International Journal of Mathematics And its Applications
Volume 4, Issue 4 (2016), 123–128. (Special Issue)
ISSN: 2347-1557
Available Online: http://ijmaa.in/

International Journal of Mathematics And its Applications

Race, Genes and Intelligence: Bayesian Methodology

Research Article*

N.John Britto[1]

1 Assistant Professor in Statistics, Joseph Arts & Science college, Villupuram, Tamilnadu, India.

Abstract: In this review paper, introduction to Race, genes and intelligence is related (correlated) influencing factor understanding, using Bayesian Methodology Hidden Markow Chain (HMC) is discussed.

Keywords: Elements of classical genetics; Definition of Race, genes and intelligence; Hardy Weinberg law; Birth Death Law (Passion Law) Survivorship model (data) Analysis Binomial Pooled Sampling.

1. Introduction

Over the pas two centuries Biomedical science has difference in achievement reflects innote differences in ability among racial groups, Broadly speaking, the view that difference in academic achievement, IQ scores (Anonymous, intelligence and genetic Deter minimum, gene Watch 19; 9-12(2006) employment status(or) wealth primarily reflect innote differences is called "biological determination" As the late Stephen J. Yould painted out, as its case, biological determinants is a "theory of limits" what is Race? Order is Heaven's first law; and this confessed, some are, and must be, greater than rest, without just gradation, could they be subjected, these to those or all to these? Alexander Pope, Essay on Man (1733)

1.1. Need of Mathematical Models in Population Genetics

A Brief History of the Role of Selection: Neutral theory

Polymorphic: This preservation of favorable variation and the rejection of injurious variation. I call natural selection, variation neither useful nor injurious wouldn't be affected by neutral selection, and would be left a fluctuating elements as perhaps we see the species called 'Polymorphic (Dawin, 1985) **Charles Darwin** was the first to formulate the concept of neutral selection and to apply it to evolution and adoption hereditary information of most organisms is encoded in deoxyribonucleic acide (DNA) and that variation is caused by Mutations; **the definition of natural selection,** the differential reproductive success of different genotypes.

Gregor Mendel (1865) carried out experiments on peas, and discovered the basic rules of inheritance, the cause of variation namely Mutation was described by Hugo de vries concluded, mutations caused Drastic, non gradual changes. The relative importance of mutuations Vs selection. Evolution proceeds by small leaps (variation) by CityplaceDarwin when, evolution precedes by large leaps caused by Mutions.

* *Proceedings : National Conference on* `Recent Trends in Applied Mathematics` *held on 22 & 23.07.2016, organized by Department of Mathematics, St. Joseph's College of Arts & Science, Manjakuppam, Cuddalore (Tamil Nadu), India.*

Neo-Darwinism (Synthetic theory of Evolution): Until early 1930'n Fisher (1930) Haldane (1932) and wright (1931) in 1960's protein sequences became allelic Kimure (1968a) and kinf and jukes (1969) Kimarel (1968b, 1977, 1979, 1983 Kiumure and ohte (1973) discussing in the following is (i) the genetic diversity is largely caused by random genetic dritt, implying that the genetic diversity seen in population is a transient (stochastic) phenomena. ii) Mutations or introduced at random and they either go to fixation or are lost soley due to stochastic forces.

Race and intelligence: Concept of race means, i.e race is correlated with income or IQ influence cognitive ability, however, single intelligence, Multiple intelligence, scholars have developed, multiple types of cognitive functions that are valuable and measurable. As pyschometricians , argue that, intelligence tests measure reasoning skills although the tests also measure.

Hereditarian: claims, are based as the alleged heritability of IQ Heritability assess the way a trait varies in population, and purports to measure how much of that variation explained by genetic differences within the population. The remaining variation is attributed to all other factors i.e the environment and non-genetic aspects of biology.

If Children in a classroom score between 90 and 130 on an IQ test a hereditarian might claim that 65% of the point difference in IQ in due to genetic differences between the students and 35% is due to other factor. Strong proponents of hereditarian theories tend to believe that genetic differences explain as much as percent of the variation in adult IQ in a population, but other scholars believe that genes explain much less than 50% of the variation in IQ.

Social environmental variables: found that a 15 point difference in IQ scoring among high school boys only explained 6% of the variability in their earning at age 35.

Genes Brains and intelligence: Scientist have vague and preliminary ideas about how brain structures correlate with thought processed (including solving problems on intelligence tests) and scientists are only beginning to study the ways in which genes influence the development brain structure. The binary formulation of **"genes Vs environment"** is misleading. Cognitive abilities are comply and will likely to influenced by a myriad of **environmental factor and genes** will play a dominant role in shaping the normal range of human cognitive abilities. It is **statistically implausible** that variants of numerous genes relating to intelligence would be distributed among racial groups in a manner that systematically conferred cognitive advantages of one groups or disadvantage in another.

Furthermore, there is no evidence to support the claim that claims that current racial difference in mean IQ scores are caused by racially distinctive pattern of genetic variation. There is evidence that, IQ scores are influenced by environmental, factor that are pervasively and systematically patterned along racial lines. Nonetheless, mean IQ differences among racial groups have been decreasing. Perhaps in improve educational opportunities, muggiest that difference in IQ scores are the result of social inequality rather than its causes.

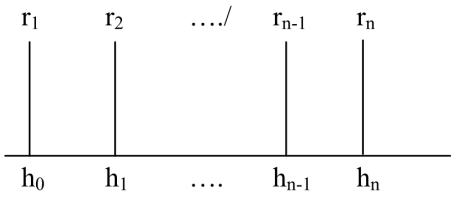
2. Methodology

2.1. Hiddern Markov Model: A General Introduction

A sequence of random variables h_1, $h_2 \ldots$. is said to follow of lth order Markow Chain of $P(h_i/n_1 = n_{i-1}) = P(h_i/h_{i-1} \ldots .h_{i-1})$ one may assume that, an observed sequence of i.i.d r.v's form a first order Markow chain with transition probabilities, $p(h_{i+1}/h_1) = \phi_i, h_i + 1$ with an observed realization of the claim, we can obtain the MLE of the ϕ by counting the frequencies of dimmer occurrences.

The basic Model of HMM, can be written as, $r_i \to f_i(r/h_i, \phi)$; $h_i \to g_i(h/h_{i-1}, \tau)$, Where f_i and g_i are probability distributions ϕ and τ are parameter and r_i are observations, The h_i form a Markow chain and are often unobservable (i.e., hidden) what is of interest is the influence of ϕ, τ and perhaps the h_i.

HMM, proposed by Churchil (1989) as follows



A graphical illustration of the hiddern Markow model.

In this model, we assume that the hidden layer $h = (h_o, h_i \ldots h_n) n_i = 0$ is a Markow chain. Each h_i takes the as only two possible values. $h_i = o$ implies that residue $r_i \rightarrow$ Multinomial (ϕ_0); and $h_i = 1$ indicates that $r_i \rightarrow$ Mullinom (ϕ_1), here $\phi k = (\phi_{ka}, \phi_{kl}, \phi_{kg}, \phi_{kt})$. $A_{2\times 2}$ transition matrix, $\tau = (\tau, k_l)$, where $\tau k_l = (p(h_i = K \rightarrow h_{i+1} = 1)$ dictates the generation of n. A similar model has been developed by Knog et .al (1994/b). Let $(\phi) = (\phi$o, ϕ_1, $\tau)$ The likelihood function of $\oplus$ can be written as L($\oplus$/R) = Σ_nP(R/n, ϕo, ϕ_1).P(n/τ) = Σ $p_o(n_o).\Pi_{i=1}{}^{n}$ $(\phi_{ni,ri}, \tau_{hi,hi-1})$ where h_o is assumed to follow a known distribution P_o(ho) This function can be evaluated using a recursive summation. $F_{k+1}(n) = \Sigma_{ni=0}.\{F_k(h_i), \tau_{ni}\ \phi_n \phi_{i+1}\}$ for i= 1...n

Hidden Markov chain: Let (x_n, n=1, 2) be a Markov chain with transition prabilities p_i,j a nd initial state prabalities $p_1 = \{px1 = i\}$ i$\geq$o, Suppose that there are finite set of I of signals and that a signed form τ is estimated each time, the Markov chain esteems a state. Further, suppose that when the Markov chain esteems state of then, independently of previous Markov chains, states and signals the signale emitted is s with probability p(s/j) Σ p(S/j)=1 that is if τ represents with signal emitted then $p\{S1 = s/x1 = j\} = p(s/j)$; $p\{S_n = s/X_1, S_1 \ldots Xn-1, Sn-1, Xn = j\} = P\{x/j\}$. A model of the preceeding type in which the sequence of signal S_1, $S_2 \ldots$ is observed while the sequence of underlying Markov chain states X_1, $X_2, \ldots, X_n$ is unobserved, is called a hidden Markov chain Model.

2.2. Bayesian Methodology

In Bayesian analysis, a joint probability distribution f(y, ϕ, τ) is employed to describe relationship among all variables under consideration those that we observe (data and knowledge?) those about which we wish to leaves (Scientific hypothesis, ϕ) and those that are needed in order to construct the model (missing data or nuissance parameters, τ) The basic probability theory that leads us to an efficient use of the available information and to precise quantification of uncertainities in estimation and prediction. The Bayesian approach has the following advantages.

i) its explicit use of probability models to formulate scientific problem.

ii) its co-herent way of incorportating all sources of informat and treating nuisance parameter and missing data.

iii) its ability to quantity uncertainties in all estimates.

Procedure: **Bayesian analysis** treats parameters ϕ, and τ as realized values of random variables that follow a "prior distribution" $f_\phi(\phi, \tau)$ typically refarded as known to the researcher independently of the data under analysis. The joint probability distribution can then be represented as Joint = likelihood X prior that is P(y, ϕ, τ) = f(y / ϕ, τ). $f_\phi(\phi, \tau)$. The theorem that combines the prior and the data, to form the conditional distribution P(ϕ, τ/y) also called the posterior distribution of ϕ

$$P(\phi, \tau/y) = [P(y, \phi, \tau)/P(y)].[f(y/\phi, \tau).f_\phi(\phi, \tau)/f(y/\phi, \tau).f_\phi(\phi, \tau)d\phi dt]$$

Where (y) Marginal likelihood, to obtain posterior distribution P(ϕ/y) = P(ϕ, τ/y)dt. Which give the point estimate of ϕ estimating Bayesian statistics i) development of a model f(y/ϕ, τ) $f_o(\phi, \tau)$ which leeds for compaling posterior distribution.

3. Analysis

Human Genetic Diversity Population genetics influence is for diploids, scaled mutation rate $\theta = 4$ Neu. Where Ne is effective population size of a stationery (or demographically stable) population n is locus Mutation rate wright Fisher Model (population). The **Wright-Fisher Model** is the basic Model for reproduction, in a finite population that can abilize several mutation models and selection schemes. Which is at the heart of may models that **describe how gene frequencies evolve in the presence of random drift, mutation and selection**.

3.1. Randon genetic drift: The Wright-Fisher Model

$P(y(n+1)) = j/y(n) = i) = (2N)p_j(1-p)2N-j)$ where $P = i/2n$, which is an example of discrete time Markov chain model. Its Analogy of continuous Markov chain is called Fokker Plamk diffusion equation (simulation of Genetic Drift). In a finite population, the random sampling of gametes alone causes changes gene frequencies. This process is known as random genetic drift. To investigate the consequences of random genetic duft, consider single locus with two alleles A_1 and A_2. Assume a randomly mating diploid population of size N (haploid population) with non overlapping generations. Each generation 2N game t is are sampled at random from the parent generation. If y(n) denotes the number of gametes of type A_1 at generation n then in the absence of mutation and selection the number of A_1 allels at time n+1is given by the binomial distribution, i.e the probability that there are j gamts of type A_1 at generation n+1 given that there are igametes of type A_1 at generation n is. Ewens (2004) model, Mutent alleles that have little effect as the phenotype of the organism may remain in the population, until they either become fixed a lost due to stochastic forces. Other alleles are maintained in or quickly eliminated from a population by selective forces. Mathemabal Models that are based on the Laws of inheirtance can illuminate the role and relative importance of stochastic and selective forces. The simplest Mathematical models track allele frequencies, in a randoming moting (panmix population) monocious population changes in allele frequencies are caused by mutation, random genetic drift and selection.

3.2. The Hardy-Weinberg Law and a Markow Chain in Genetics

Consider a large population of individuals, each of whom possess a particular pair of genes, of which each individual gone is classified as being of type A type a. Assume that the preparation of individuals whose gene parts are AA, aa, Ac are respectively p_o, q_o and r_o (po+lo+r_o=1). When two individuals mate, each contributes one of his or her genes, chosen at random to the resultant offspring. Assuming that the mating occurs at random, in that eacher individual is equally likely to mate with any other individual, we are interested in determining the proportion of individuals in the next generation whose genes are AA, aa or Aa calling these proportions P, O and r they are easily obtained by focusing attention am an individual of the next generation and them determining the probabilities for the gene pari of that individual.

for type 'A' $p\{A\} = p_o\{A/AA\}P_o + p_{o\{}A/aa\}(o + P\{A/Aa)ro = Po + ro/2i.eP = P\{A\}P(A) = (Po + ro/2)2$

type 'a' $P\{a\} = Q_o + r_o/2; p\{a)P\{a) = (Q_o + r_o/2)2$

Type Aa $r = 2P\{A\}p\{a\} = 2(p_o + r_o/2)(p_o o + r_o/2)$.

The fraction of the gene pod that are A and a are the same in the initial generation. From then it follows that, under random moting, in all successive generation after the initial are the percentage of the population having gene pa AA, aa and Aa will remain fixed at the values P, q and r.

3.3. Departure from Hardy-weinberg Equilibrium

HWE can be disrupted by population sub division or inbreeding and specific attempted to detect discrepancy between observed and expected geneotype frequencies.

Generation time: Discrete Generation: Although generation interval is commonly thought of the average age of the parents when therein offspring are born. Demographers use formula that relate generation time to the age of reproducing females, the reproductive level of each age group and the probability of each age group.

Non randon mating : inbreeding : How does affect allele frequencies ? P_{n+1}be the frequency of the allele frequencies is calculated as the frequencies of homozygotes for one allel plus half the frequency of the heterozygote for all allele plus (After one generation of inbreeding)

$$\begin{aligned} p_n + 1 &= P^2(1-F) + PF + 1/2(2pq(1-F) \\ &= P^2(1-F) + PF + 1/2Pq(1-F) \\ &= P^2 + P4 + F(P - P^2 - Pq) \\ &= P(P+q) + P(F(1-P-q) \\ &= P(1) + PlF(0) \\ &= \text{P The individuals the bottom of the pedigree is inbread with the one genotype.} \end{aligned}$$

Pedigree Analysis Path diagram construction: The inbreeding co-efficient, F of an individual (the probability of autorgosity) can be determined by pedigree analysis.

Genotype proportions in a population with inbreeding: Genotype Due to Random Due to Observed moting inbreeding proportion

Genotype	Due to Random moting		Due to Observed inbreeding		proportion
AA	(1-F)	+	PF		P^2 +FPq
Aa	P^2(1-F)	+		=	2Pq(1-F)
aa	4^2(1-F)	+	qF	=	q^2+Fpq
Total	(p2+2P4+q2)-(1-F) (1-F)		(P+q)F F		1

To obtain the total probability of inbreeding, the values of each path must be added. F1= Σ (1/2)n (1+F1). Where F1 is the probability that the two alleles in 1 are identical, by decent n is the number of ancestors as a given path F_1 is the inbreeding co-efficient of the offspring of first cousin.

Fitness: The model for a diploid population is defined as follows. Generation are non overlapping and the population size N is held constant. A_s before, we consider a single locus with two alleles A1 and A2 with genotypes A1, A_1, A_1A_2 and A_2A_2 and reproduction fitness $(W = 1 - s)W_{11}$, W_{12} and W_{22}.

Selection against the recessive Homozygotic one locus with two alleles A and a

$In + 1 = P2(1 - s) + Pq/(1 - sq2)(1 - sq2) = q(4 - s + p)/(1 - sq^2)q(1 - Sq)/(1 - Sq)$. We assume that population is randomly mating and in Hardy weinbery suppose that there are i genes of type A1 and 2 N-i genes of type A_1. Then assessing selection affects survival between the rygote and adult stage as before, and denoting by p(a) = i/2N the gene frequency of Ai at generation n, the gene frequency of A1, at generation n, the gene frequency of A1 after selection becomes

		Genotypes			Total
		AA	Aa	aa	
Intial genotype frequencies		P^2	2Pq	q2	
Fitness (W)		1	1	1-s	
Ratio after selection	P^2	2Pq	P^2 (1-s)		1-s P^2=W^2
Genotype frequencies	P2	2P4	q2(1-s)		
After selection	W	W	W		

$\phi(n) = p(n)(p(n)W11 + (1 - pn)W12)/w(n)$, Where $W(n) = p(n)^2W11 + 2p(n)(1 - p(n))W12 + (1 - p(n)2W22$ is the average fitness. If mutation follows, selection than assuming systematic mutation with probability u, the gener frequency of A_1 after mutation between. $\phi(n) = \phi(n)(1 - u) + (1 - \phi(n)G$. The N individuals of the next generation are formed by sampling 2 N independent game t form the binomial pooling scheme. That is, if ϕ is the frequency of A_1 after selection and mutation, then the probability that there are j genes of type A_1 in the following generation is

$$P(y(n+1)) = \begin{pmatrix} 2N \\ J \end{pmatrix} \phi^j (1 - \phi i)^{2N-j}$$

The gene frequency at generation n+1 is thus j/2N is called transition probability.

Speciation $F = 1 - H/2pq$ (i.e 1-F(n) = F(n) Reliability Analysis). i.e to define the **inbreeding co-efficient F** of a population as the relative reduction in heterozygote in the population due to inbreeding. In an individual f(t) = $\lambda e^{-\lambda t}$ (Density function) its distribution is $f(t) = 1 - e^{-\lambda t}$, its lead to sensored survival data analysis of frailty model.

4. Conclusion

i) Fishers exact test an R X C contingency tables (Permutation)

ii) Random walk (Markov chain)

iii) Likelihood Ratio (L.R.) Method

iv) Bayessian Methods

Acknowledgement

Dr.(Miss). S.Muthalagi, Ph.D.(Zoo)Asst prof in zoology, Kolangiapper Govt Arts college, Virudhachalam, Dr.S.Thobias, Head of the Department in Statistics in Loyola college (Aut.) Madras for their valuable suggestions and moral support.

References

[1] John Felix, *Survival Data Analysis*, P.hD. submitted in Annamalai University, (2007).

[2] S.Thilagam, *Frailtey Models*, M.Phil., Thesis (unpublished) submitted in Madras University, (2003).

[3] N.John Britto, M.Phil., Thesis, Manonmaniam Sundranar University, (2010).

[4] Henry C.Tuckwell, *Elementary Application of probability theory*, 2^{nd}, Chapmen & Hall Publication, (1995).

[5] M.Rose, *Probability Models*, Academic Press.

[6] N.JohnBritto, *A Study on stochastic Models for HIV infection and Aids*, International journal of computing Algorithm, 03(3)(2014), 297-299.