Rev. Brasil. Genet. VII, 4, 689-707 (1984)
(Brazil. J. Genetics)

THE HARDY-WEINBERG EQUILIBRIUM UNDER A BAYESIAN PERSPECTIVE

Carlos Alberto de Bragança Pereira and André Rogatko

ABSTRACT

This paper presents Bayesian counterparts for the trinity of the classical Statistical analysis (test of hypothesis, estimation, and confidence regions) of the Hardy-Weinberg populational equilibrium. Numerical results are presented to illustrate the techniques and the main advantages of using these alternative methods.

1. INTRODUCTION

One of the concerns of a geneticist when analyzing a population is to know if it follows Hardy-Weinberg (HW) equilibrium (Hardy, 1908; Weinberg, 1908; Li, 1976). Several methods were developed to test the equilibrium hypothesis (Hogben, 1946; Levene, 1949; Haldane, 1954; Cannings and Edwards, 1969; Smith, 1970; Vithayasai, 1975; Emigh and Kempthorne, 1975; Chapco, 1976; Elston and Forthofer, 1977), all of them based on a chi-square test for goodness of fit where the maximum likelihood estimate is substituted for the unknown parameter or on variations of Fisher's exact test for small samples. Frequentely the null hypothesis is not rejected but further analysis is hindered since the power of the test cannot be evaluated.

The analysis presented in this work leads to a Bayesian counterpart for the trinity of the classical statistical analysis (test of hypothesis, estima-

Instituto de Matemática e Estatística, Universidade de São Paulo, Cidade Universitária, Caixa Postal 20.570 (Agência Iguatemi), 01000 São Paulo, SP, Brasil.

tion, and confidence regions) and allows the computation of the two kinds of errors.

Let A_1A_1, A_1A_2, and A_2A_2 be the genotypes related to a pair of alleles, A_1 and A_2; n_1, n_2, and n_3 their sample frequencies ($n_1 + n_2 + n_3 = n$); and p_1, p_2, and p_3 their populational proportions ($p_1 + p_2 + p_3 = 1$ and $p_i \geqslant 0$, $i = 1,2,3$) respectively. The likelihood function may be written as

$$L_1(p_1, p_3) = \frac{n!}{n_1!\,n_2!\,n_3!}\; p_1^{n_1}\, p_2^{n_2}\, p_3^{n_3},$$

since $p_2 = 1 - p_1 - p_3$ and the parameter space may be represented by the set

$$\Theta = \{(p_1 p_3); p_1 \geqslant 0, p_3 \geqslant 0, \text{and } p_1 + p_3 \leqslant 1\}.$$

A population follows HW equilibrium if and only if there is a number p ($0 \leqslant p \leqslant 1$) such that $p_1 = p^2$, $p_2 = 2p(1-p)$, and $p_3 = (1-p)^2$. Therefore, under HW equilibrium, the likelihood function may be written as

$$L_0^*(p) = \frac{n!}{n_1!\,n_2!\,n_3!}\; 2^{n_2}\, p^{2n_1+n_2}\, (1-p)^{2n_3+n_2}$$

where $0 \leqslant p \leqslant 1$.

The above equilibrium restrictions hold if and only if $p_1 = (1 - \sqrt{p_3})^2$ [this may be written as $p_3 = (1 - \sqrt{p_1})^2$] which is equivalent to $\dfrac{p_2^2}{p_1 p_3} = 4$. It is clear that, under the equilibrium hypothesis, the parameter space is reduced to the set

$$\Theta_0 = \{(p_1, p_3); p_3 = (1 - \sqrt{p_1})^2 \text{ and } 0 \leqslant p_1 \leqslant 1\}$$

which is a subset of Θ. These two sets Θ and Θ_0 are shown in Figure 1.

The Bayes factor discussed in Section 2 is a Bayesian counterpart of the HW equilibrium chi-square test. As a Bayesian counterpart for the estimation problem, a Bayes estimator for the parameter $\theta = p_2^2/p_1 p_3$ together with its variance is introduced in Section 3. The construction of credible regions (which may be viewed as Bayesian confidence regions) is discussed in Section 4. Finally, numerical examples illustrating the previous methodological results are shown in Section 5.

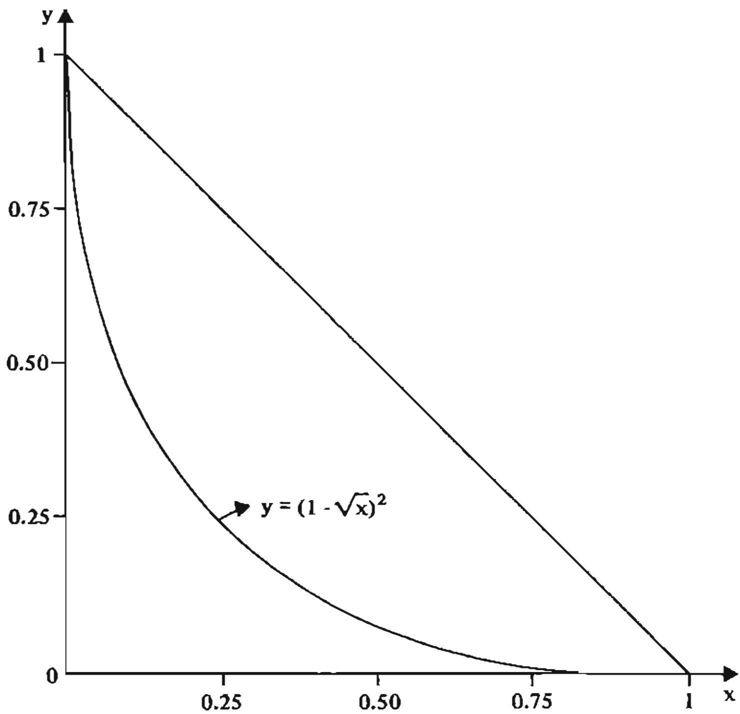

Figure 1 - The set Θ is the region inside the triangle. The set Θ_0 is the region covered by the curve $y = (1 - \sqrt{x})^2$. Clearly, $\Theta_0 \subset \Theta$

The choice of priors for (p_1, p_2, p_3) is restricted, in this paper, to the class of Dirichlet distributions because it is a conjugate class and enables the authors to obtain general analytic expressions.

2. THE BAYES FACTOR

The analysis discussed in this section is based on the Bayes factor for sharp hypotheses (Jeffreys, 1939). Although the choice of priors for (p_1, p_3) was restricted to the class of Dirichlet (generalized Beta) distributions, the technique used here is general and may be used for any assessed prior. The notation $(p_1, p_3) \sim D(\alpha_1, \alpha_2, \alpha_3)$ where $\alpha_i > 0$ $(i = 1, 2, 3)$, indicates that (p_1, p_3) is distributed as Dirichlet with parameters α_1, α_2, and α_3. Its density is given by

$$g(p_1,p_3) = \frac{\Gamma(\alpha)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1} p_3^{\alpha_3 - 1}$$

where $(p_1, p_3) \in \Theta$, $\alpha = \alpha_1 + \alpha_2 + \alpha_3$, and $\Gamma(\cdot)$ is the gamma function.

Combining the above prior with likelihood L_1, the predictive probability function (the marginal probability function of the data) under hypothesis H_1 of no equilibrium restriction is given by

$$f_1(n_1,n_2,n_3) = f(n_1,n_2,n_3 \mid H_1) =$$

$$= \int_0^1 \int_0^{1-x} g(x,y)L_1(x,y)dy\,dx = \frac{n!\,\Gamma(\alpha)}{\Gamma(n+\alpha)} \prod_{i=1}^{3} \frac{\Gamma(n_i+\alpha_i)}{n_i!\,\Gamma(\alpha_i)}$$

which is the probability function of a Dirichlet-Multinomial (DM) distribution with parameter $(n;\alpha_1,\alpha_2,\alpha_3)$. This distribution has been extensively studied (see, for example, Basu and Pereira, 1982).

Under hypothesis H_0 of HW equilibrium, the likelihood as a function of (p_1,p_3) may be expressed in two "different" reparametrizations:

$$L_0^*(\sqrt{p_1}) = \frac{n!}{n_1!n_2!n_3!} 2^{n_2} (\sqrt{p_1})^{2n_1+n_2} (1-\sqrt{p_1})^{2n_3+n_2}$$

and

$$L_0^*(1-\sqrt{p_3}) = \frac{n!}{n_1!n_2!n_3!} 2^{n_2} (1-\sqrt{p_3})^{2n_1+n_2} (\sqrt{p_3})^{2n_3+n_2}.$$

Although these two expressions lead to the same inference (under H_0) for every classical non-Bayesian analysis based on likelihoods, when a Bayesian view is considered this does not necessarily occur. For a Bayesian statistician, those likelihood functions are considered as conditional probabilities of the data given the parameters - the first expression above has p_1 as the conditioning argument while the second has p_3. Therefore, equivalent analysis based on $L_0^*(\sqrt{p_1})$ and on $L_0^*(1-\sqrt{p_3})$ may produce different results. This fact is known as the Borel-Kolmogorov paradox (Kolmogorov, 1950; Lindley, 1982): The information that an event has happened may be improved if one tells how one learned about it. In the case discussed here, there is no reason for a choice between $L_0^*(\sqrt{p_1})$ and $L_0^*(1-\sqrt{p_3})$. Since a joint distribution for (p_1,p_3) is being considered to measure prior information, the following likelihood function - which is an average of the previous ones and thus depends on both p_1 and p_3 - was chosen here to play the role of likelihood under H_0:

$$L_0(p_1,p_3) = \frac{1}{2}[L_0^*(\sqrt{p_1}) + L_0^*(1-\sqrt{p_3})].$$

The predictive probability function under H_0, obtained by combining $L_0(p_1,p_3)$ with the prior for (p_1,p_3), is given by

$$f_0(n_1,n_2,n_3) = f(n_1,n_2,n_3 \mid H_0) =$$

$$= \frac{n!}{n_1!n_2!n_3!} 2^{n_2-1}\{E[(\sqrt{X})^{2n_1+n_2}(1-\sqrt{X})^{2n_3+n_2}] +$$

$$+ E[(1-\sqrt{Y})^{2n_1+n_2}(\sqrt{Y})^{2n_3+n_2}]\},$$

where X, used for p_1, and Y, used for p_3, have Beta distributions with parameters $(\alpha_1, \alpha-\alpha_1)$ and $(\alpha_3, \alpha-\alpha_3)$, respectively, and $E[\cdot]$ is the expectation operator. By using $(1-x) = (1-\sqrt{x})(1+\sqrt{x})$, an alternative expression for this probability function is written as:

$$f_0(n_1,n_2,n_3) =$$

$$= \frac{n!}{n_1!n_2!n_3!} 2^{n_2}\{\frac{B(\alpha'_1+\alpha_1;\alpha'_3+\alpha_2)}{B(\alpha_1;\alpha_3+\alpha_2)} E[(1+Z)^{\alpha_2+\alpha_3-1}] +$$

$$+ \frac{B(\alpha'_1+\alpha_2;\alpha'_3+\alpha_3)}{B(\alpha_1+\alpha_2;\alpha_3)} E[(2-W)^{\alpha_1+\alpha_2-1}]\}$$

where $\alpha'_1 = 2n_1+n_2+\alpha_1$, $\alpha'_3 = 2n_3+n_2+\alpha_3$, $B(a;b)$ is the Beta function at point (a,b), and Z and W have Beta distributions with parameters $(\alpha'_1+\alpha_1;\ \alpha'_3+\alpha_2)$ and $(\alpha'_1+\alpha_2;\ \alpha'_3+\alpha_3)$, respectively. For the symmetric prior $(\alpha_1 = \alpha_2 = \alpha_3)$, the above expression is reduced to

$$f_0(n_1,n_2,n_3) =$$

$$= \frac{n!\,2^{n_2}}{n_1!n_2!n_3!} \frac{B(\alpha'_1+\alpha_1;\alpha'_3+\alpha_1)}{B(\alpha_1;2\alpha_1)} \{E[(1+Z)^{2\alpha_1-1}] +$$

$$+ E[(2-Z)^{2\alpha_1-1}]\},$$

where Z now has Beta distribution with parameter $(\alpha_1' + \alpha_1; \alpha_3' + \alpha_1) = (2n_1 + n_2 + 2\alpha_1; 2n_3 + n_2 + 2\alpha_1)$. Finally, for the uniform prior $(\alpha_1 = \alpha_2 = \alpha_3 = 1)$, the predictive probability function under H_0 is:

$$f_0(n_1, n_2, n_3) = \frac{6n!\, 2^{n_2}}{n_1!\, n_2!\, n_3!} B(2n_1 + n_2 + 2; 2n_3 + n_2 + 2).$$

Let ξ be the probability that H_0 (the HW equilibrium hypothesis) holds prior to the observation of the data (that is, ξ is the prior probability of H_0) and denote by $\xi(n_1, n_2, n_3)$ the posterior (after the data have been observed) probability of H_0. By applying the Bayes formula,

$$\xi(n_1, n_2, n_3) = \frac{\xi f_0(n_1, n_2, n_3)}{\xi f_0(n_1, n_2, n_3) + (1 - \xi) f_1(n_1, n_2, n_3)}.$$

The decision criterion to accept or reject H_0 is based on the odds of H_0. The prior odds are $0 = \frac{\xi}{1-\xi}$ and the posterior odds are $0(n_1, n_2, n_3) = \frac{\xi}{1-\xi} \frac{f_0(n_1, n_2, n_3)}{f_1(n_1, n_2, n_3)}$ which is the prior odds times the likelihood ratio. For the particular case of uniform priors the posterior odds have the following sample expression:

$$0(n_1, n_2, n_3) = 3 \frac{\xi}{1-\xi} \frac{(n+2)!\, 2^{n_2}}{n_1!\, n_2!\, n_3!} B(2n_1 + n_2 + 2; 2n_3 + n_2 + 2).$$

Usually, when $0(n_1, n_2, n_3) < 1$ the decision is against H_0 and conversely $0(n_1, n_2, n_3) \geqslant 1$ favors H_0. However, a decision criterion may be properly reached by a suitable choice of a number c such that if $0(n_1, n_2, n_3) < c$ then H_0 is rejected and it is accepted otherwise. Corresponding to this rule, there is a region $R_c = \{(n_1, n_2, n_3); 0(n_1, n_2, n_3) < c\}$ of sample points leading to the rejection of H_0. Based on R_c, the probabilities of the first kind and the second kind of errors are defined respectively as:

$$\alpha = \Sigma_c f_0(n_1, n_2, n_3) \text{ and}$$

$$\beta = 1 - \Sigma_c f_1(n_1, n_2, n_3)$$

where Σ_c indicates the summation over all points of R_c. The analogy with significance level and power of a test is natural. Table I shows the values of α and β for some different sample sizes (n), where c = 1 and the uniform prior were chosen.

Table I - Values of α and β for different sample sizes (n). Uniform prior and c = 1 were chosen.

n	10	20	30	40	50	60	70	80	90	100
α	.1878	.1164	.0968	.0900	.0786	.0720	.0658	.0587	.0566	.0534
β	.4545	.4113	.3669	.3310	.3107	.2924	.2786	.2689	.2568	.2475

These probabilities may explain why in most situations, where one expects to reject hypothesis H_0, the data support non-rejection of H_0.

In order to illustrate the advantages of the method described in this section, numerical results are presented in Section 5.

If one decides to use $L_0^*(\sqrt{p_1})$ [or $L_0^*(1-\sqrt{p_3})$] as the likelihood, one needs to use a prior distribution for p_1 [or p_3] coherent with the prior for (p_1, p_3) already assessed. In the particular case of this work, p_1 and p_3 have (in the prior) marginal Beta distributions with parameters $(\alpha_1;\alpha_2+\alpha_3)$ and $(\alpha_3;\alpha_2+\alpha_1)$, respectively. Nevertheless, as already pointed out, the predictive distribution under H_0 when $L_0^*(\sqrt{p_1})$ is used differs from that obtained with $L_0^*(1-\sqrt{p_3})$. The study of Dickey and Lientz (1970) on sharp hypothesis problems presents alternative choices of likelihoods for the case where the parameter space is contained in the real line.

3. BAYES ESTIMATION

Another situation that characterizes the HW equilibrium is $\theta = \dfrac{p_2^2}{p_1 p_3} = 4$. Therefore θ is a parameter of interest.

The two following facts simplify the derivation of the Bayes estimator of θ:

1. If (p_1, p_3) is distributed as $D(\alpha_1;\alpha_2;\alpha_3)$ prior to data, then its posterior distribution is $D(\alpha_1+n_1;\alpha_2+n_2;\alpha_3+n_3)$.

2. If X_1, X_2, and X_3 are mutually independent gamma variables with

parameters $(\alpha_1 + n_1 ; \beta)$, $(\alpha_2 + n_2 ; \beta)$, and $(\alpha_3 + n_3 ; \beta)$, respectively, then $\left[\frac{X_1}{X}, \frac{X_2}{X}, \frac{X_3}{X}\right]$, where $X = X_1 + X_2 + X_3$, is distributed as $D(\alpha_1 + n_1 ; \alpha_2 + n_2 ; \alpha_3 + n_3)$. Here β is the scale parameter.

These two results imply that θ is distributed as $\frac{X_2^2}{X_1 X_3}$ and then a simple proof for the following proposition is obtained.

Proposition. If $\alpha_1 + n_1 > 1$ and $\alpha_3 + n_3 > 1$, then the Bayes estimator (the posterior mean) of θ is $\hat{\theta} = \frac{(\alpha_2 + n_2)(\alpha_2 + n_2 + 1)}{(\alpha_1 + n_1 - 1)(\alpha_3 + n_3 - 1)}$. In addition, if $n_1 + \alpha_1 > 2$ and $n_3 + \alpha_3 > 2$, then the posterior variance of θ is

$$\mathrm{Var}[\theta \mid (n_1, n_2, n_3)] = \hat{\theta}\left[\frac{(\alpha_2 + n_2 + 2)(\alpha_2 + n_2 + 3)}{(\alpha_1 + n_1 - 2)(\alpha_3 + n_3 - 2)} - \hat{\theta}\right].$$

Proof. Since $\hat{\theta} = E\left[\frac{X_2^2}{X_1 X_3}\right]$, $\mathrm{Var}[\theta \mid (n_1, n_2, n_3)] = \mathrm{Var}\left[\frac{X_2^2}{X_1 X_3}\right]$, and X_1, X_2, and X_3 are mutually independent, the proof is completed by properly using the following moments:

(i) For any positive integer, k

$$E[X_2^k] = \frac{(\alpha_2 + n_2)(\alpha_2 + n_2 + 1)\ldots(\alpha_2 + n_2 + k - 1)}{\beta^k}$$

(ii) for $\alpha_i + n_i > 1$, $i = 1, 2, 3$

$$E[X_i^{-1}] = \frac{\beta}{\alpha_i + n_i - 1}$$

(iii) For $\alpha_i + n_i > 2$, $i = 1, 2, 3$

$$E[X_i^{-2}] = E[X_i^{-1}] \frac{\beta}{n_i + \alpha_i - 2} \quad \square.$$

By looking at (ii) and (iii) it becames clear that, changing the restrictions on $\alpha_i + n_i$, general expressions for other moments of θ may be obtainned.

Numerical illustrations of the above results are presented in Section 5.

4. CREDIBLE REGIONS

The Bayesian counterpart for confidence regions of level α is the smallest subset, Θ_α, of the parameter space Θ such that, in the posterior distribution the event $(p_1, p_3) \in \Theta_\alpha$ has probability α.

Definition. The credible region of level α for the parameter $\theta = \frac{p_2^2}{p_1 p_3}$ is the image Θ'_α of the function $\theta: \Theta_\alpha \rightarrow \Theta'_\alpha$ where $\theta(x,y) = \frac{(1-x-y)^2}{xy}$ and Θ_α is defined above.

Usually, it is not very simple to obtain this region Θ'_α. However, a good approximation may be obtained by using the mean and the variance of the posterior distribution of θ presented in Section 3.

Another alternative way to test hypothesis H_0 of HW equilibrium is to verify whether the set $\Theta_\alpha \cap \Theta_0$ is empty or not. The intuition here follows the classical test based on confidence regions.

To solve the equation leading to a credible region α for (p_1, p_3) a double integration is required. However, one of the integration limits cannot be set in an explicit form because it is a solution of

$$x^{n_1} y^{n_3} (1-x-y)^{n_2} = c,$$

where c is a constant. Hence, an evaluation of the credible region cannot be reached by means of numerical methods. Nevertheless, a good approximation can be obtained by simulation.

Recall that if X_1, X_2, and X_3 are independent variables with gamma distribution with parameters $(n_1+1;1), (n_2+1;1)$, and $(n_3+1;1)$ respectively, then for $X = X_1 + X_2 + X_3$, $(\frac{X_1}{X}, \frac{X_2}{X}, \frac{X_3}{X})$ is distributed as

$$D(n_1+1; n_2+1; n_3+1).$$

Since methods to simulate the gamma distribution are available (Robinson and Lewis, 1975; Schmeiser and Lal, 1979) a set *P* of N points (p_1, p_3), $p_1 = X_1/X$ and $p_3 = X_3/X$, was obtained through the simulation of N values of X_1, N values of X_2, and N values of X_3. The values of $L_1(p_1, p_3)$ were evaluated for each point of *P* and then sorted in a descending way. The αN points of *P* which correspond to the greatest values of $L_1(p_1, p_3)$ were plotted. This procedure is numerically illustrated in the following section.

5. NUMERICAL ILLUSTRATIONS

In this section, numerical applications of the results discussed in the previous sections are presented. For simplicity, the uniform distribution was chosen to be used as the prior in all illustrations.

The triangular arrays presented in the appendix show the integer part of the Bayes factor for each possible value of (n_1, n_3) in the cases of $n = 10i$ where $i = 1,2,3$ and 5. For large sample sizes the agreement wlth the equilibrium curve in Figure 1 is suggestive.

Tables II and III present, for the sample size $n = 20$, the integer part of the Bayes estimator, $\hat{\theta}$, multiplied by 10 and the posterior standard deviation of θ multiplied by 10, respectively.

Figure 2 shows the likelihood function $L_1(p_1, p_3)$ for the sample $(n_1, n_2, n_3) = (25, 50, 25)$. Figures 3 and 4 show the 95% credible regions (zeros) for samples (25, 50, 25) and (49, 2, 49), respectively. They were obtained by the procedure described in Section 3 for $N = 1300$ generated points in each case. In Figure 3 the intersection between the credible region and the equilibrium curve ("E" characters) is not empty. Conversely, in Figure 4, the set of equilibrium points, represented by the equilibrium curve, and the credible region are disjoint. It is concluded that in the first case the population is in HW equilibrium, and in the second case it is not. Both conclusions are based on a "credible level" of $\alpha = .95$.

6. FINAL REMARKS

The methods presented in this article to test the HW equilibrium hypothesis have the advantage that they can be applied to samples of any size. However, the magnitude of the α error and especially of the β error indicate that satisfactory decisions can only be reached for samples of substantial size.

This paper exemplifies how the Bayesian methodology may produce powerful tools to help the researcher to reach good decisions which can be accurately evaluated.

Table II - Integer part of the Bayes estimator $\hat{\theta}$ multiplied by 10 for sample size n = 20.

20																				
19	1																			
18	3	1																		
17	7	2	0																	
16	13	4	1	0																
15	20	7	3	1	0															
14	30	11	5	2	1	0														
13	43	16	8	4	2	1	0													
12	60	23	12	6	3	2	1	0												
11	82	33	17	10	5	3	2	1	0											
10	110	45	24	14	8	5	3	2	1	0										
9	147	61	33	20	12	8	5	3	1	1	0									
8	195	83	46	28	18	12	8	5	3	2	1	0								
7	260	111	63	39	26	17	11	8	5	3	2	1	0							
6	350	152	87	55	37	25	17	12	8	5	3	2	1	0						
5	480	210	121	78	53	37	26	18	12	8	5	3	2	1	0					
4	680	300	175	114	78	55	39	28	20	14	10	6	4	2	1	0				
3	1020	453	267	175	121	87	63	46	33	24	17	12	8	5	3	1	0			
2	1710	765	453	300	210	152	111	83	61	45	33	23	16	11	7	4	2	1		
1	3800	1710	1020	680	460	350	260	195	147	110	82	60	42	30	20	13	7	3	1	
$n_1 \rightarrow$ 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
↑ n_3																				

Table III - Integer part of the posterior standard deviation of θ multiplied by 10 for sample size n = 20.

20																				
19																				
18		2																		
17		4	1																	
16		8	3	1																
15		12	5	2	1															
14		19	7	3	2	1														
13		27	11	5	3	1	1													
12		37	16	8	5	3	1	1												
11		50	22	12	7	4	2	1	1											
10		68	30	16	10	6	4	2	1	1										
9		90	40	23	14	9	6	4	2	1	1									
8		121	55	31	20	13	9	6	4	2	1	1								
7		163	75	43	27	18	12	9	6	4	2	1	1							
6		223	103	60	39	26	18	13	9	6	4	3	1	1						
5		314	147	87	57	36	27	20	14	10	7	5	3	2	1					
4		465	221	132	87	60	43	31	23	16	12	8	5	3	2	1				
3		755	365	221	147	103	75	55	40	30	22	16	11	7	5	3	1			
2		1524	755	465	214	223	163	121	90	68	50	37	27	19	12	8	4	2		
1																				
$n_1 \rightarrow$ 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
↑ n_3																				

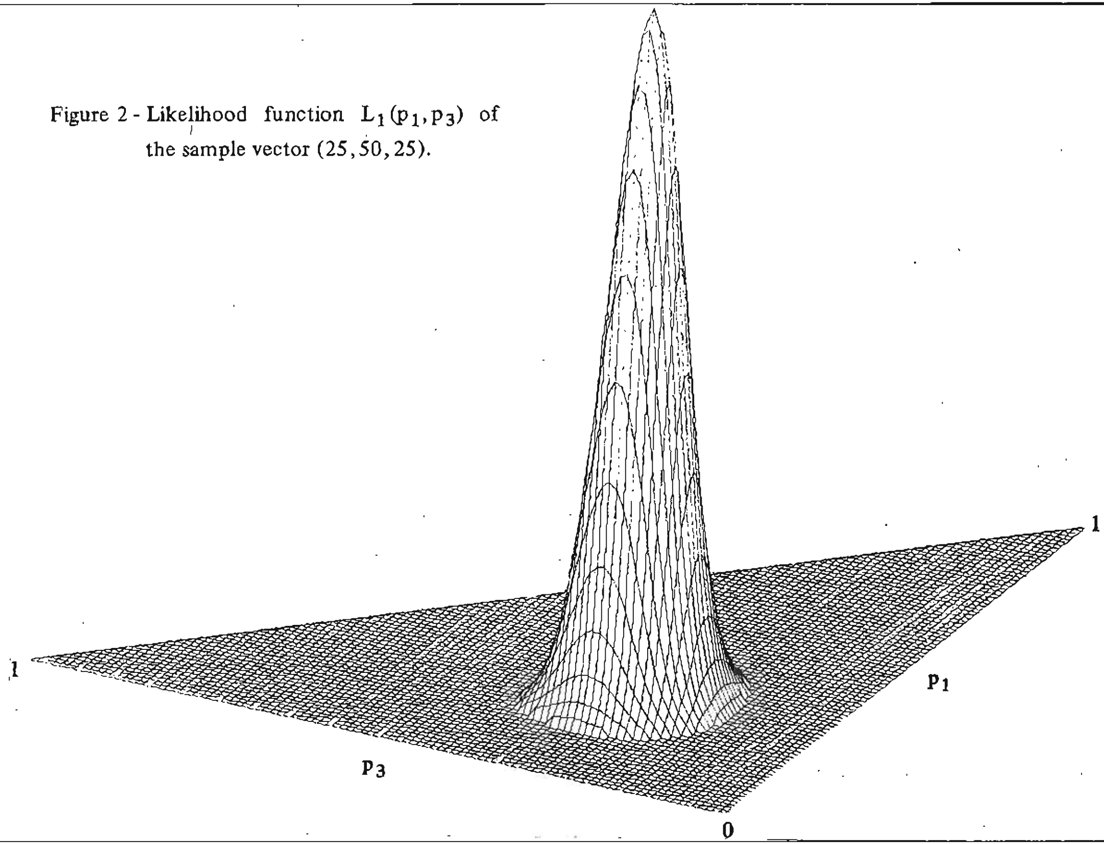

Figure 2 - Likelihood function $L_1(p_1, p_3)$ of the sample vector (25, 50, 25).

Figure 3 - 95% credible region (zeros) obtained by simulation, where 1300 points were generated for the sample vector (25, 50, 25). The equilibrium curve is represented by "E" characters.

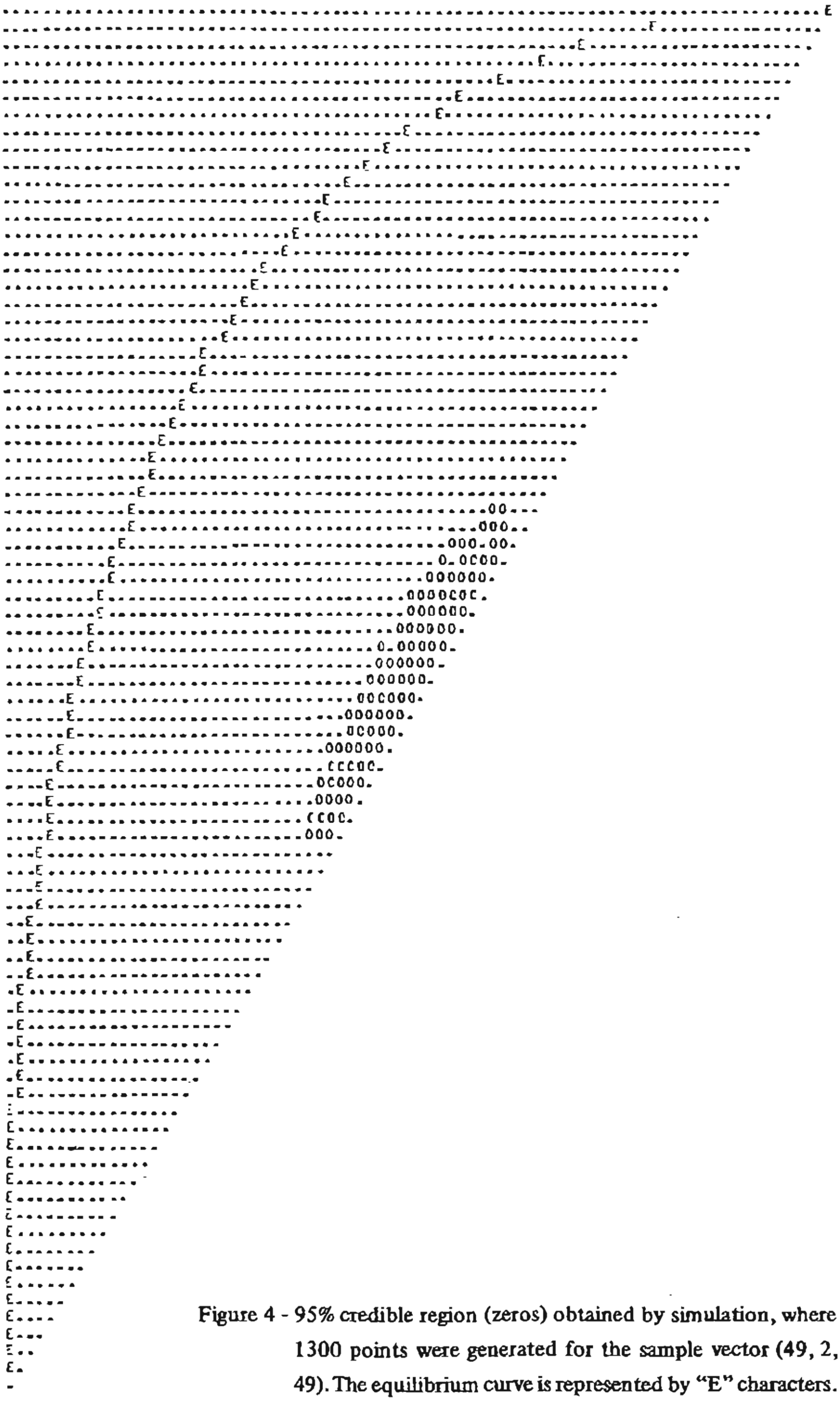

Figure 4 - 95% credible region (zeros) obtained by simulation, where 1300 points were generated for the sample vector (49, 2, 49). The equilibrium curve is represented by "E" characters.

ACKNOWLEDGMENTS

We are grateful to FAPESP for financial support.

RESUMO

Este trabalho descreve uma alternativa Bayesiana para a trindade da Análise Estatística Clássica (teste de hipótese, estimação e região de confiança) do equilíbrio populacional de Hardy e Weinberg. Resultados numéricos são apresentados para ilustrar esta metodologia e suas principais vantagens em relação à clássica.

REFERENCES

Basu, D. and Pereira, C.A.B. (1982). On the Bayesian analysis of categorical data: the problem of nonresponse. *J. Stat. Plan. Inf. 6*: 345-362.

Cannings, C. and Edwards, A.W.F. (1969). Expected genotypic frequencies in a small sample: deviation from Hardy-Weinberg equilibrium. *Am. J. Hum. Genet. 21*: 245-247.

Chapco, W. (1976). An exact test of the Hardy-Weinberg law. *Biometrics 32*: 183-189.

Dickey, J.M. and Lientz, B.P. (1970). The weighted likelihood ratio sharp hypothesis about chances, the order of a Markov chain. *Ann. Math. Stat. 41*: 214-226.

Elston, R.C. and Forthofer, R. (1977). Testing for Hardy-Weinberg equilibrium in small samples. *Biometrics 33*: 536-542.

Emigh, T.H. and Kempthorne, O. (1975). A note on goodness of fit of a population to Hardy-Weinberg structure. *Am. J. Hum. Genet. 27*: 778-783.

Haldane, J.B.S. (1954). An exact test for randommess of mating. *J. Genet. 52*: 631-635.

Hardy, G.H. (1908). Mendelian propositions in a mixed population. *Science 28*: 49-50.

Hogben, L. (1946). *An Introduction to Mathematical Genetics.* Northon, New York.

Jeffreys, H. (1939). *Theory of Probability.* Clarendon, Oxford.

Kolmogorov, A.N. (1950). *Foundations of the Theory of Probability.* Chelsea, New York.

Levene, H. (1949). On a matching problem arising in genetics. *Ann. Math. Stat. 20*: 91-94.

Li, C.C. (1976). *First Course in Population Genetics.* Booxwood, California.

Lindley, D.V. (1982). The Bayesian approach to statistics. In: *Some Recent Advances in Statistics* (Oliveira, T. de and Epstein, B., eds.). Academic Press, New York, pp. 65-87.

Robinson, D.W. and Lewis, P.A.W. (1975). Generating gamma and Cauchy random variables: an extension to the Naval Postgraduate School random number package. Naval Postgraduate School, Monterey, California.

Schmeiser, B.W. and Lal, R. (1979). Squeeze methods for generating gamma variates. Southern Methodist University, Dallas, Texas.

Smith, C.A.B. (1970). A note on testing the Hardy-Weinberg law. *Ann. Hum. Genet. 33*: 277-283.

Vithayasai, C. (1975). Exact critical values of the Hardy-Weinberg test statistics for two alleles. *Comm. Stat. 1*: 229-242.

Weinberg, W. (1908). Uber den Nachweis der Vererbung beim Menschen. *Jahresh. Verein Vaterl. Naturk. in Wurttemberg 64*: 368-382.

(Received April 6, 1984)

APPENDIX

This Appendix presents tables for the integer part of the Bayes factor for four different sample sizes. The columns correspond to the values of n_1 and the rows to the values of n_3.

n = 10

10	0										
9	1	0									
8	2	0	0								
7	2	0	0	0							
6	2	1	0	0	0						
5	1	2	1	0	0	0					
4	1	2	1	0	0	0	0				
3	0	1	2	1	0	0	0	0			
2	0	1	1	2	1	1	0	0	0		
1	0	0	1	1	2	2	1	0	0	0	
0	0	0	0	0	1	1	2	2	2	1	0
	0	1	2	3	4	5	6	7	8	9	10

n = 20

20	0																				
19	1	0																			
18	2	0	0																		
17	2	0	0	0																	
16	2	0	0	0	0																
15	3	1	0	0	0	0															
14	2	2	0	0	0	0	0														
13	2	2	1	0	0	0	0	0													
12	2	2	1	0	0	0	0	0	0												
11	1	2	2	1	0	0	0	0	0	0											
10	1	2	2	1	0	0	0	0	0	0	0										
9	1	2	2	2	1	0	0	0	0	0	0	0									
8	0	1	2	2	2	1	0	0	0	0	0	0	0								
7	0	1	2	2	2	1	1	0	0	0	0	0	0	0							
6	0	0	1	2	2	2	1	1	0	0	0	0	0	0	0						
5	0	0	1	2	2	2	2	1	1	0	0	0	0	0	0	0					
4	0	0	0	1	2	2	2	2	2	1	0	0	0	0	0	0	0				
3	0	0	0	0	1	2	2	2	2	2	1	1	0	0	0	0	0	0			
2	0	0	0	0	0	1	1	2	2	2	2	2	1	1	0	0	0	0	0		
1	0	0	0	0	0	0	0	1	1	2	2	2	2	2	2	1	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	1	1	1	2	2	2	3	2	2	2	1	0
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

n = 30

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

n = 50

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
50	0																																																		
49	1	0																																																	
48	2	0	0																																																
47	2	0	0	0																																															
46	3	0	0	0	0																																														
45	3	0	0	0	0	0																																													
44	4	1	0	0	0	0	0																																												
43	4	1	0	0	0	0	0	0																																											
42	4	1	0	0	0	0	0	0	0																																										
41	4	2	0	0	0	0	0	0	0	0																																									
40	4	3	0	0	0	0	0	0	0	0	0																																								
39	4	3	1	0	0	0	0	0	0	0	0	0																																							
38	4	3	1	0	0	0	0	0	0	0	0	0	0																																						
37	3	4	2	0	0	0	0	0	0	0	0	0	0	0																																					
36	3	4	2	1	0	0	0	0	0	0	0	0	0	0	0																																				
35	3	4	3	1	0	0	0	0	0	0	0	0	0	0	0	0																																			
34	2	4	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0																																		
33	2	4	4	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0																																	
32	1	3	4	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																
31	1	3	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0																															
30	1	3	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																														
29	0	2	4	4	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																													
28	0	2	3	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																												
27	0	1	3	4	4	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																											
26	0	1	3	4	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																										
25	0	1	2	3	4	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																									
24	0	0	2	3	4	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																								
23	0	0	1	2	4	4	3	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																							
22	0	0	1	2	3	4	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																						
21	0	0	0	1	3	4	4	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																					
20	0	0	0	1	2	3	4	4	3	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																				
19	0	0	0	1	2	3	4	4	4	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																			
18	0	0	0	0	1	2	3	4	4	3	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																		
17	0	0	0	0	1	2	3	3	4	4	3	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																	
16	0	0	0	0	0	1	2	3	4	4	4	3	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																
15	0	0	0	0	0	1	1	2	3	4	4	4	3	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0															
14	0	0	0	0	0	0	1	2	2	3	4	4	4	3	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0														
13	0	0	0	0	0	0	0	1	2	3	3	4	4	4	3	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0													
12	0	0	0	0	0	0	0	1	1	2	3	3	4	4	4	3	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0												
11	0	0	0	0	0	0	0	0	1	1	2	3	3	4	4	4	3	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0											
10	0	0	0	0	0	0	0	0	0	1	1	2	3	3	4	4	4	3	3	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0										
9	0	0	0	0	0	0	0	0	0	0	1	1	2	3	3	4	4	4	3	3	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0									
8	0	0	0	0	0	0	0	0	0	0	0	1	1	2	2	3	4	4	4	4	3	3	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0								
7	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	2	3	3	4	4	4	4	3	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0							
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	3	3	4	4	4	4	3	3	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	2	3	3	4	4	4	4	3	3	3	2	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0					
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	2	3	3	4	4	4	4	4	3	3	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0				
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	2	2	3	3	4	4	4	4	3	3	3	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0			
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	2	3	3	3	4	4	4	4	4	3	3	2	2	1	1	0	0	0	0	0	0	0	0	0		
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	2	2	3	3	3	4	4	4	4	4	3	3	3	2	1	1	1	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	2	2	3	3	3	4	4	4	4	4	4	4	3	3	2	2	1	0