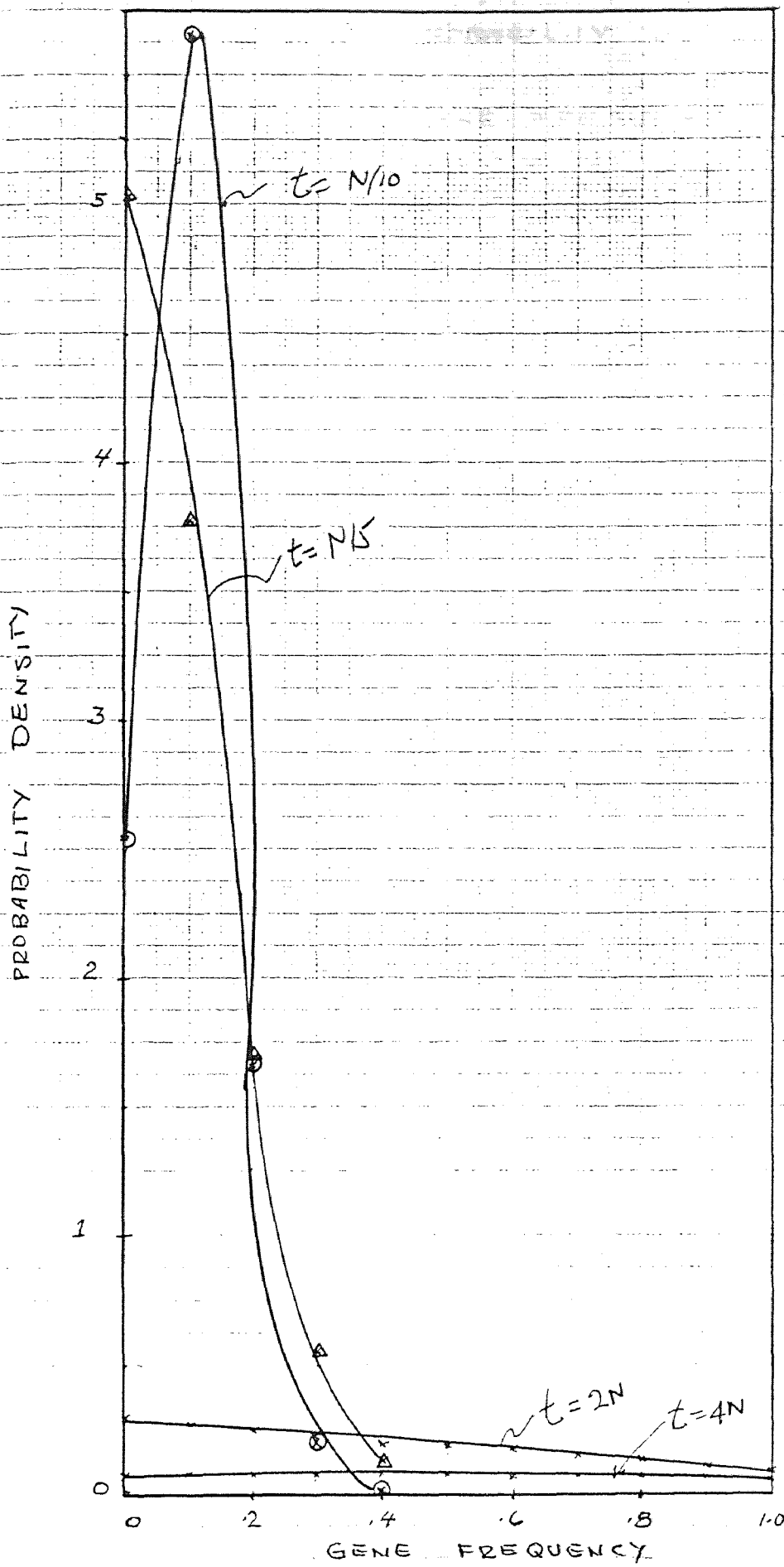


FIGURE 1



$x_0 = 0.1$

FIGURE 2

PROBABILITY DENSITY AS A FUNCTION OF GENE FREQUENCY (at $t=0$ $x_0 = 0.3$)

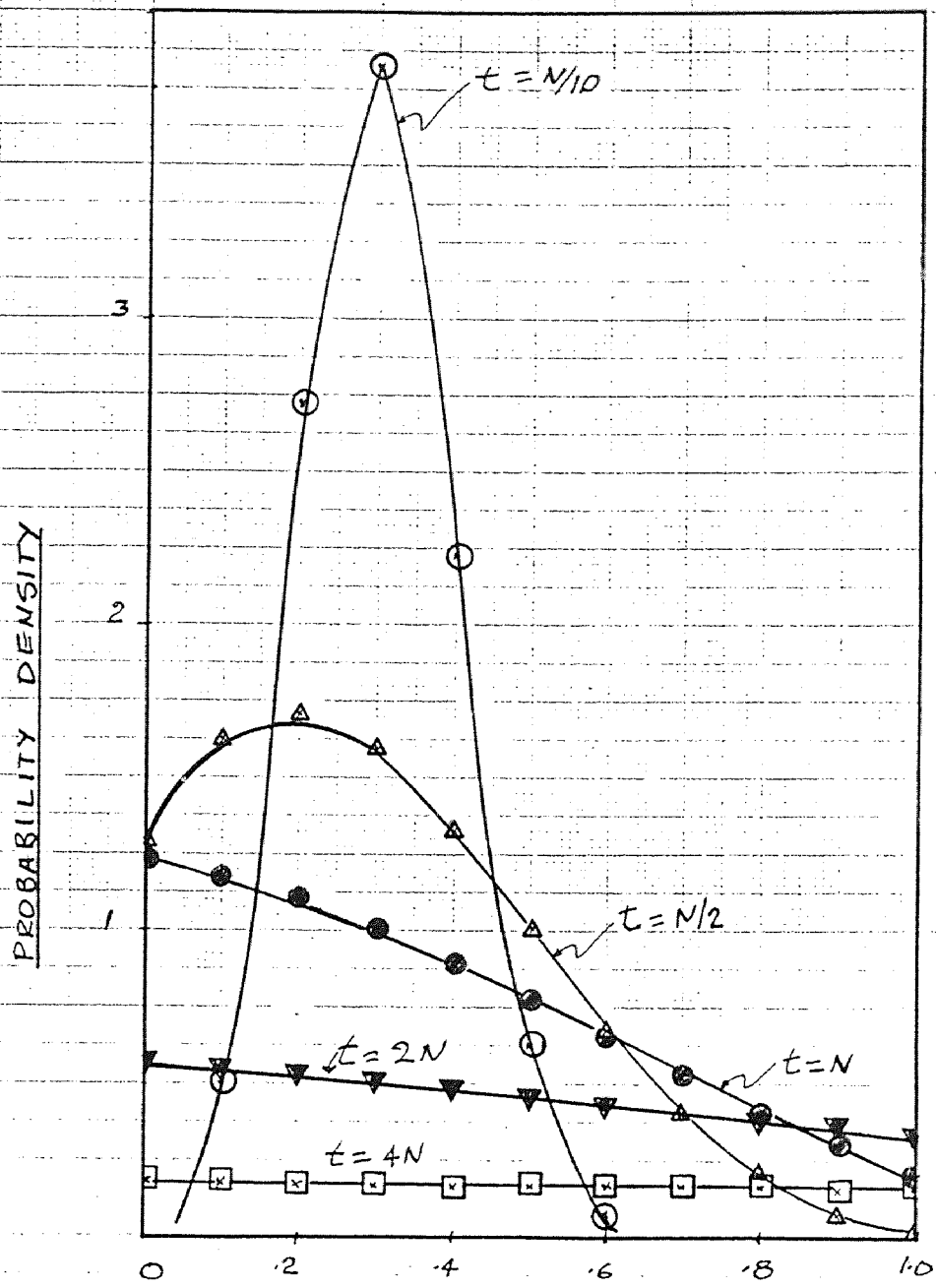
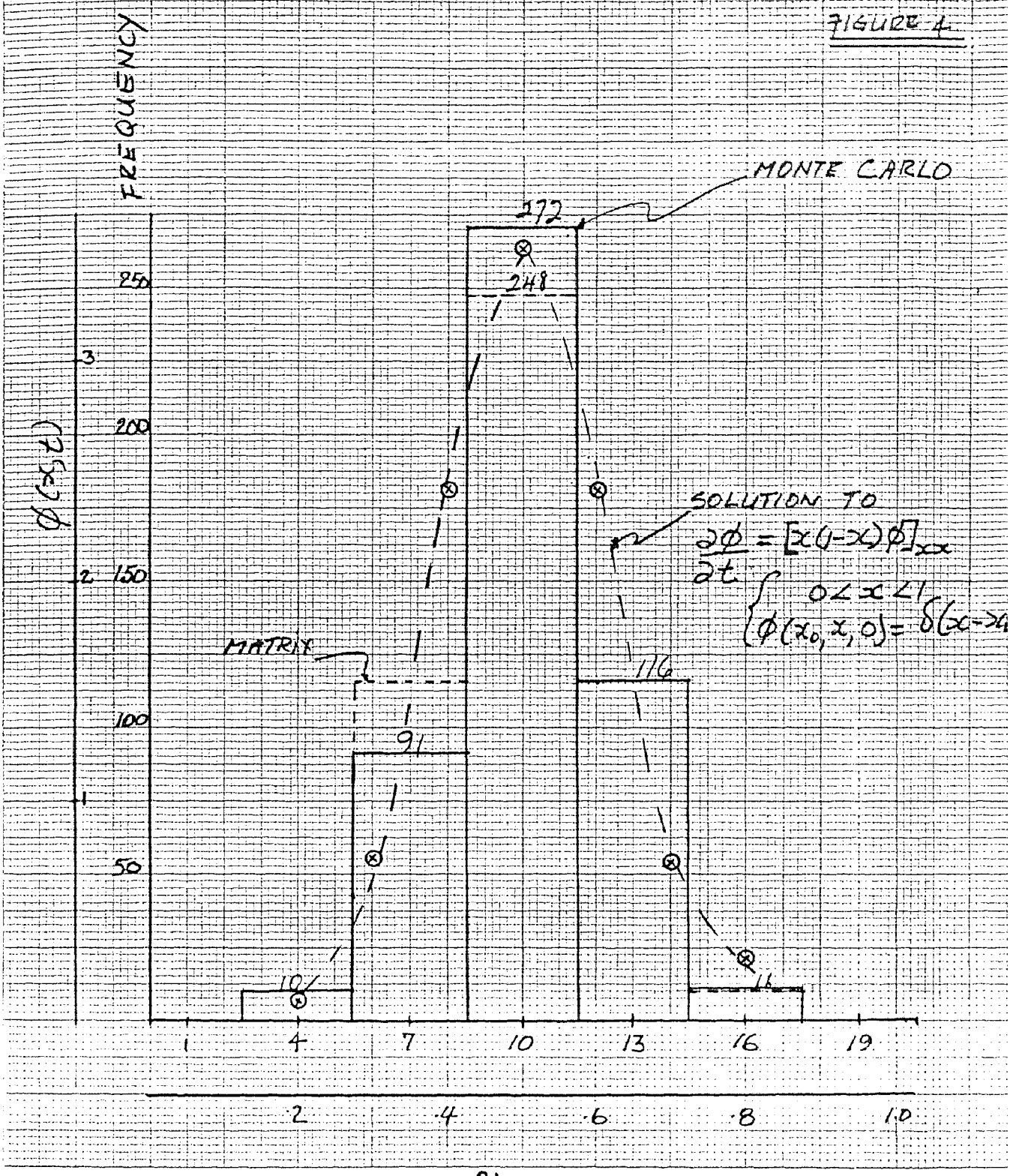


FIGURE 3

FREQUENCY HISTOGRAM OF ONE ALLE
IN A TWO-ALLELE POPULATION (20)
OBTAINED FROM 500 TRIALS OF AN
EXPERIMENT (1st GENERATION de t=

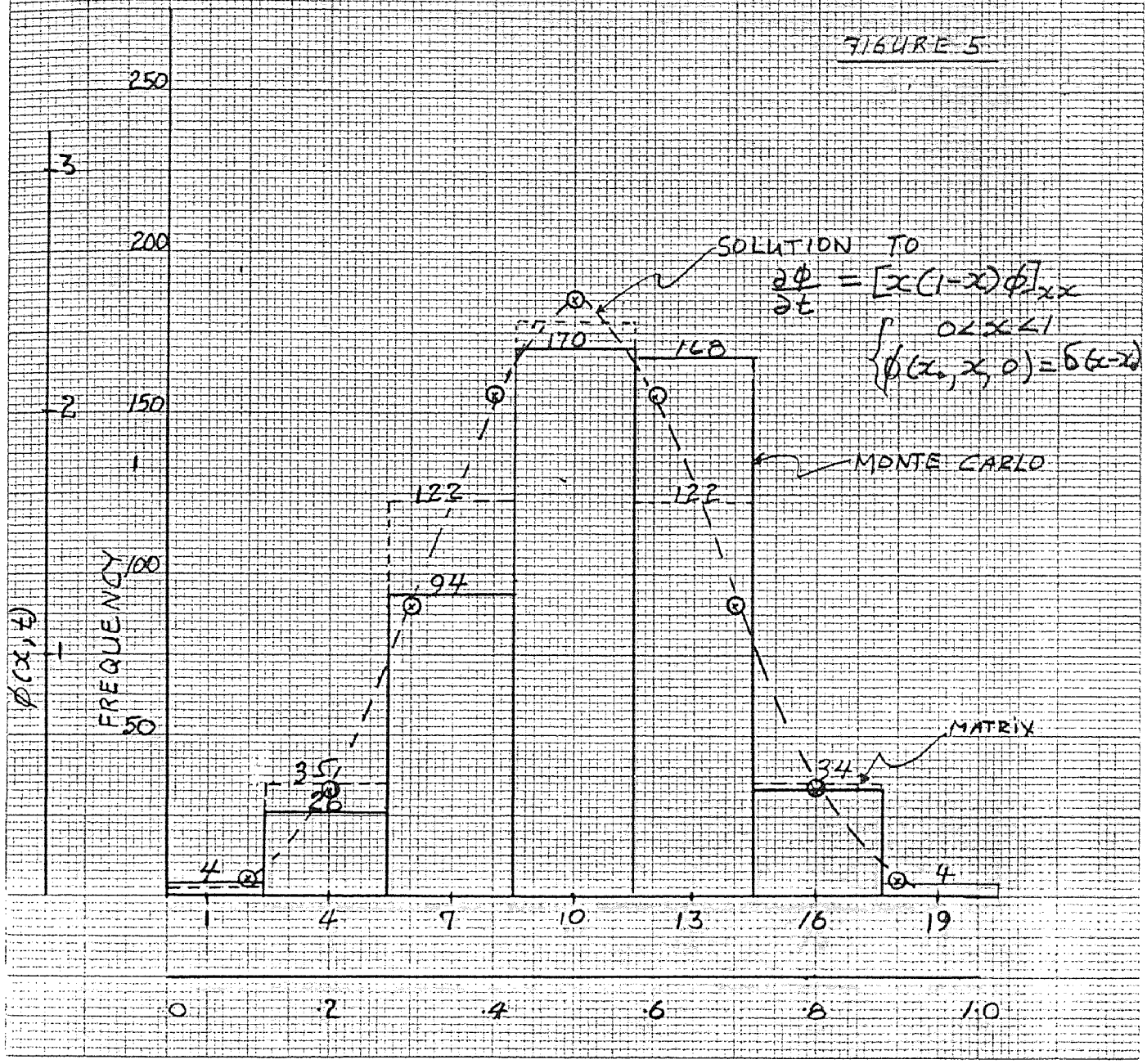
FIGURE 4.



's to the inch

FREQUENCY HISTOGRAM OF ONE ALLELE IN A TWO-ALLELE POPULATION (20) OBTAINED FROM 500 TRIALS OF AN EXPERIMENT (2ND GENERATION IE, t=N/5)

FIGURE 5

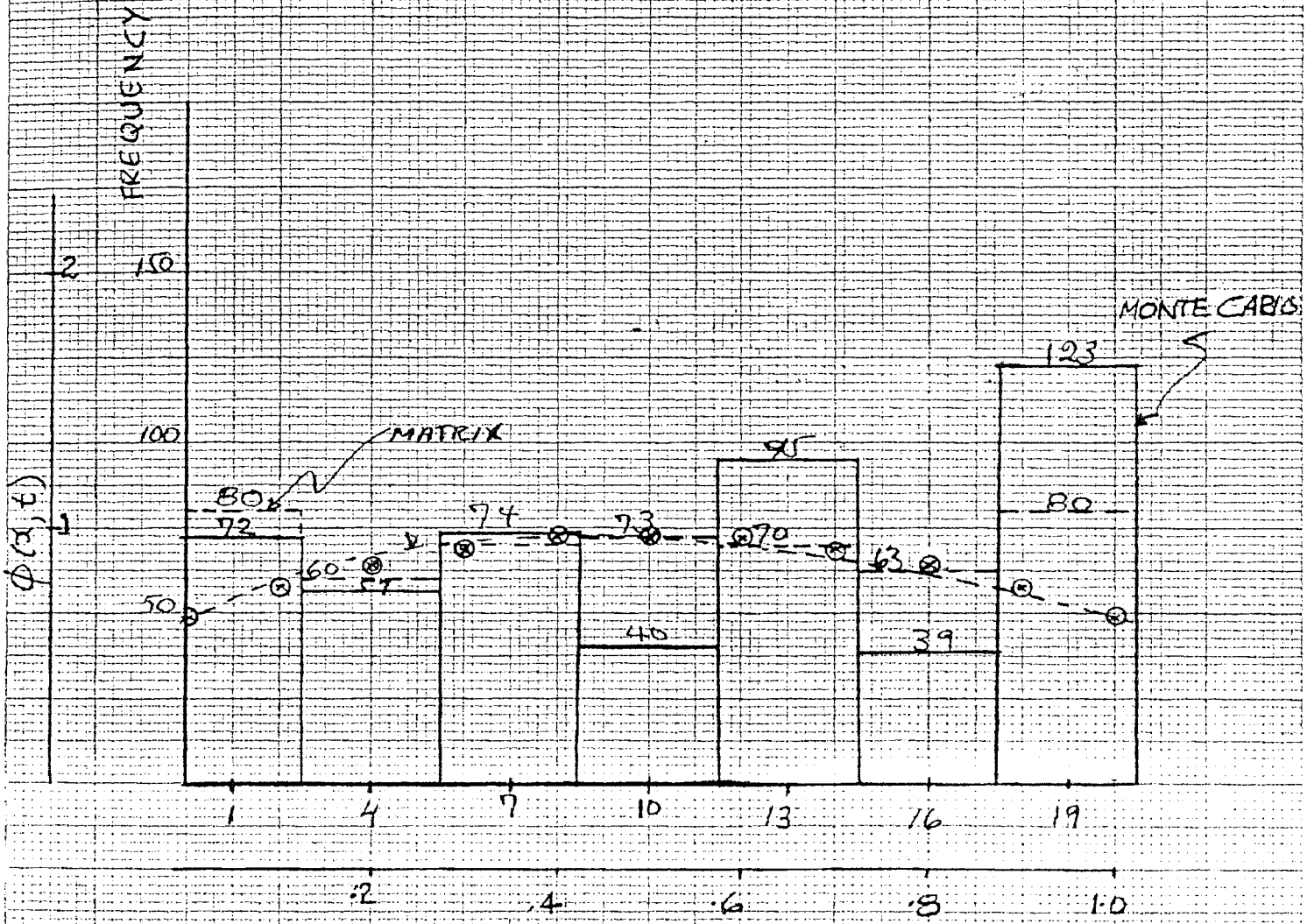


FREQUENCY HISTOGRAM OF ONE ALLELE IN A TWO-ALLELE POPULATION (20) OBTAINED FROM 500 TRIALS OF AN EXPERIMENT (10th GENERATION ie. $t = N = 10$)

FIGURE 6

⊗ SOLUTION TO $\frac{\partial \phi}{\partial t} = [x(1-x)\phi]_{xx}$

$$\begin{cases} 0 < x < 1 \\ \phi(x_0, x, 0) = \delta(x - x_0) \end{cases}$$

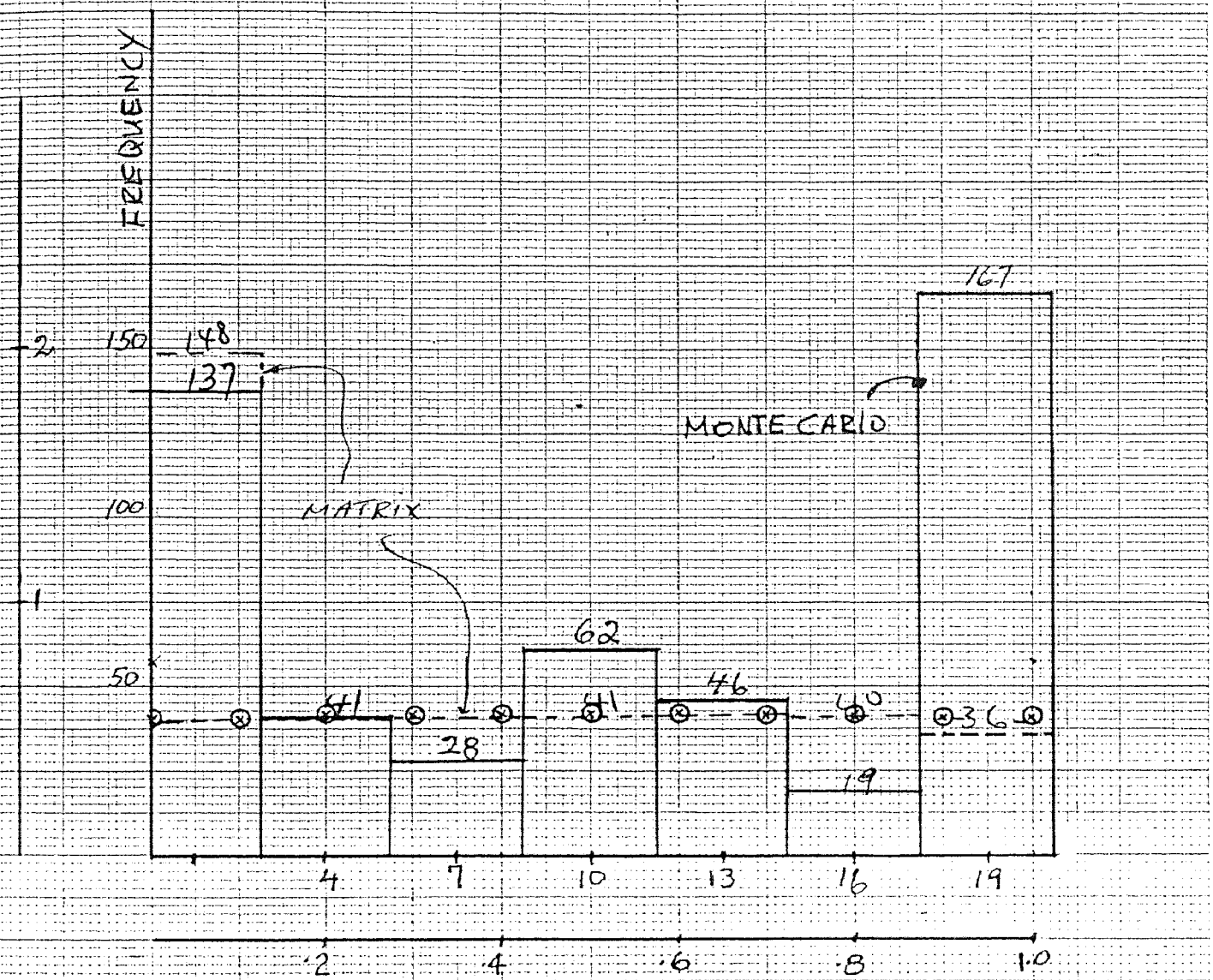


FREQUENCY HISTOGRAM OF ONE ALLELE IN A TWO-ALLELE POPULATION OBTAINED FROM 500 TRIALS OF AN EXPERIMENT (20^{TH} GENERATION i.e. $t=2n = 2N$).

FIGURE 7

⊕ SOLUTION TO $\frac{\partial \phi}{\partial t} = [x(1-x)\phi]_{xx}$

$$\begin{cases} 0<x<1 \\ \phi(x_0, x, 0) = \delta(x-x_0) \end{cases}$$



References

1. Ewens,W. J. "Poulation Genetics" London Metheu(1969)

2. Sobol, I.M. "The Monte Carlo Method" University of Chicago Press.

3. Crow,J .F. and Kimura,M. "An Introduction to Population Genetics" Harper and Row (1970).

4. Ludwig,D. (1974).

CHAPTER II

EPIDEMICS

Over the years , models of various degrees of mathematical complexities have been developed to study a variety of epidemics. Such studies are complicated for various reasons. The differing etiologies of the diseases lead to some difficult to analyse non-linear models. Nevertheless there exist some models of both deterministic and stochastic nature which possess characteristics associated with many diseases. We will examine some of these models.In both the deterministic and stochastic cases, we have the following five assumptions :

1) Following introduction of the disease into the community, the total population size remains fixed.
2) Everyone in the community is initially susceptible to the disease.
3) Everyone who has contracted the disease and has recovered is immune.
4) The disease is spread by direct contact between a susceptible person (susceptible) and an infected person (infective)
5) The infectives are introduced into the community independently.

DETERMINISTIC THEORY

The simplest deterministic model that we consider first, already possess a characteristic that plays a dominant role in most models. Unless the size of the

infective group reaches a certain "treshold" level, the disease is not likely to spread. We consider a community of N individuals, all susceptible, into which an infective is introduced. The population size remains fixed at N + 1 , x is the number of susceptible, y is the number of infective and z those removed are all continuous variables as is t time. A relationship that holds for all time is :

$$x + y + z = N + 1 \qquad (1)$$

On the basis of assumption (5) the following sketch shows the three classes to which an individual can belong.

where S = susceptible [x(t)]

I = infective [y(t)]

R = removed (isolated or immune) [z(t)]

The length of time a member of this closed population belongs to one of these classes is not fixed. We assume the " law of mass action " where the rate at which new infectives are generated is proportional to the product of both the susceptible and infective population, diminished only by those that are removed. The removal class is increasing at the rate proportional to the infective class and the susceptible are diminished by the factor βxy where x and y are as defined earlier .

Thus the following equations govern this process.

$$dx/dt = -\beta xy \qquad (A)$$

$$dy/dt = \beta xy - \gamma y \qquad (B)$$

$$dz/dt = \gamma y \qquad (C)$$

where B is infection rate and is removal rate. To obtain a solution to the above system of differential equations we make the substitution

$$\rho = \gamma/\beta \qquad (3)$$

in (2) after deviding (2B) by (2A) thus :

$$dy/dx = (\beta xy - \gamma y)/-\beta xy = \rho/x - 1$$

where ρ is the relative removal rate. Then

$$dy = (\rho/x-1)dx + c$$

reduces to

$$y = \rho \ln x - x + c$$

A relationship for c is deduced by noting that at t = 0

$$x_0 + y_0 = N \quad ; \; z = 0$$

therefore

$$y_0 + x_0 - \rho \ln x_0 = n - \rho \ln x_0 = c$$

such that

$$y = N - x - \rho \ln x_0/x \qquad (4)$$

Substitution of (4) into (2A) gives

$$dx/dt = -\beta x[N-x-\rho \ln x_0/x] \qquad (4A)$$

and the substitution of (4) into (2C) gives

$$dz/dt = \rho\beta y = \rho\beta(\rho \ln x/x_0 - x - N) \qquad (4C)$$

We can rescale time by letting

$$T = \beta t$$

such that

$$dT/dt = \beta = \gamma/\rho .$$

Thus T is the the new time scale. If each of equations 4A, 4B and 4C is rescaled we obtain

$$dx/dT = -[Nx - x^2 - \rho x \ln(x_0/x)] \qquad (5A)$$

$$dy/dT = \rho \ln(x_0/x)(\rho - x) + (N+\rho-x) - \rho N \qquad (5B)$$

and

$$dz/dT = \rho(N-x-\rho \ln x_0/x) \qquad (5C)$$

result. Equation (5C) provides an independent check on the results of (5A) and (5B), since relation (1) must always be satisfied. Further, noting that each of 5A,5B and 5C are of the form

$$x' = f(x)$$

$$y' = g(x)$$

$$z' = h(x)$$

the system can be solved numerically. Since there is only one independent variable we can use Simpsons rule to integrate each of the equations. Figures 1 and 2 were constructed with results from this integration technique.

II STOCHASTIC THEORY - CHAIN BINOMIAL MODELS

In the following models we assume that into a homogeneously mixing population of susceptibles an infective is introduced. We choose as a unit of time,

the mean length of the infectious period. Since at each stage of the epidemic there are susceptibles and infectives, we assume that at the next stage the new crop of cases is binomially distributed. Possible chains in a household of 3 (2 susceptibles and one infective) are :

$$1,\ 1^2,\ 1^3,\ \text{and}\ 12.$$

The case 1 is the case where at time 1 there is one infective and at subsequent times none, 1^2 is the case stage 1 , 1 infective and stage 2 also 1 infective , 12 is the case of 1 infective at stage 1 and 2 infectives at stage 2. We consider two different models, the Reed Frost and Greenwood models. We let I_t and S_t be the number of infectives and susceptibles at time t and p = 1-q is the probability of adequate contact between any two members of the group at time t.

To derive the binomial distribution we observe that since p is the probability of contact between any two members of the population ,q is the probability that these two members will not meet and $q^I t$ the probability that a given susceptible will not meet with any of the I_t infectives. Thus the probability that a given susceptible will meet with at least one

of the I_t infectives is $(1-q^{I_t})$. Reed- Frost model is the following binomial chain :

$$P(I_{t+1}/S_t, I_t) = \frac{S_t!}{I_{t+1}!\,S_{t+1}!}\left(1-q^{I_t}\right)^{I_{t+1}}\left(q^{I_t}\right)^{S_{t+1}}$$

For the Greenwood model we asume that the chance of infection is not influenced by the size of the infectious population. We assume that the probability of a given susceptible being infected is p. Thus the Greenwood model is the binomial chain given by:

$$P(I_{t+1}/S_t, I_t) = \frac{S_t!}{I_{t+1}!\,S_{t+1}!}\,p^{I_{t+1}}(1-p)^{S_{t+1}}$$

The tables below show that the possible chain and probabilities for the Reed-Frost and Greenwood models are indistinguishable for the case of a small household (household of three) while there is a difference if the household is greater than three.

TABLE I

Type of Introduction	Frequency Reed-Frost	Greenwood
Single	q^2	q^2
	$2pq^2$	$2pq^2$
	$2p^2q$	$2p^2q$
	p^2	p^2

TABLE II

Type of Introduction	no of cases	Frequency	
		Reed-Frost	Greenwood
Single	1	q^2	q^2
	2	$2pq^2$	$2pq^2$
	3	$p^2(1+2q)$	$p^2(1+2q)$

In addition the last table shows one initial case followed by one new case ($2pq^2$). Like occurences are combined to facilitate examination of the total size of an epidemic such as $\langle 1^3 \rangle$ and $\langle 12 \rangle$. In each case two new cases follow the initial case, thus a total of three , giving a frequency :

$$p^2 + 2p^2q = p^2(1 + 2q).$$

We used a Monte Carlo method to simulate epidemics in a population of sizes 2 and 4 into which an infective was introduced. The results are tabulated in tables 1 and 2.

DISCUSSION OF RESULTS

Following numerical integration of equations 5A and 5B, figures 1 & 2 were constructed. It is evident from figure 1 that there exists a relative removal rate (ρ)

below which epidemic occurs and above which epidemic does not occur. For our case ($x(0) = 30$), this relative removal rate is 20. It is reasonable to assume that no true epidemic will occur if the relative removal rate is larger than the initial available number of susceptibles. Therefore for an epidemic to occur the relative removal rate must be smaller than the initial number of susceptibles (i.e. $\rho < x(0)$). However it must be understood that eventhough we may know ρ relative to x at time zero we cannot directly predict ρ for a given population size analytically. Some factors preventing analytic prediction of ρ are the difference in types of diseases and the variability of ρ itself.Therefore as evidenced in figure 1 a small relative removal rate ($\rho = .1$) gives a pandemic whereas a large relative removal rate ($\rho = 40$) gives no epidemic. Figure 2 is a set of epidemic curves based on different relative removal rates.Again it is evident that a total epidemic will occur for the case $\rho = 1$. Further, at this small relative removal rate, approximately 87 percent of the susceptibles will become infected within a very small period of time following contact. What this means is that during the epidemic there will be a majority of infectives and a minority of susceptibles which is enough of a factor to guarantee a pandemic

eventhough there was only one initial case (infective) introduced. On the other hand a large relative removal rate (ρ = 20) would have only 2 percent of the susceptibles becoming infected within the same time period (as for the case ρ = 1) following contact. Tables 1 & 2 are the results of the simulation of the Reed-Frost and Greenwood models. In these tables the frequencies listed in Bailey's book are used for comparison with the respective models result. As is clear from the tables a small probability of contact between infective and susceptible results in no true epidemic. This is expected,since small frequency of contact between individuals implies smaller contact frequency between infective and susceptible, and is similar in effect as a large relative removal rate (previously discused). As the intimacy is improved within the household ($p \rightarrow 1$) total epidemic occurs as is indicated by both models.

Comparison of Models

For small population sizes (including infective) both the Reed-Frost and the Greenwood models are expected to produce identcal results. This is verified in table 2 for a population size of 3. However, table 1 demonstrates differences between the models which are due to the assumption concerning the influence of

chance infection due to the number of infectives available at a time t+1 (see introduction). Individual chains were crosschecked with those expected in Bailey's book [1] and were found to be satisfactory. At this point only the chain type can be determined,for example a chain $\{ 1\ 2\ 1^2 \}$ was generated (among others) for the household of 5 (4 susceptibles).

Since each of these chains occur with a definite frequency, depending on which model is examined, one may estimate the total frequency in a given number of trials by taking the sum of each type of chain generated. As an illustration , the household of three (Bailey[1]) gives :

no epidemic	=	q^2	(frequency)
1 new case	=	$2pq^2$	"
2 " cases	=	$2p^2q$	"
total epidemic	=	p^2	"

Then in an experiment of 500 trials with n_1 single cases occurring this would yield $n_1/500$ to be compared with $2pq^2$, and so on.

To statistically compare the experiments with the expected results we used the " χ^2 goodness of fit test". Since there are eight sets of data for each model we try to decide on the two worst cases over the range of contact frequency for each model. For example

at a p of 0.25 (low end) and at a p of 0.45 (high end). The quantity :

$$\chi^2 = \sum_{i=1}^{n} (O_i - e_i)^2 / e_i$$

for the Greenwood model is determined ; where O_i is the observation (simulation) while e_i is the expected (predicted).

$p = 0.25$

Greenwood		Model
O_1	= 51	e_1 = 49.
O_2	= 102	e_2 = 94.5
O_3	= 108	e_3 =109.5
O_4	= 89	e_4 = 89.
O_5	= 150	e_5 =158.

From the above values χ^2 is 1.102; and from statistical tables with a degree of freedom of 4

$$\chi^2 \ll \chi^2_{.95} = \chi^2_{crit}$$

This indicates no significant difference between the expected and predicted values.
Similarly at p = 0.45, χ^2 = 8.6 ; no significant difference between the observed and predicted results.

Below is the result of a similar analysis on the Reed-Frost model

Reed-Frost model

p	χ^2	χ^2_{crit}	Deg of freedom
0.35	2.65	9.49	4
0.45	2.30	9.49	"

From the result presented in the above tables both models are well represented by their respective simulation.

Conclusion

The epidemic curve in figure 2 and the phase portrait in figure 1 both emphasize the influence of the relative removal rate. In our case (n = 31) the relative removal rate is 20 , above which no epidemic would seem to occur. These figures also demonstrate the existence of a treshhold population size.

Both the Reed-Frost and Greenwood stochastic models are well represented (χ^2-goodness of fit) when the Monte Carlo method is used. The models indicate that in order for serious epidemic to occur the frequency of contact between susceptibles and infective must be relatively large .

Table 1

Greenwood		no. of cases	p	Reed-Frost	
Model	Prediction			Model	Prediction
15.7	15.6	3	0.25		
28.7	28.1	2			
55.3	56.3	1			
22.9	21.6	3	0.30	21.8	21.6
29.5	29.4	2		31.8	29.4
47.3	49.0	1		46.4	49.0
31.7	28.2	3	0.35	28.2	28.2
28.7	29.6	2		28.4	29.6
39.3	42.2	1		43.4	42.2
34.7	35.2	3	0.40	36.8	35.2
32.7	28.8	2		29.6	28.8
32.3	36.0	1		33.6	36.0
43.7	42.5	3	0.45	47.4	42.5
26.3	27.2	2		24.2	27.2
29.7	30.3	1		28.4	30.3
47.1	50.0	3	0.50	48.8	50.0
26.5	25.0	2		27.6	25.0
26.1	25.0	1		23.6	25.0
10.5	10.4	3	0.20	10.0	10.4
22.5	25.6	2		27.8	25.6
66.7	64.0	1		62.2	64.0

Household of 3 - Monte Carlo simulation of the models

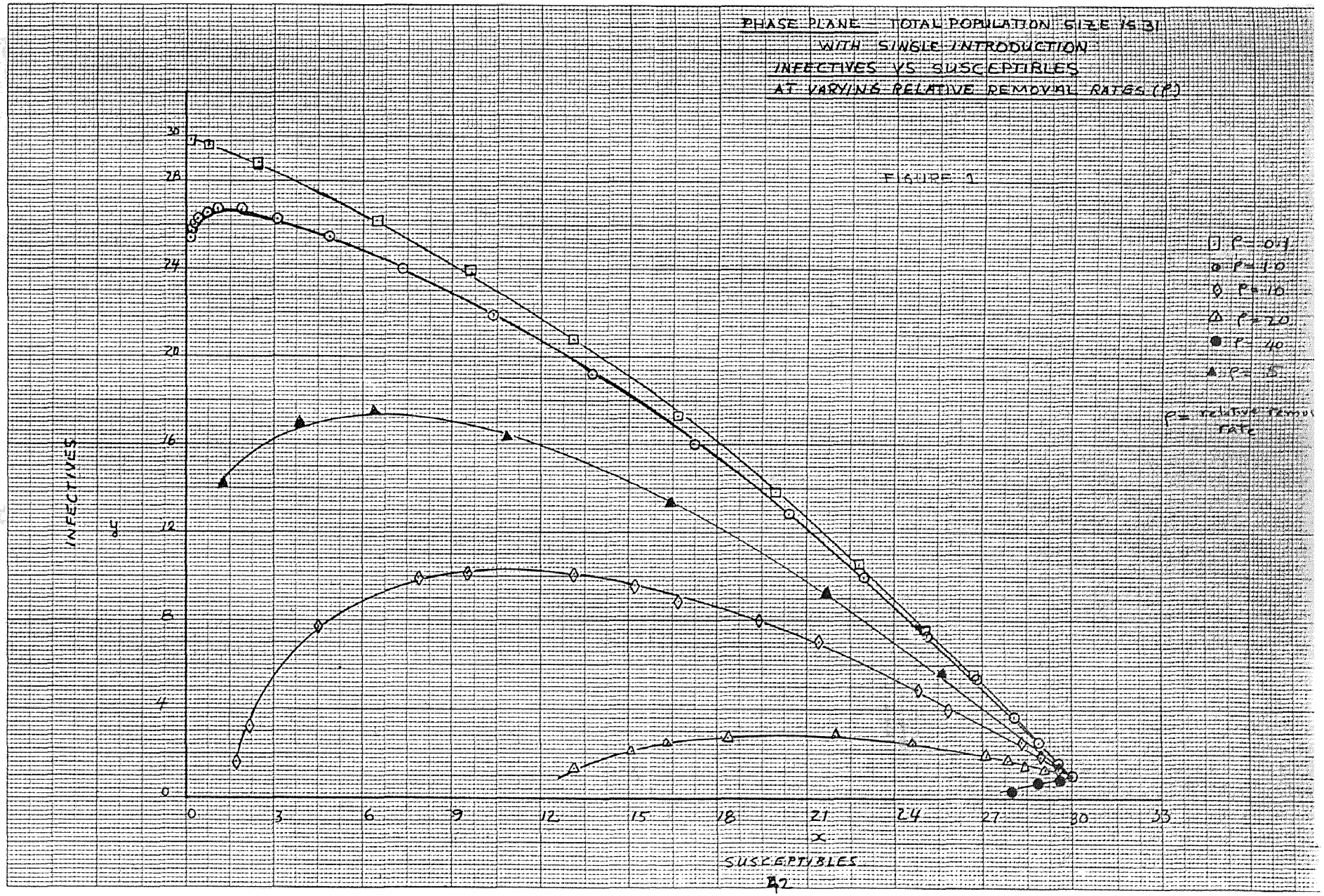
Table 2

Greenwood		no of cases	p	Reed-Frost	
Model	Prediction			Model	Prediction
2.0	1.98	5	0.15	2.4	3.3
7.2	7.5	4		6.4	8.2
15.8	15.7	3		16.2	13.7
19.6	22.6	2		22.2	22.6
55.4	52.2	1		52.8	52.2
10.2	9.8	5	0.25	13.4	15.8
20.4	18.9	4		19.0	18.1
21.6	21.9	3		18.2	16.7
17.8	17.8	2		18.4	17.8
30.0	31.6	1		31.0	31.6
25.0	23.8	5	0.35	39.0	37.0
28.0	27.4	4		23.0	21.8
20.6	20.3	3		10.8	12.7
11.2	10.6	2		10.4	10.6
15.2	17.9	1		16.8	17.9
31.8	32.3	5	0.40	46.2	48.9
31.2	29.4	4		21.2	20.8
15.8	17.8	3		9.8	9.9
8.4	7.5	2		8.4	7.5
12.8	12.9	1		14.4	12.9
38.0	41.2	5	0.45	63.4	60.6
33.8	29.9	4		16.2	18.2
17.0	14.8	3		7.4	7.0

Table 2 - cont'd

4.2	4.9	2		4.6	4.9
7.0	9.2	1		8.4	9.2
51.0	50.0	5	0.50	72.8	71.1
26.8	28.9	4		14.0	14.8
12.4	11.7	3		5.2	4.6
4.0	3.1	2		3.4	3.1
5.8	6.3	1		4.6	6.3
60.2	58.4	5	0.55	81.4	80.0
25.4	26.8	4		11.2	11.2
8.4	8.8	3		2.0	2.9
2.4	1.8	2		2.4	1.8
3.6	4.1	1		3.0	4.1
61.4	66.2	5	0.60	86.8	87.0
25.8	23.9	4		8.0	7.8
8.6	6.2	3		1.6	1.6
1.6	0.98	2		1.0	0.98
2.6	2.6	1		2.6	2.6

Household of 5 - Monte Carlo simulation of the models



References

1.Sobol, I. M. "The Monte Carlo Method" University of Chicago Press.

2.Bailey, N. T. "Mathematical Theory of Epidemics"

APPENDIX

RANDOM NUMBER TABLE

Table A. 400 random numbers

86 515	90 795	66 155	66 434
69 186	03 393	42 502	99 224
41 686	42 163	85 181	38 967
86 522	47 171	88 059	89 342
72 587	93 000	89 688	78 416
52 452	42 499	33 346	83 935
76 773	97 526	27 256	66 447
04 825	82 134	80 317	75 120
87 113	84 778	45 863	24 520
84 754	57 616	38 132	64 294
56 558	12 332	94 377	57 802
88 955	53 758	91 641	18 867
33 181	72 664	53 807	00 607
67 248	09 082	12 311	90 316
27 589	99 528	14 480	50 961
79 130	90 410	45 420	77 757
25 731	37 525	16 287	66 181
45 904	75 601	70 492	10 274
19 976	04 925	07 824	76 044
15 218	49 286	89 571	42 903

Table B. 88 random normal variables **

0.2005	1.1922	—0.0077	0.0348
1.1609	—0.6690	—1.5893	0.5816
0.5864	—0.9245	0.0904	1.5068
0.1425	—0.2863	1.2809	0.4043
0.9516	—1.7708	2.8854	0.4686
—0.5863	0.8574	—0.5557	0.8115
1.1572	0.9990	—0.1032	0.5405
—1.4428	—0.5564	—0.5098	—1.1929
—0.3924	1.7981	0.6141	—1.3596
0.8319	0.4270	—0.8888	0.4167
0.9780	—0.7679	0.8960	0.5154
1.0423	—1.8149	1.1803	0.0033
1.8818	0.7390	—0.2736	1.0828
—1.1147	0.2776	0.1012	—1.3566
0.6379	—0.4428	—2.3006	—0.6446
1.4664	1.6852	—0.9690	—0.0831
—0.2676	—1.2496	—1.2125	1.3846
—0.6022	0.0093	0.2119	—1.4647
—0.0572	—0.5061	—0.1557	—1.2384
1.4943	—0.4406	—0.2033	—0.1316
—0.8513	1.1054	1.2237	—0.7003
0.7165	0.8563	—1.1630	1.8800

* Random numbers simulate the values of a random variable with distribution (22) (see 3.1)
** Normal variables simulate the values of a normal (Gaussian) random variable ζ with $a = 0$, $\sigma = 1$

```
15  FOR I = 0 TO 500
16 A(I) = 0:BB(I) = 0:C(I) = 0
17 E(I) = 0:D(I) = 0
20 N1 = N2N / 6:TIME = 1
21  PRINT
25 TRIAL = TRIAL - 1
30 MR = NI / N2N
40 J = 0:P = MR
42  IF TRIAL = 0 THEN 207
45  GOSUB 5010
50 B(1) = PRB
60  FOR L = 1 TO N2N
70 J = L:P = MR
71  GOSUB 5010
80 B(L + 1) = B(L) + PRB
100  NEXT L
110 X =  RND (1)
120  REM  COMPARE
130  FOR LL = 1 TO N2N + 1
140  IF X < B(LL) GOTO 165
160  NEXT LL
165  PRINT "  ";LL - 1;
168 RANVARX = LL - 1
170 R = TIME
176  IF ((R = 1) OR (R = 2) OR (R
      = 3)) THEN  GOSUB 6000
177  IF ((R = 4) OR (R = 5)) THEN
      GOSUB 6000
178  IF ((LL = 1) OR (LL = N2N +
      1)) THEN 20
180 TIME = TIME + 1
200 NI = LL - 1
205  GOTO 30
207 BB(N2) = A(N2) + BB(N2):C(N2)
      = C(N2) + BB(N2):D(N2) = D(
     N2) + C(N2):E(N2) = E(N2) +
     D(N2)
208 BB(0) = BB(0) + A(0):C(0) = C
     (0) + BB(0):D(0) = D(0) + C(
     0):E(0) = E(0) + D(0)
210  PRINT A(0),BB(0),C(0),D(0),E
     (0)
211  FOR I = 1 TO N2N - 1
212  PRINT A(I),BB(I),C(I),D(I),E
     (I)
215  NEXT
216  PRINT A(N2),BB(N2),C(N2),D(N
     2),E(N2)
217  DATA  5,501
220  END
5000  REM  BINOMIAL PROB([illegible])
5010  [illegible] 507
5020  [illegible] N2N [illegible]
     [illegible]
[illegible]
5030 PRB = PR[illegible]
[illegible]
[illegible]
[illegible]
[illegible]

6085  RETURN
6090  FOR I = 0 TO N2N
6100  IF (RAN [illegible] = 0 THEN BB(I)
      = BB(I) + 1
6110  NEXT
6115  RETURN
6120  FOR I = 0 TO N2N
6130  IF (RAN - I) = 0 THEN C(I) =
     C(I) + 1
6140  NEXT
6145  RETURN
6150  FOR I = 0 TO N2N
6160  IF (RAN - I) = 0 THEN D(I) =
     D(I) + 1
6170  NEXT
6175  RETURN
6180  FOR I = 0 TO N2N
6190  IF (RAN - I) = 0 THEN E(I) =
     E(I) + 1
6200  NEXT
6210  RETURN
```

Routine using The MonteCarlo METHOD FOR The Discrete MODEL.

Appendix A-2

```
6000  REM [illegible]
      [illegible]
      [illegible]
6005  REM [illegible]
      OCCURRENCE
6010  IF [illegible]
6020  IF [illegible]
6030  [illegible] THEN 6120
[illegible]
[illegible]
[illegible]
[illegible]
```

```
200  READ X0,TIME,X
205  PRINT "X0,SUM,X,I"
206  PRINT X0
210  DIM P(20)
220  DIM Q(20)
222  DIM C(20)
224  DIM CT(20)
226  DIM FX0(20)
228  DIM FFX(20)
229  DIM SUM(20)
230  FOR J = 1 TO 20
231 P(J) = 0
232 Q(J) = 0
233 C(J) = 0
234 CT(J) = 0
235 FX0(J) = 0
236 FFX(J) = 0
237 SUM(J) = 0
238  NEXT J
240 K = X0 * (1 - X0)
245 CT(1) =  EXP ( - 2 * TIME)
250 FX0(1) = 1
260 FFX(1) = 1
270 C(1) = 6
275 I = 1
280 SUM(1) = C(1) * K * CT(1)
285  PRINT SUM(I),X,I,time
290  FOR I = 2 TO 20
295 X0 = 0.5
300 C(I) = I * (I + 1) * (2 * I +
     1)
310 CT(I) =  EXP ( - I * (I + 1) *
     TIME)
320  GOSUB 5000
330 FC = P
340  IF I > 2 THEN  GOTO 410
350 FX0(I) = FX0(I - 1) + P(I - 1
    )
360 X0 = X
370  GOSUB 5000
375 FC = P
380 FFX(I) = FFX(I - 1) + P(I - 1
    )
390 SUM(I) = SUM(I - 1) + C(I) *
     CT(I) * FX0(I) * FFX(I) * K
395  PRINT SUM(I),X,I,Time
400  NEXT I
410  GOSUB 6000
420 FC = QSUM
430 FX0(I) = QSUM + 1
440 X0 = X
450  GOSUB 5000
460 FC = P
470  GOSUB 6000
480 FC = QSUM
490 FFX(I) = QSUM + 1
500 SUM(I) = SUM(I - 1) + C(I) *
     CT(I) * FX0(I) * FFX(I) * K
510  PRINT SUM(I),X,I,Time
520  NEXT I
530 X = X + 0.1
535 X0 = 0.5
540  IF X < 1.1 GOTO 230
550  DATA  0.5,0.05,0
560  END
5000  FOR N = 1 TO (I - 1)
5010 P(N) = ((N - I) * (I + N + 1
     ) * X0) / (N * (N + 1))
5020  NEXT N
5030  RETURN
6000 Q(1) = P(1)
6005 QSUM = Q(1)
6010  FOR M = 2 TO N
6020 Q(M) = P(M) * Q(M - 1)
6030 QSUM = QSUM + Q(M)
6040  NEXT M
6050  RETURN
```

Diffusion Approximation

A-2-1

```
200  READ A, B, N, LPHA, PEOP, RHO
210 H = (B - A) / N
230 X = LPHA
232 YI = 1
234 DTAU = 0
240  PRINT "H , X, Y, DTAU"
250  PRINT H, X, YI, DTAU
260  FOR I = 1 TO N / 9
270 FX0 = X * (X - PEOP + RHO *  LOG
     (LPHA / X))
290 TWOK = (X + H / 2) * (X + H /
     2) - PEOP * (X + H / 2) + RH
     O * (X + H / 2) *  LOG (LPHA
      / (X + H / 2))
295  IF X + H (  = 0 GOTO 370
300 THREK = (X + H) * (X + H) - P
     EOP * (X + H) + RHO * (X + H
     ) *  LOG (LPHA / (X + H))
320 X = X - (H / 6) * (FX0 + 4 *
     TWOK + THREK)
322 YF0 = RHO *  LOG (LPHA / X) *
     (RHO - X) + X * (PEOP + RHO -
     X) - RHO * PEOP
324 YTWOK = RHO *  LOG (LPHA / (X
      + H / 2)) * (RHO - (X + H /
     2)) + (X + H / 2) * (PEOP +
     RHO - (X + H / 2)) - RHO * P
     EOP
325  IF X + H (  = 0 GOTO 370
326 TYHRK = RHO *  LOG (LPHA / (X
      + H)) * (RHO - (X + H)) + (
     X + H) * (PEOP + RHO - (X +
     H)) - RHO * PEOP
328 YI = YI - (H / 6) * (YF0 + 4 *
     YTWOK + TYHRK)
330 D0TAU =  - 1 / (PEOP * X - X *
     X - X * RHO *  LOG (LPHA / X
     ))
332 D2TAU =  - 1 / (PEOP * (X + H
      / 2) - (X + H / 2) * (X + H
      / 2) - (X + H / 2) * RHO *
      LOG (LPHA / (X + H / 2)))
334 D3TAU =  - 1 / (PEOP * (X + H
     ) - (X + H) * (X + H) - (X +
     H) * RHO *  LOG (LPHA / (X +
     H)))
336 DTAU = DTAU + (H / 6) * (D0TA
     U + 4 * D2TAU + D3TAU)
340  PRINT X, YI, DTAU
350  NEXT I
360  DATA  9, 0, 180, 9, 10, 3
370  END
```

Computation for System of ODE "Epidemic"

```
200  READ SSPTBL,TIME,P
201 "SSPTBL=SUSSCEPTIBLE,P=ASSUME
     D MIXING FREQUENCY"
202 "PEOP=TOTAL POPULATION SIZE"
203 "PSRBSS=PROBABILITY OF SUCCES
     S"
204 "PRBFL=PROBABILITY OF FAILURE
     "
205 "RNVARX=PROBABILITY INTERVAL
     WITH WHICH GENERATED RANDOM
     NUMBERS ARE COMPARED205"RNVA
     RX = PROBABIL. INT ER VAL WI
     THWHICHGENER AT EDR AND OMNU
     MBERSARECOPARED"
206 "TIME=0 (ONLY AT START OF RUN
     )"
210 TRIAL = 1
211 00CUR = 0
212 01CUR = 0
214 02CUR = 0
216 03CUR = 0
218 04CUR = 0
219 NET = 1
220  FOR N = 1 TO 10
230 RINVARX(N) = 0
240 RNVARX(N) = 0
250  NEXT N
300 N = 1
310 PEOP = SSPTBL + 1
320 PRBFL = (1 - P) ^ (SSPTBL - T
     IME)
330 PSRBSS = P ^ TIME
320  GOSUB 5000
330 FC = G
340 PBROB = Q * PSRBSS * PRBFL
350 RINVARX(N) = TIME
360 RNVARX(N) = PBROB
370 TIME = TIME + 1
380 N = N + 1
390  IF N <= PEOP GOTO 320
410 SUM(1) = RNVARX(1)
420  FOR N = 2 TO PEOP
430 SUM(N) = RNVARX(N) + SUM(N -
     1)
440  NEXT N
450  FOR I = 1 TO PEOP
455  NEXT I
460 G =  RND (1)
470  FOR N = 1 TO PEOP
485  IF  G <= SUM(N) AND G  = N GOTO
     500
490  NEXT N
500 NET = RINVARX(N) + NET
505  IF RINVARX(K) = 0 GOTO 560
520 SSPTBL = SSPTBL - RINVARX(K)
530  IF SSPTBL = 0 GOTO 560
540 TIME = 0
550  GOTO 220
560 SSPTBL = 4
562 PEOP = SSPTBL + 1
566  IF (PEOP - NET) = 0 THEN 00C
     UR = 00CUR + 1
567  IF (PEOP - NET) = 1 THEN 01C
     UR = 01CUR + 1
568  IF (PEOP - NET) = 2 THEN 02C
     UR = 02CUR + 1
569  IF (PEOP - NET) = 3 THEN 03C
     UR = 03CUR + 1
570  IF (PEOP - NET) = 4 THEN 04C
     UR = 04CUR + 1
575 TIME = 0
580 P = 0.15
590 TRIAL = TRIAL + 1
591 NET = 1
600  IF TRIAL <= 500 GOTO 220
605  PRINT "FINAL SIZES OF EPIDEM
     ICS"
606  PRINT PEOP,00CUR"TIMES"
607  PRINT PEOP - 1,01CUR"TIMES"
608  PRINT PEOP - 2,02CUR"TIMES"
609  PRINT PEOP - 3,03CUR"TIMES"
610  PRINT PEOP - 4,04CUR"TIMES"
611  PRINT "TOT TRIAL="TRIAL,"P-G
     UESS="P,"POPSIZE="PEOP
612  DATA  4,0,0.15
620  END
5000  IF TIME <= 1 GOTO 5070
5010 Q = SSPTBL / TIME
5020 M = TIME - 1
5030  FOR J = 1 TO M
5040 Q = Q * (SSPTBL - J) / (TIME
     - J)
5050  NEXT J
5060  RETURN
5070  IF TIME > 0 GOTO 5100
5080 Q = 1
5090  RETURN
5100 Q = SSPTBL
5110  RETURN
```

GREENWOOD MODEL

```
200  READ SSPTBL,TIME,P
205 NFCTVE = 1
210 TRIAL = 1
211 00CUR = 0
212 01CUR = 0
214 02CUR = 0
216 03CUR = 0
218 04CUR = 0
219 NET = 1
220  FOR N = 1 TO 10
230 RINVARX(N) = 0
240 RNVARX(N) = 0
250  NEXT N
280 N = 1
290 PEOP = SSPTBL + 1
295 NQ = (1 - P) ^ NFCTVE
300 PRBFL = NQ ^ (SSPTBL - TIME)
310 PSRBSS = (1 - NQ) ^ TIME
320  GOSUB 5000
330 FC = Q
340 PBROB = Q * PSRBSS + PRBFL
350 RINVARX(N) = TIME
360 RNVARX(N) = PBROB
370 TIME = TIME + 1
380 N = N + 1
390  IF N <  = PEOP GOTO 300
410 SUM(1) = RNVARX(1)
420  FOR N = 2 TO PEOP
430 SUM(N) = RNVARX(N) + SUM(N -
    1)
440  NEXT N
450  FOR I = 1 TO PEOP
456  NEXT I
460 G =  RND (1)
470  FOR K = 1 TO PEOP
480  IF (G - SUM(K)) <  = 0 GOTO
    505
490  NEXT K
495  PRINT " "
500  PRINT RINVARX(K),"TRIAL ",TR
    IAL
505 NET = RINVARX(N) + NET
510  IF RINVARX(K) = 0 GOTO 560
520 SSPTBL = SSPTBL - RINVARX(K)
530  IF SSPTBL = 0 GOTO 560
535 NFCTVE = RINVARX(K)
540 TIME = 0
550  GOTO 220
560 SSPTBL = 4
562 PEOP = SSPTBL + 1
565 NFCTVE = 1
568  IF (PEOP - NET) = 0 THEN 600
    02 = 02CUR + 1
570  IF (PEOP - NET) = 1 THEN 600
    03 = 03CUR + 1
572  IF (PEOP - NET) = 2 THEN 600
    04 = 04CUR + 1
575 TIME = 1
580 P = 0.4
590 TRIAL = TRIAL + 1
591 NET = 1
600  IF TRIAL <  = 500 GOTO 220
605  PRINT "FINAL SIZES OF EPIDEM
    ICS"
606  PRINT "# OF 00CUR"TIMES"
607  PRINT PEOP - 1,01CUR"TIMES"
608  PRINT PEOP - 2,02CUR"TIMES"
609  PRINT PEOP - 3,03CUR"TIMES"
610  PRINT PEOP - 4,04CUR"TIMES"
611  PRINT "TOT TRIAL="TRIAL - 1,
    "P-ENTERED="P,"POPSIZE="PEOP

612  DATA  4,0,0.4
620  END
5000  IF TIME <  = 1 GOTO 5070
5010 Q = SSPTBL  TIME
5020 M = TIME - 1
5030  FOR J = 1 TO M
5040 Q = Q * (SSPTBL - J) / TIME
    - J
5050  NEXT J
5060  RETURN
5070  IF TIME > 0 GOTO 5100
5080 Q = 1
5090  RETURN
5100 Q = SSPTBL
5110  RETURN
```

REID-FROST Model.

A-4-2.