Conference Proceeding Issue Published in International Journal of Trend in Research and Development (IJTRD), ISSN: 2394-9333, www.ijtrd.com

Introduction to Stochastic Process in Biology Markov Chain Application in Genetics

[1]N. John Britto, [2]Dr. V. Rajagopalan, [3]Dr.S.Sasikala and [4]Dr. S. Muthalagi,
[1]Faculty in Statistics, King Nandhivarman College of Arts and Science, Thellar-604 406, Thiruvannamalai Dt, Tamil Nadu
[1]Research Scholar in Statistics, Annamali University, Tamil Nadu
[2]Profesor, Department of Statistics, Annamalai University, Tamil Nadu
[3]Asst. Prof. in Statistics, Periyar EVR (Govt.) Arts College, Trichy, Tamil Nadu
[4]Asst. Prof. in Zoology, Kolanjiappar Government Arts College, Virudhachalam, Cuddalore District, Tamil Nadu

Abstract: The British Mathematician Hardy and the German Physician Weinberg, in evolution the genetic make up of species and population may change over time, some traits may be lost, other new one arises, while some persist unchanged. Though chance plays a larger role in the inheritance of traits in a single parent cross, understanding how this plays out in the evolution of a population requires a Stochastic modeling. Let forces on a single gene in a large population. To describe the variability of this gene among the population members, we use allels that are certain type of frequency, few nuermical examples reviewed.

Keywords: *Hardy-Weinberg Law, Branching Process, Genetic Threshold, Eigen values and vectors, Fokker-Plank Diffusion equation.*

I. INTRODUCTION

1.1Gene Frequency in population; MN blood type

The presence of each of the allels M and N be detected through antigene tests. A person with genotype MM has type M blood and a person with genotype NN has type N blood (The two allels M and N are thus codominant, as both are equally expressed in the phenotype.

Numerical Example (1);

Suppose in, a population that 60 individuals have type M blood, 101 individuals type MN blood and 53 individuals type N blood for a total population size of 214. Because each person carries two allels of the gene, there are a total of 2 (214) = 428 allels in this data. To determine the frequency of M allels, we note that each person of M blood type carries 2 those of type MN say 1 and those of N carry 0. Thus the frequency of the allel is

M: 2(60) + 1 (101) / 428 = 0.52

N: 1(101) + 2 (53) / 428 = 0.48

Adding gives to 1. Notice, the genotype frequencies in the population are

MM: 60/214 = 0.28, MN:101/214 = 0.47, NN = 53/214 = 0.25

We can use these to calculate allel frequencies also, but because each genotype involves, 2 allels, we have to divide by 2 to account for the change in the number of objects

M: 2 (0.28)/2 + 1 (0.47)/2 = 0.28 + 1/2(0.47) = 0.52

With a similar calculation, giving the frequency of N ½(.49992) = .48

Eggs \Sperms	$A(p)$	$a(q)$
$A(p)$	$AA(p^2)$	$Aa(pq)$
$a(q)$	$Aa(pq)$	$aa(q^2)$

The results, showing frequencies of three genotypes produced due to random mating in a population having allels 'A' with a frequency 'p' and allel 'a' with frequency 'q'. Types of random Mating combinations and their relative frequencies in a population containing $p^2AA, 2pqAa$ and q^2aa genotypes.

Males	Genotypes and frequencies			
	AA		Aa	aa
Females	p^2		$2pq$	q^2
Genotypes and Frequency	AA p Aa pq	$1/2AA+1/2Aa$ $2p^3q$	$1/4AA+1/2Aa+1/4aa$ $4p^2q^2$ $2pq$	$1/2Aa+1/2aa$ aa
Aa				

1.2 Hardy-Weinberg Law and a Markov Chain in Genetics

In a large population of individuals, each of whom posses a particular pair of genes is classified as being of type A or type a. Assume the proportion of individuals whose gene pair are AA, Aa, aa are respectively p_0, q_0 and $r_0\,(p_0+q_0+r_0=1)$, when two individuals mate, each contribute one of his or her genes chosen at random, to the resultant offspringconsidering these proportions *p*, *q* and *r* they are easily obtained by focusing our attention on an individuals of the next generation and then determining the probabilities for the gene pair of that individual.To being, note that randomly choosing a parent and then randomly choosing one of it gene is equivalent to just randomly choosing a gene from the total gene population. By conditioning on the gene pair of the parent, we see that, a randomly choosen gene will be type *A* with probability

$$P\{A\} = p\{A/AA\}\,p_0 + P\{A/aa\}\,q_0 + P\{A/aa\}\,r_0$$

$$p_0 + r_0/2$$

Similarly, it will be of type a with probability

$$P\{a\} = q_0 + r_0/2$$

Thus, under random mating a randomly choosen number of the next generating will be type *AA* with probability *p*, when

$$P = p\{A\}\,p\{A\} = \{p_0 + r_0/2\}^2$$

Similarly, the random choosen member will be type aa with probability

$$q = p\{a\}\, p\{a\} = \{q_0 + r_0/2\}^2$$

and will be type *Aa* with probability

$$r = 2p\{A\}\, p\{a\} = 2\{p_0 + r_0/2\}\{q_0 + r_0/2\}$$

Since, each member of the next generation will be independent of each of the three gene types with probabilities *p*, *q*, *r* it follows that the percentage of members of the next generation that are of type AA, Aa or aa respectively *p*, *q* and *r*.

$$\begin{aligned} p + r/2 &= \{p_0 + r/2\}^2 + \{p + r_0/2\}\{q_0 + r_0/2\} \\ &= \{p_0 + r_0/2\}\{p_0 + r_0/2 + q_0 + r_0/2\} \\ &= p_0 + r_0/2 (\text{since } p_0 + q_0 + r_0 = 1) \\ &= P\{A\} \end{aligned}$$

The fraction of the gene pools that are *A* and *a* are the same as in the initial generation. From this it follows that, under random mating, in successive generation, after the initial the percentage of the population having gene pairs *AA*, *aa* and *Aa* will remain fixed at the values p, q and r. Suppose now, that the gene pair population has stabilized in the percentages p, q, r and let us follow the genetic history of a single individual and her decendents (for simplicity, assumes that each individual has exactly one offspring 1, so for a given individual).

Let x_n denote the genetic stage of her decendents in the n^{th} generation, the transition probability Matrix of the Markov chain namely.

	AA	*aa*	*AA*
AA	$p + r/2$	0	$q + r/2$
aa	0	$q + r/2$	$p + r/2$
Aa	$p/2 + r/4$	$q/2 + r/4$	$p/2 + q/2 + r/2$

Is easily verified by condition the stable if the randomly choosen mate, it is quite intituitve, that the Limiting probabilities for this Markov Chain (which is also equals the fraction of the individual descendants that are in each of the three genetic statutes), should just be p, q and r.

Theorem 1:

For an irreudicible ergodic Markov Chain $\lim_{n\to\infty} p^n$ exists, and is independent of *i*, $\pi_j = \lim_{n\to\infty} p_{ij}^n,\ j \geq 0$

Then *j* is the unique nonnegative solution of

$$\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij},\ j \geq 0$$

Reduced to

$$p = p(p + r/2) + r(p/2 + r/4) = (p + r/2)^2$$

$$q = q(q + r/2) + r(q/2 + r/4) = (q + r/2)^2$$

$p + q + r = 1$ result is established.

1.3 Deviation from Hardy-Weinberg Equilibrium

The changes in gene frequency can be produced by reducing in (i) population size (ii) selection (iii) Mutation (iv) Genetic drift (v) Migration in Human genetics, the process of formation of new species. Each population is consists of two or more subpopulation with *n* different gene frequencies.

Numerical Example (2)

While count, the number of AA, Aa and aa in a single locus in the population defined by the A and a allels and we shall assume incomplete dominance, so that, Aa heterozygote may can be distinguished Phenotypically from AA and aa homozygotes, we then count the numbers of AA, Aa and aa individuals in a given generation immediately before and immediately after some selective event, then in two decision of point we can calculate (i) A Survival Rate (ii) Relative Fitness (iii) Selection coefficient

Number of individuals in the population according to the genotype

	AA	*Aa*	*aa*
Before Selection	4100	5000	2200
After Selection	3900	4000	1200

(i) Survival Rate $AA = 3900/4100 = 0.95$

$Aa = 4000/5000 = 0.80$

$aa = 1200/2200 = 0.55$

(ii) Relative Fitness (W) $W_{AA} = 0.95/0.95 = 1.00$

$W_{Aa} = 0.80/0.95 = 0.84$

$W_{aa} = 0.5/0.95 = 0.58$

(iii) Selection coefficient $S_{AA} = 1 - W_{AA} = 0$

$S_{Aa} = 1 - W_{Aa} = 0.16$

$S_{aa} = 1 - W_{aa} = 0.42$

The genotype with the largest Survival rate is defined as the Fittest and is used as the standard for the relative fitness (w) of all other genotypes.

The action of selection in a statistical problem, Haldene, Fisher, Wright, Li (1948) and Lernes (1950), suppose that a dominant gene A has the frequency *q* and its recessive allel a the frequency $(1-q)$, in the gene pool of a sexually random breeding population. According Hardy Weinberg Rule (Binomial Square Rule), the population consists of three genotype with frequencies $q^2 AA + 2q(1-q)^2\, Aa + (1-q)^2\, aa = 1$. Let the adaptive values (w) of the dominants AA and Aa, be equal to unity 1 and that of the recessive to $(1-s)$.

In other words, for every unit of offspring produced by the dominant, the recessive produces $(1-s)$ offspring of the average. The value is called Selection co-efficient, the frequencies of the three genotypes before and after selection will be

Genotypes	*AA*	*Aa*	*aa*	Total population
Adaptive value (W)	1	*s*	$(1-s)$	$\overline{W}$

Initial frequenc y	q^2	$2q(1-q)$	$(1-q)^2$	1
Frequenc y after selection	q^2	$2q(1-q)$	$(1-s)(1-q)^2$	$1-s(1-q)^2$

The frequency *q* of the gene *A* in the next generation will be

$$q_1 = \left[q^2 + q(1-q)\right] / \left[1 - s(1-q)2\right] = q\left[1 - s(1-q)^2\right]$$

The increment, of the frequency of the gene *A* is one generation will be

$$\Delta q = sq(1-q) / \left[1 - s(1-q)^2\right]$$

Numerical Example (3)

Adaptive value (w)	0	0.4	0.9	0.99	1.5
Selection coefficient (S)	1.0	0.6	0.1	0.01	-0.5
Frequency after one Generation of Selection (q)	0.67	0.58	0.5128	0.5012	0.441
Increment of gene Frequency (Δq)	0.17	0.08	0.0128	0.0012	-0.056

1.4 Effect of Selection

The genes *A* and *a* be equally frequent, in the original population, so that $q = (1-q) = 0.5$ let the adaptive value of the dominants (AA and Aa) be unity, and suppose that, recessive allels aa have the adaptive values of 0 (a recessive lethal) or of 0.419 semilethal) or 0.9 or 0.99 (subvitals) or 1.5 (Supervital). The frequencies *q* of the gene a in the next generation, and the increments of the gene frequency, will be then for small selection coefficient(s) an approximate formula for the number of generation (n) necessary to change the frequency of a deleterious recessive gene from q_0 to q_n is as follows:

$$ns = q_0 q_n / q_0 q_n + \log_e \left(q_0 / 1 - q_0, 1 - q_n / q_n\right)$$

Under the Hardy Weinberg equilibrium assumptions, states a population having genotypic frequencies $p(\text{of } A_1A_1), 2q(\text{of } A_1A_2)$ and $R(\text{of } A_2A_2)$ achieves after one generation of random mating, stable genotypic frequencies $p^2, 2pq, q^2$ where $p = p + q$ and $q = q + r$. If the initial frequencies $p, 2q, r$ are already of the form $p, 2pq, q^2$ then these frequencies are stable for all generation of the assumption doesn't hold, the law itself may not hold. Suppose as a continuous time analogue to the above that in a smallest time dt a fraction dt of the population dies and is replaced, by random sampling from the population at large, under this system, the frequency *p* of A_1 doesn't change with time, but if $p(t)$ is the frequency A_1 a at time *t*, then

$$P(t + dt) = p(t)(1 - dt) + p^2 dt$$

Passing to the limit in this equation

$$dp(t)/dt = -p(t) + p^2$$

So that, $p(t) = \left\{p(0) - p^2\right\} \exp(-t) + p^2$

Clearly, a population initially in H.W. equilibrium will remain in equilibrium, but for non equilibrium population, the equilibrium state is approached asymptotically, when

Hardy Weinberg proportion strictly apply, later on, when considering, finite population, the population will be counted at the age of sexual maturing. The differential reproductive rates may be due to several causes, including the particular different survival rates and different offspring distributions, the quantities w_{11}, w_{12} and w_{22} will be called 'fitness' of the threeGenotypes, when these are operating and genotypic frequencies will usually change from one generation to the next, with the fitness given above, the frequency the various genotypes in the following generation now satisfies the equation

$$P^1 : 2Q^1 : R^1 = \left(W_{11}P + W_{12}Q\right)^2 ; 2\left(W_{11}P + W_{12}Q\right)\left(W_{12}Q + W_{22}R\right); \left(W_{12}Q + W_{22}R\right)^2 = \left(P^1\right)^2 ; 2p^1q^1; \left(q^1\right)^2$$

Where $p^1 = \left(W_{11}P + W_{12}Q\right)\left(W_{11}P + 2W_{12} + W_{22}R\right)$

Clearly one generation, of random mating, H.W, proportion are achieved. In the same argument shows that

$$P_{11}, 2q_{11}R_{11} = \left(P^{11}\right)^2 ; 2P_{11}Q_{11}; q_{11}$$

$$= \left\{W_{11}\left(p^1\right)2 + W_{12}\left\{P1q^1\right\}^2 ; 2\left\{W_{11}\left(p^1\right)2 + w_{12}p^1q^1 + w_{22}\left(q_1^2\right)\right\}; \left\{w_{12}p^1q^1 + w_{22}\left(q^1\right)^2\right\}\right\}^2$$

It follows that the equation $p^{11} = p^1$ no longer holds good in general constant but for non equilibrium populations the equilibrium state is approached asymptotically.

II. MATERIALS AND METHODS

2.1 Branching Process: Population genetics

Branching process is evolution of aggregate of systems where components can reproduce, be transformed and die the transition being governed by stochastic law the word 'Branching process' coined by Kolmogorov and Diminintiev (1947) to describe the stochastic process of Population Genetics. William Watson (1873) formulated the problem of Extinction families Fisher (1930) studied the survival of the progeny of mutant gene, Haldane (1927) applied in Population genetics.

Theorem 2:

Bienayme-Galton-Watson process is a Markov Chain $\{Z_n\}, n = 0,1,2,...$ the non-negative integers

$$Z_{n+1} = \sum_{r=1}^{Z_n} K = k_1 + k_2 + ... + k_{z_n}$$

And $Kr;\ r \geq 1$ are identically independent random variables with probability distribution

$$P\{k_j = k\} = p,\ k = 0,1,2,...$$

$$\sum P_i = 1$$

We interpret Z_n as the number of objects in the n^{th}generation of a population of family unless the contrary is states we always assume that $z_0 = 1$.

The probability generating function $h(s)$ of *Z*.

$$H\left(S^{2_j}\right)=E\left(S^{2_j}\right)=\sum_{k=0}^{\infty}p_kS^k,|s|\le 1$$

will be referred to offspring generating function.

Defines its iterates, $h_{n+1}(s)=h\left(h_n(s)\right)$

$$h_1(s)=h(s)$$

$$h(0)=s$$

Banayne Galton and Watson (1874) proved that the generating function of Z is the n^{th} iterate $h_n(s)$ of $h(s)$, in order to avoid trivariate, we assume throughout $p_j\neq 1,$ for any $j, p_0, p_1^n<1,$ so that $h(s)$ will be strictly convex on the unit interval.

The process Z is called subcritical, critical, supercritical according as

$$M_1=E(R_1)=h^1(1)$$

Is less than, equal to or greater than 1,

$$E\left(Z_{n+1}/Z_n\right)=E\left(\sum_{r=1}^{Z_n}kr/Z\right)$$

$$=ZnE(K)$$

$$=Zm_1 \text{ and}$$

Hence $E(z_n)=m^n$

Simple Calculation yield

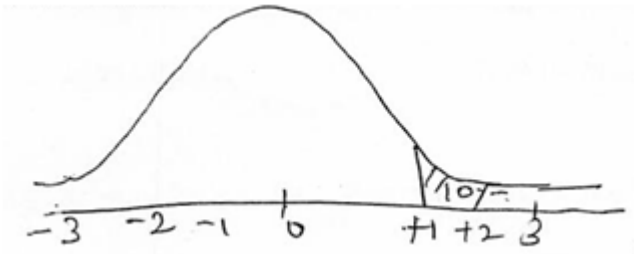
$$V(z_n)=m^{n-1}\left(m_j^n-1\right)/(m-1)\sigma^2 \quad m_1\neq 1$$

$$m=1$$

Where $\sigma^2=v(z)$

2.2 Incidence (Survival); Threshold Characters

Quasi continuous variations (1952), the phenotype values are discontinuous but the mode of inheritance is like that of continuously varying characters



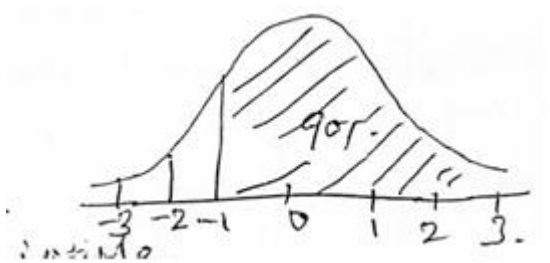


Illustration of a threshold character with two visible classes. The vertical line marks threshold between the two phenotypic classes, one of which is cross-hatched. The population depicted on the left has an incidence of 10% that on the right an incidence of 90%.Incidence (survival) on the visible scale individuals, can have only two values 0 or 1 groups of individuals, however such as families or the population has a whole can have any value, in the form of proportion or percentage of individuals in one or other class. This may be referred the incidence of character. The interpretation of genetic analysis of threshold character is therefore facilitated by the transformation of incidence to values on the underlying scale. The transformation is easily made by normal cases. The standard deviation and population Mean of acharacters with three visible classes may be in the general form the following way.

2.3 Population Genetics: Fokker-Plank Diffusion Equation

Inbreeding, Natural Selection, Genetic recombination the theory of Stochastic Process to genetical and evolutionary theory should therefore have already became apparent, we suppose that there is a population of size Nt, each individual of which carries a pair of genes AA, Aa or aa, the number of A genes in the total population will be denoted by At, and we write further $A_t\left[2N_t\right]=p_t$. We shall assume that the Bivariate process A_t, N_t is a Markov one in time *t*. This can't be strictly true, for the exchange genes will be related in individual genealogical lines to generation times, which we have seen must age-dependent birth and death rates, but over the entire population it is a more realistic assumption, except perhaps for short-term effects if generation times are strictly periodic, Any effect of their genes has of course been ignored.

The detailed evolution of At and Nt will also depend on the Phenotypic combination at any instant, and its breeding characteristics. We can only hope to be able neglect such factor in detail if we review the evolutionary changes in a rather broad manner, rather than of the simple Additive model. We have seen that the behavior of a finite population can be fundamentally different according to whether its total size is expected to increase or decrease, and we may expect this to apply also to a population genetic constitution. In many situation it is reasonable to suppose that the total size N_t is fairly restricted by the environment, and keeps approximately constant, but as Felier (1951) has stressed, this is a drastic assumption, which severely limits the class of process studied. It would be useful, as in the process of the last selection, to study the joint distribution of A_t and N_t under conditions, under general Mutation and Selection changes this is, however not easy and we consider here the solution under the selection, but usual assumptions of N constant, following a treatment of due to Malecot (1952).

As $N_t=N$ is constant the Markov process is one in A_t only or equivalently in $p_t=\frac{1}{2}\frac{A_t}{N}$ equation for $M_t(Q)=E\left(e^{\theta t}\right)$ is

$$\frac{\partial M_t(\theta)}{\partial t}=\psi\left(\theta,\frac{\partial}{d\theta}\right)M_t(\theta) \quad \ldots(1)$$

where $\psi(\theta, P_t) = \underset{\Delta t \to 0}{\text{lt}} \log E \dfrac{\left\{ \dfrac{e^{\theta \Delta P_t}}{P_t} \right\} / t^2}{\Delta_t}$

We now make the following simplifying assumptions.

(1) For small t_t, there is a chance KNPtQtΔ_t where $Q_t = 1 - P_t$, of a single gene transition $a \to A$.Due to the random shufulling of generation the offspring of any mating (Strictly speaking, these changes depend on the phenotypic male and female frequencies of the AA, Aa and aa, gene pairs, even under random mutation rate u from A to a i.e the chances of such a mutation Δt is $2NP_t$ similarly the rate from a to A is v.

(2) Selection is assumed to operate on the ratio Pt/Qt of A to genes, changing its value deterministically (again for Nt (constant, it seems difficult to formulate this more realistically by amount of $\sigma\Delta_t$ Pt/Qt in t; this implies in Pt of

$$\sigma P_t \sigma_t P_t Q_t \Delta_t + \left[V Q_t - \mu P_t \right] \Delta_t$$

Assumption (ii) and (iii) gives a change in Means of Pt of change in mean but a variance $\frac{1}{2} K P_t Q_t \frac{\Delta t}{N}$ the variance due to (ii) is

The third-Cumulants contribution from (i) and (iii) is $0(1/N^2)$

Hence $Q(Q, Pt) = \theta \left[\sigma P_t Q_t + U \phi_t - \mu P_t \right]$

$$+ \frac{1}{2} \theta^2 \left[\frac{1}{2} K P_t \theta_t + \frac{1}{2} U \phi_t + \mu P_t \right] / \left[N + 0\left(\theta^3 / v^2 \right) \right]$$

….(2)

The equation (1) then becomes a Fokker plank diffusion equation

The equation (1) thus becomes a diffusion equation, we may neglect the terms of $0\left(1/N^2\right) = 0$ that, Q only appears explicitly in as a Quadratic equation (1) is more, a partial differential equation of the second order in d/Dq. If we assume that the sol. Of (1) has a limit distribution, their must satisfy the equation

$$\Psi\left(\theta, \frac{\partial}{\partial \theta} \right) \frac{M}{\theta} = 0 \ldots (3)$$

For example, if the mutation rates *v*, and U are zero, we have the caution

$$\left[\sigma\theta + \frac{1}{K} \frac{\theta^2}{N} \right] \left\{ \frac{\partial}{\partial \theta} - \frac{\partial^2}{\partial \theta^2} \right\}_M \ldots (4)$$

This has the solution Aθ+B, moreover, from the full equation containing $\partial M_t / \partial t$, we see that $\partial M_t / \partial t = 0$, when $Q = 0$ or $4N\sigma / t$.

Hence for these values of Q, the limiting $M(Q)$ is equal to its initial value are, so that

A+B=1, Ae-4N$^{\sigma}$/k+B=e-4N$^{\sigma}$/k

where M=1-e4N σ_r /k/1-e-4N σ /k(e$^{\theta}$-1)+1 … (5)

Giving the complementary chances of extinction of the A or a genes in the absence of mutation, Malecot has shown further that this distribution is infact the limiting solution of (1) and (2) for u and V zero.

As a second fairly tractable care, we suppose $\sigma = 0$ and u, V are $0(1/N)$, then expanding M_t in powers of u, We find the equation for $Ms = E(p_t)$.

$$\partial ms / dt = -sv\left(m_s - m_{s-1} \right) u_f m_f - ks(s-1) / kN\left(ms - ms - 1 \right) + 0\left(\frac{1}{N^2} \right)$$

And in particular, as $m_0 = 1$

$$\partial m_1 / dt = -v\left(m_1 - 1 \right) - u m_{11}$$

When

$$m_1 = v\left[1 - e - (u+v)t / U + V + re - (U+v) \right] t - v / (U+v) \ldots (6)$$

The equation (6) may be solved successively for m_2, m_3 and it is evident that m_s has a limit as $t \to \infty$ for allels. This must therefore be given by the recurrence relation.

$$m_s \left[s - 1 + 4N(u+V) \right] / k + 0\left(1/N \right)$$

$$= m_{s-1} \left[s - 1 + 4Nv / k + 0(1/N) \right] \ldots (7)$$

Which defines the continuous distribution from P (in depending of r)

$$f\left(p / dp \alpha P 4Nv / k - 1(1-P) 4NU / k - 1 dp \right) \ldots (8)$$

This result may be obtained more simply from a general formula due to wright

$$F(P)_{\alpha} 1 / \sigma^2(P) \exp\left(2 / m(p) / 2t(p) dp \right) \ldots (9)$$

Where m(p) is the change in means per unit time, and (p) the change variance, we put

M(p) = v(1-p) and $\sigma^2(p) = 1/2KP(1-P)/N$ and (8) follows. The formula (9) Qa, however obtained merely by requiring the mean and variance of any stationary distribution with probability density to remain constant and the solution (5) shows that it doesn't give the correct solution in all cases.

The effect of selection, A gene which is selectively advantageous against one genetic background is disadvantageous against another, it will be shown later that, such interaction effects, can have major evolutionary consequences, to define, at the time of conception of any generation the frequencies of the genotypes are P, $2Q_t$, R then these contributes gamets to form the individuals in the following generations in the proportion W_{11}P: $2W_{12}$Q:W_{22}R (Note that the population is being considered at the time of form formulation of Zygote the gametes of the previous generation, when selective differences exist this is only time by extinction we mean the event that the random variable (Zn) consists of Zeros for all but a finite number values of n and we define the probability q of extinction $q = \underset{n \to \infty}{\text{lt}} hn(0)$

From the convex nature of the probability generation function we have the following fundamental theorem due Steffiemson (1932)

Theorem (2) The extinction probability q of the Biename Galton Watson process is the smallest non-negative root of the equation $n(s)=s$. It is 1 if $m_1=1$ and <1 if $m_1>1$. Book (1966) has hown hat if we can find an upper bound $g(s)$ for $h(s)$ the q will be bounded by the smallest non-negative root q_1 of $g(s)=s$ and q_1 will be the upper bound for the probability of extinction. Since $h(s)$ is strictly convex on the unit interval as $n\to\infty, h(s)$ increases to q in $0\le s\le q,$ decreases to q in $q\le s\le 1$ and $h_n(s)=s$ if $s=q$ or 1. Also since, $p(Z_{n+1}\ne k, i\ge 1/Z=k)>0$ all the states $k\ge 1$ are transient and hence with probability 1, $Z_n 0$ or as $n\to\infty$ then we have

Theorem (3) Whatever be the value of m_1, the sequence $\{Z_n\}$ converges as $n\to\infty$ to or 0 respectively with probability $1-q$ or q.

Churchil (1967) has proved that $\lim_{n\to\infty} P(Z_n=k)=ak$ exist for all *k*, if the generating function governing production in the n^{th} generation, $n\to\infty$ converges sufficiently to the degenerating function, this result is generated by Kaplan (1973) for Multiprocess, when H.W proportions strictly apply later on, when considering finite population, the population will be counted at the age of sexual maturity, the differential reproduction rates may be due to several causes, including the particular different survival rates and different offspring distributions, the quantities W_{11}, W_{12} and W_{22} will be called the 'fitness' of three genotypes, when these are not equal selective forces are operating and genotype frequencies will usually change from one generation to the next. With the fitness, given above, the frequencies will usually change from one generation to the next.

With the fitness given above, the frequencies of the various genotypes in the following generation now satisfy the equation

2.4 Selection: The response and its prediction, genetic properties of the population by the choice (1) individuals to be used as parents, which constitute selection and the (ii) second by control of the way in which the parents are matted, which embrass inbreeding, cross breeding, the basic effect of selection is to change the array of frequency.Non-recessive A_1 mutant is introduced into a presumly pure A_2A_2 population, then homozygotes A willn't usually appear until the frequency of a is quite high. We can't normally, Multiple alleles, $A_1, A_2, ..., A_k$ each with non-negligible frequency. Suppose that the fitness of A_iA_j individuals is W_{ij} the frequency of A_1 is P_1 and that genotype frequencies are in Hardy-Weinberg form. Thus, the mean fitness of the population may be taken as unity with this convention, if the new mutant is denoted by A_{k+1} we denote by A_jA_{k+1}. We denote the fitness of A A by U. In deriving the survival probability of formation of homozygotes mutants can be ignored and will then describe an A_jA_{k+1} individuals as a Mutant of type. If in generation *t* the number of type i mutants is $N_i(t)$, then clearly

$$E_{nj}(t+1)=P_j\sum_i u_iN_i(t),\quad i=1,2,...,k \qquad \text{.... (10)}$$

In Matrix term, this may be written

$$E_n(t+1)=Un(t) \qquad \text{.... (11)}$$

when $U=PU_1^1=P^1=\{p...p_k\}$

$$u^1=\{u_f...u_k\}$$

to obtain survival probabilities it is necessary to know not only the number of mutants, but the complete joint distribution of the number of A_{k+1} offspring from A_iA_{k+1} parent.

If the joint generating function of this distribution $f_i(Q_i...Q_k)$ then $f\{1,1,...\}=1$ and writing $f_i\{Q,...\theta\}=f_i(q)\ f_1^1(1)=u_i$.Now any mutant offspring from an A_iA_{k+j} parent will be A. A with probability P_j from this it follows, that $f_i(Q_1,...,Q_k)$ assumes the specified form $f_i(Q_1,...,Q_k)-f_r(P_iQ_i+...+P_1(Q_k))$, Harris (1963) Theorem can be applied.

Theorem (4)

For a multiple type Branching process governed by equation (10) is that is for a set of objects of various types, each of which can produce offspring of any type according to Branching process law, the probability of the survival of Mutant A_{k+1} is positive of largest eigenvalue of M exceeds unity. In this case, if *Q* is the probability of extinction mutant given a single initial mutant type *i*, the *Q* are the unique positive selection (<1) of two set of equation.

$$Q_i=f_i(Q_1,...,Q_k)\quad i=1,2,...,k$$

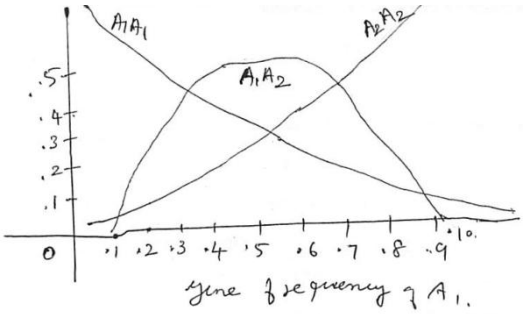
It is useful to consider first the condition that the various Q_i be less than unity. The show that *M* is of rank unity and thus has only one non-zero eigenvalues. This eigenvalue must then, be identical to the sum of all eigenvalues, which trace (*M*) of $\sum P_iU_i$. Thus the necessary and sufficient condition that Survival probabilities be positive $\sum P_iU_i>1$. When holds, the easiest way to solve the system of equation is to multiply the equation p_i and add over *i*.Defining, this leads $\theta=\sum P_i\theta_i$ Often *Q* is easily the quantity of interest, since if we don't know the exact genotype formed by initial mutant, it is reasonable to assume that this genotype is A_iA_{k+1} with probability *Pet*. If any event, individuals θ_i values are best found by solving for θ and calculating each θ_i directly

Selection index $I=Px+WP_y$

Where *I* is the true index by Means of which individual are to be chooses. *W* is a factor by which the Phenotypic value character *Y* is to be Multiplied by Selection differential

$$R_{\sigma p}=S_{\sigma p}/h^2$$

Relationship between genotype frequencies and gene frequencies for allels in a population of H.W. equilibrium.



CONCLUSION

Finally linear Models of Structured population $P = fP - dp = \{f - d\}P$ which means given a current population *P* and the fecundity and death rates *f* and *d*, we can predict the changes in the population. $P_t = P(t)$ the size of the population measured on day *t* is $P = P - P$ is the differences or change in population between two consecutive days, now we ultimately care about is understanding populationp, not just, Bp, but $p = p_t + \Delta p = p_t + (f - d)p_t = (1 + f - d)pt,$ some constant together by letting $= (1 + f - d)^{t+1} \delta t$ our model of population growth has below simply

$$p_{t+1} = \lambda p$$

population ecologist often refer to the constant of the finite growth rate of the population, from this model we can now predict population an any future times. $P_{t+1} = (1 + f - d)pt,$ a difference equation, when the difference doesn't appear and $\Delta p = (f - d)dp$ notice that in this model, we ignore the males, the female population is the important one track to understand the long term growth or decline of the population.The interpretation of genetic analysis of threshold character is therefore facilitated, underlying scale. The transformation is easily made by reference to a table of probabilities of the normal curve.

The threshold is a point of truncation whose deviation from the population mean can be found from the proportion of the population falling beyond it.

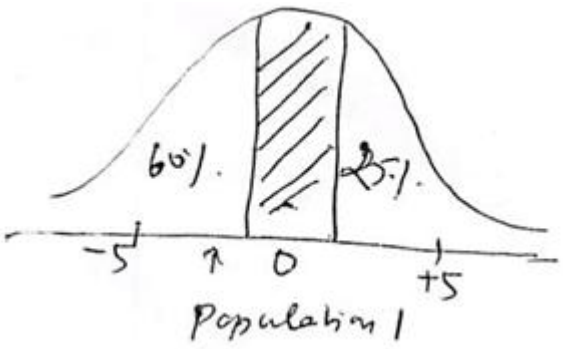
Fisher and Yates (1943) is convenient to use because it refers to a single tail of the distribution and deviates confusion over the sign of the deviation. The transformation from the visible to the underlying scale enables us to state the mean phenotypic value of a population or family in terms of its standard deviation, and to compare the means of different populations or families provided they have the same standard deviation. It is convenient to take the position of the threshold as the origin or zero point, on the underlying scale and to express the Mean as a deviation from that threshold. Thus if the incidence of the character is, for example 10 percent a table of the normal curve shows, that the threshold exceeds the mean by 1.28, the population Mean refer to the threshold as origin is therefore -1.28 or if the incidence curve 90 percent than the population Mean would be 1.28 for any comparison of means, however it is necessary to assume that the population compared have the same variance on the underlying scale. If reasons are known for the variances not being equal in comparisons for example between inbreeds, F1's and F2's, then the means can't be expressed on a common scale that allows a valid comparison to be made. This is as we can go with a character that is simply expressed only two classes. The mean of a population or group can be stable, but not the variance because the Mean has to be started in terms of the standard deviation. We can however, subjects the observed means of families to analyse and compute the heritability of the character. The heritability in threshold character is treated by A. Robertson and Lerner (1949) and Dempster and Lerner (1950). If the character has three classes in its visible scale then comparison can be made the variances of populations as well as between the Means simple class.

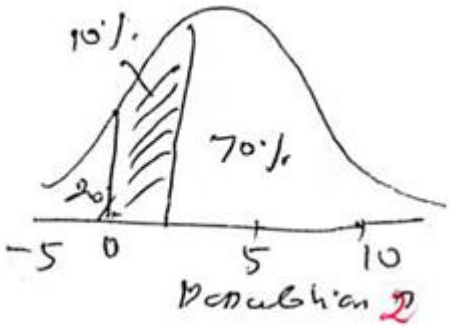
For example threshold character with three visible classes, in two populations with incidence as shown the axes are marked in threshold units and the population means are indicated by arrow.

	Class		
	X	I	Z
Incidence % Population (1)	60	15	25
Population (2)	20	10	70

The deviation of the thresholds from the population means from a table of normal curve are as follows

	X/I Intervals	I/Z	Threshold
Population (1)	–0.25	0.67	0.42
Population (2)	–0.84	–0.52	0.32





The intervals between the two thresholds, given above on the right are found by subtraction of the deviations of the two thresholds in each population. These threshold intervals are supposed by hypothesis to be equal on the underlying scale. By assigning the threshold intervals the value of one 'threshold unit'. We can therefore express the standard deviation of the two population on a common

Basis in terms of threshold units the std. deviation then becomes

$$\sigma_1 = 2.38 \text{ threshold units}$$

$$\sigma_2 = 3.12 \text{ threshold units}$$

The mean of the population can also be expressed in threshold units, denoted from X/I threshold on origin, they are $M = -0.25 = -0.60$ threshold units

$$M_1 = -0.84 = 2.62 \text{ threshold units}$$

The Standard deviation and population Mean of a Character with three visible classes may be put in general form in the following way, Let 'X' be the incidence in one visible class and Y the incidence in this class with the intermediate class. Let the threshold between these two classes be the origin of the underlying scale. Let *x* and *y* be deviations of the two threshold corresponding to the incidence *x* and *y* respectively, then the standard deviation is $= 1/x - y$ threshold units and the Mean is $M = -\sigma$.

$$\sigma = -x/x - y \text{ threshold units}$$

The comparison of variances in this way depends entries as we have pointed out, on the assumption that, the interval between the two threshold is constant from one population to antoher. If we think again of the hypothetical substances (or) processes, whose concentrate or rate determines the values in the underlying scale, the assumption is that the intermediate class express the same difference of concentration or of rate in the two population compared.

Acknowledgement

Dr.S.Thobias Ph.D., Associate Professor& Head, Department of Statistics, Loyola College (Aut.), Madras, Dr.(Mrs).Nancy, Ph.D., Asst. Professor in Statistics, Presidency College, Chennai. Dr.(Mrs).M.Selvam, Ph.D., Asst. Prof. in Statistics, Govt. Arts College, Salem, Dr.K.Harikumar, Ph.D (Mathematics)., Principal, Prof.(Miss).E.Radhika, M.Sc., M.Phil., B.Ed., HOD& Faculties in PG Dpartment of Mathematics, King Nandhivarman Arts & Science College, Mr.N.Selvaraj, B.Sc., B.Ed., (Mathematics), Dr.Felix, Associate Professor in Community Medicine & Bio Statistics, Department, Rajah Muthiah (Govt.) Medical College, Chidambaram for their valuable suggestions to prepare this paper.

References

[1] M.Rose, Probability Models, (9/e), Academic Press

[2] Elizabeth S. Allman-John A. Rhode, Mathematical Model in Biology, Cambridge Press

[3] Henry C.Tuckwell, Elementary Application of probability theory (2/e), Chapmen & Hall (1995)

[4] N.John Britto (2018) Introductory Statistics, ISBN 978-93-87756, www.mayaspublications.com

[5] John Felix (2007) unpublished Ph.D. thesis, Survival Data Analysis, Annamalai University

[6] Mrs.S.Thilagam (2003) unpublished M.Phil., thesis, Frailty Models Madras University

[7] N.John Britto (2010), Unpublished M.Phil., Thesis, Application of Stochastic process in Evolutionary Genetics, Manonmaniam Sundaranar University

[8] N.John Britto, Dr.S.Muthalagi 'Introduction to Stochastic Models in Human Genetics http://ijar.com, e-ISSN 2348-1264, print ISSN 2349-5138, (i.f 3.25) UGC Serial No. 43602, special issue, PP:167-170, Proceedings of the National Conference on Recent Advances and Innovations in Pure and Applied Mathematics held on 22nd&23rd Feb 2019 organized by Department of Mathematics, Sri Akilandeswari Women College, Wandhiwash, Thiruvannamalai Dt, Tamil Nadu.

[9] N.John Britto, Dr.V.Rajagopalan, Dr.S.Sasikala, 'Stochastic Process in Evoluationary Genetics; Fokker Plank Diffusion Model" ISSN 2349-5162, www.ietir.org, April, 2019, Vol.6, Issue 4, PP 631-639, Proceedings in the international conference held on 15th and 16th, Feb 2019, organized by Department of Mathematics, St.Joseph's Arts and Science College (Aut.), Cuddalore

[10] N.John Britto 'Race, Genes and Intelligence: Bayesian Methodology' ISSN: 2347-1557, Vol. 04, Issue 4 (2016), PP:123-128 (Speical Issue), International Journal of Mathematics and its Application: http://ijmaa.in Proceedings in the National Conference on Recent trends in Applied Mathematics (NCRTAM'16) held on 22nd& 23rd July 2016, organized by PG and Research Department of Mathematics, St.Joseph's Arts and Science College (Autonomous), Cuddalore.

[11] M.S.Bartlett, F.R.S An introduction to Stochastic Process with Special reference to methods and application, Cambridge (1966) University Press

[12] D.J.Finney (1978):Statistical Methods in Biological Assays

[13] A.P.Corse and S.A.Paranji (2000): A course in Mathematical & Statistical Ecology, Kluwer

[14] R.C.Elandt Johnson (1975): Probability Models and Statistical Methods in Genetics: Wiley Publication

[15] C.C.Ti (1976): First Course in Population genetics, Boxwood Press

[16] Falconor, D.S. (1977) Introduction to Quantitative genetics

[17] Kempthrone Oscar (1957) – An Introduction to genetic statistics, John Wiley & Sons, New York

[18] Rao C.R (1952) Advanced Statistical Methods in Bioimetric Research, John Villy & Sons, New York

[19] Selected Statistical Techniques (2006), Dr.M.Rajagopalan, Vikas Publication

[20] Karlin S and Tylor H.M. (1975), A First Course in Stochastic Process Vol I, Academic Press

[21] J.Medhi, Stochastic Process (2nd Edition), New Age International, 1992

[22] G.Sankaranarayan: Branching process and its estimation theory.

[23] An introduction to probability theory and its Application, Vol II 2/C, Wiley, Publication

[24] Applied Stochastic Process: A biostatistics and Population oriented approach

[25] Population genetics with Fluctuating population rises, (2017)

[26] Probability and Random process (2/e) Scott Miller Donald Childrens with Application to Signla Processing and Communication, Academic Press

[27] Probability and Stochastic Process (2005) 104, D.Yates, David J.Goodman, Wiley Publication.