Brazilian Journal of Genetics, 19, 3, 485-491 (1996)
(Revista Brasileira de Genética)

METHODOLOGY

Properties of estimators of the inbreeding coefficient and the rate of cross fertilization obtained from gene frequency data in a diploid population

Joel Augusto Muniz[1], *Décio Barbin*[2] *and Roland Vencovsky*[3]

ABSTRACT

The properties of the estimators of the inbreeding coefficient and the cross pollination rate were investigated in a diploid population. The estimates, obtained by the moments method, were based on the analysis of variance of the gene frequency of individuals from random samples. Since these estimators were obtained from the ratio of two random variables, approximations were produced by the Taylor series function. The formulas obtained were checked using simulation data. The results indicated that the estimators of the inbreeding and cross pollination rates are positively and negatively biased, respectively. The expression of these tendencies is a function of 1/n, which becomes smaller as the sample size increases. The simulation confirmed those results and the validity of the expressions to calculate the error of the estimates.

INTRODUCTION

The methodology of gene frequency data analysis (Weir, 1990; Vencovsky, 1992) uses the indicator variable Y, which has a value of one when a given allele, say A_1, is present in the individual and zero when it is absent and alleles A_2, A_3, ..., A_u are present. The estimation of the genetic parameters involved is possible from the frequencies of these variables analyzed by an analysis of variance using the procedures of Experimental Statistics in association with a statistical model, which is a function of the structure of the problem. The inferences concerning the parameters are affected by non-normality of the Y variable and, consequently, the ratio between two mean squares will not have the Snedecor F distribution. Furthermore, in some situations, the parameters are defined as a ratio among random variables and their estimator, obtained by the moments method, uses an approximation which considers the estimator of a ratio as the ratio of the estimates (Nei and Chakravarti, 1977; Weir and Cockerham, 1984).

METHODS

The statistical model associated with the analysis of variance of each allele is:

[1] Departamento de Ciências Exatas, Universidade Federal de Lavras, Caixa Postal 37, 37200-000 Lavras, MG, Brasil. Send correspondence to J.A.M.

[2] Departamento de Matemática e Estatística, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, 13418-900 Piracicaba, SP, Brasil.

[3] Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, 13418-900 Piracicaba, SP, Brasil.

$$y_{ij} = p + a_i + g_{j(i)}$$

where

$\mathbf{y_{ij}}$ is the frequency of the j-th gene within the i-th individual, corresponding to the values of a binary variable which assumes the value one, if the j-th gene within the i-th individual is the allele A_1 and assumes the value zero, otherwise;
p is the parametric frequency of the **A** allele in the population;
$\mathbf{a_i}$ is the effect of the i-th individual, with i = 1, 2, ..., n;
$\mathbf{g_{j(i)}}$ is the effect of the j-th gene within the i-th individual, with j = 1, 2.

The model is random (Cockerham, 1969) and Table I shows the analysis of variance.

Table I - Analysis of variance of the gene frequencies of an allele, in samples of n individuals of a population.

Source of variation	d.f.	MS	E[MS]
Individual	n-1	MSI	p(1 - p)(1 - F) + 2p(1 - p)F
Genes/individual	n	MSG	p(1 - p)(1 - F)

The variance component due to genes can be identified in Table I as

$$\sigma_G^2 = p(1 - p)(1 - F)$$

whose estimate by the moments method gives

$$\hat{\sigma}_G^2 = MSG.$$

The variance component due to individuals is

$$\sigma_I^2 = p(1 - p)F$$

which is estimated by the moments method as

$$\hat{\sigma}_I^2 = \frac{MSI - MSG}{2}.$$

The total variance is obtained by the sum of the variance components due to genes and individuals

$$\sigma^2 = \sigma_G^2 + \sigma_I^2 = p(p - 1)(1 - F) + p(1 - p)F = p(1 - p)$$

being estimated by

$$\hat{\sigma}^2 = \hat{\sigma}_G^2 + \hat{\sigma}_I^2 = \frac{MSG + MSI}{2}.$$

The ratio among the variance components due to individuals and the total variance gives the parametric value of the mean inbreeding coefficient (F) of the individuals. Using an approximation which takes the estimator of a ratio as the ratio of the estimates, the estimator of F is given by,

$$\hat{F} = \frac{\hat{\sigma}_I^2}{\hat{\sigma}^2} = \frac{MSI - MSG}{MSI + MSG}.$$

For the model under study, if the population is in Wright's equilibrium, that is, in balance with inbreeding, the rate of cross pollination, t, is related to F through the expression

$$F = \frac{1 - t}{1 + t},$$

which, when solved for t, gives

$$t = \frac{1 - F}{1 + F}.$$

Its estimate by the moments method is given by

$$\hat{t} = \frac{1 - \hat{F}}{1 + \hat{F}}.$$

The estimators of F and t are ratios of random variables. Thus the expressions proposed for the mean and variance of the estimators are based on development through Taylor's series, at the point $x = \mu_x$ and $x = \mu_y$ (Kendal and Stuart, 1963; Mood *et al.*, 1974) which define the mean and the variance of a ratio, respectively as

$$E[X/Y] \approx \mu_x/\mu_y - 1/\mu_y^2 \, Cov[X,Y] + \mu_x/\mu_y^3 \, Var[Y]$$

and

$$Var[X/Y] \approx \frac{\mu_y^2 Var[X] + \mu_x^2 Var[Y] - 2\mu_x \mu_y Cov[X,Y]}{\mu_y^4}.$$

The estimator expressions and their variances were numerically assessed using the results of the calculation in 100 experiments simulated using the SAS package (Statistical Analysis System), taking the sample sizes n = 10, 30, 100 and 200, extracted from 25 diploid populations defined by combination among the gene frequencies 0.1, 0.3, 0.5, 0.7, and 0.8 and inbreeding coefficients 0.1, 0.2, 0.3, 0.4, and 0.5. In the simulation experiment the RANUNI function (SAS/GRAPH, 1990) was used to generate random numbers. This function produces values of a random variable uniformly distributed in the interval limited by zero and one.

In each experiment the values of the estimators F and t and their variances were calculated. In the N = 100 experiments, the estimator means and variances were calculated: $\bar{\hat{F}}$, $\bar{\hat{t}}$, $\bar{V\hat{a}r}(\hat{F})$ and $\bar{V\hat{a}r}(\hat{t})$, respectively, as well as the variances $Var(\hat{F})$ and $Var(\hat{t})$. The comparison among the results obtained in the 100 experiments with the parametric values $E(\hat{F})$, $E(\hat{t})$, $Var_p(\hat{F})$ and $Var_p(\hat{t})$ allowed the assessment of the properties of the estimators.

RESULTS AND DISCUSSION

The expected value of the F estimator was obtained from the expression proposed (Kendal and Stuart, 1963; Mood *et al.*, 1974) for the ratio among two random variables: X = MSI - MSG and Y = MSI + MSG.

If the number of individuals in the sample, the gene frequency and the inbreeding rate of the population allow the assumption of normality of the gene frequency distribution of the sample (Mood *et al.*, 1974; Robertson and Hill, 1984), then the expression

$$\frac{v\,MS}{E[MS]} \approx \chi^2_v,$$

according to Searle (1971). After algebraic development this relationship becomes

$$E[\hat{F}] \approx E\left[\frac{X}{Y}\right] = E\left[\frac{MSI - MSG}{MSI + MSG}\right] \approx F - \frac{F(1-F^2)}{n}.$$

The result shows that the F estimator obtained by the moments method was biased. The bias was negative, indicating that the F estimates tend, on average, to be lower than the parametric value. The bias is a function of 1/n and approximates zero when n becomes large.

Considering the same χ^2 approximation as before and the expression proposed (Kendal and Stuart, 1963; Mood *et al.*, 1974) for the variance of a ratio among random variables, the following variance of the F estimator was obtained:

$$Var[\hat{F}] \approx \frac{(1-F^2)^2}{n},$$

which is the equivalent of the expression shown in (Fisher, 1970; Johnson and Kotz, 1970) for the variance of the estimate of the correlation coefficient when working with large, medium or small samples. According to the authors, these considerations allow the assumption of normality for the distribution of the sample correlation coefficient. A $Var[\hat{F}]$ estimator was obtained replacing F with its estimate $\hat{F}$, giving the expression:

$$V\hat{a}r[F] \approx \frac{(1-\hat{F}^2)^2}{n}.$$

These procedures when used for the **t** estimator give:

$$E[\hat{t}] \approx \frac{1-F+F/n(1-F^2)}{1+F-F/n(1-F^2)} + \frac{2}{n}\,\frac{(1-F^2)^2}{[1+F-F/n(1-F^2)]^3}$$

showing that the t estimator is also biased. The bias of t, however, was positive indicating that this estimate tends to be superior to the parametric value.

The variance of the **t** estimator is

$$Var(\hat{t}) \approx \frac{(2t)^2}{n},$$

and its estimate is given by

$$Var(\hat{t}) \approx \frac{(2t)^2}{n}.$$

Tables II and III contain results of simulation studies which allowed the evaluation of the estimators and their variances.

The mean values of $\hat{F}$ in Tables II and III, on the whole, are lower that the parametric values, confirming the negative bias of the estimator. The mean values of $\hat{t}$ show the positive bias of the estimator for the rate of cross pollination. In the two cases there is a reduction in the bias, when the size of the sample is increased and also when the frequency of the A allele in the population approaches 0.5.

The proposed expression for estimating the variances of the estimators showed satisfactory results. The mean values of $V\hat{a}r(\hat{F})$ and $V\hat{a}r(\hat{t})$ are close to the parametric values and to those calculated with the estimates in 100 experiments as the sample size was increased in the population with lower inbreeding and frequency of the A allele near 0.5.

In general, the results of Tables II and III indicate that the estimates of F and t, obtained by the moments method, and their variances may be used when working with at least 30 individuals in populations with an inbreeding coefficient lower than 0.4, and where the loci are sufficiently polymorphic with the most frequent allele having frequencies between 0.3 and 0.7.

Table II - Mean and variance of the F estimator in N = 100 experiments with different size samples, taken from a population with two alleles with diverse combinations of gene frequency and inbreeding coefficients(*).

N	p	F	$E(\hat{F})$	$\bar{\hat{F}}$	$Var_p(\hat{F})$	$Var(\hat{F})$	$\bar{V\hat{a}r}(\hat{F})$
10	0.1	0.1	0.0901	0.0440	0.0980	0.0610	0.0832
30			0.0967	0.0888	0.0327	0.0520	0.0299
100			0.0990	0.1096	0.0098	0.0157	0.0095
200			0.0995	0.0862	0.0049	0.0086	0.0048
10	0.1	0.2	0.1808	0.1220	0.0922	0.0842	0.0662
30			0.1936	0.1713	0.0307	0.0568	0.0285
100			0.1981	0.1893	0.0092	0.0211	0.0089
200			0.1990	0.1847	0.0046	0.0120	0.0046
10	0.1	0.3	0.2727	0.1890	0.0828	0.1341	0.0565
30			0.2909	0.2471	0.0276	0.0842	0.0255
100			0.2973	0.2753	0.0083	0.0216	0.0082
200			0.2986	0.2958	0.0041	0.0083	0.0041
10	0.1	0.4	0.3664	0.2771	0.0706	0.1595	0.0549
30			0.3888	0.3455	0.0235	0.0949	0.0219
100			0.3966	0.3844	0.0071	0.0246	0.0070
200			0.3983	0.4030	0.0035	0.0137	0.0034
10	0.1	0.5	0.4625	0.2362	0.0563	0.1491	0.0508
30			0.4875	0.4472	0.0188	0.0846	0.0189
100			0.4963	0.5009	0.0056	0.0168	0.0055
200			0.4981	0.4892	0.0028	0.0104	0.0029
10	0.3	0.1	0.0901	0.1069	0.0980	0.1019	0.0804
30			0.0967	0.0935	0.0327	0.0337	0.0307
100			0.0990	0.1037	0.0098	0.0139	0.0095
200			0.0995	0.0851	0.0049	0.0046	0.0049
10	0.3	0.2	0.1808	0.1980	0.0922	0.0983	0.0769
30			0.1936	0.2146	0.0307	0.0467	0.0279
100			0.1981	0.2000	0.0092	0.0099	0.0090
200			0.1990	0.1964	0.0046	0.0059	0.0046
10	0.3	0.3	0.2727	0.2728	0.0828	0.1135	0.0711
30			0.2909	0.2927	0.0276	0.0281	0.0265
100			0.2973	0.2893	0.0083	0.0098	0.0083
200			0.2986	0.2988	0.0041	0.0049	0.0041
10	0.3	0.4	0.3664	0.3646	0.0706	0.1325	0.0611
30			0.3888	0.3837	0.0235	0.0364	0.0229
100			0.3966	0.4053	0.0071	0.0084	0.0069
200			0.3983	0.3885	0.0035	0.0052	0.0036
10	0.3	0.5	0.4625	0.4955	0.0563	0.1007	0.0504
30			0.4875	0.4757	0.0188	0.0391	0.0192
100			0.4963	0.4950	0.0056	0.0089	0.0056
200			0.4981	0.4781	0.0028	0.0043	0.0030
10	0.5	0.1	0.0901	0.0375	0.0980	0.1203	0.0796
30			0.0967	0.0949	0.0327	0.0293	0.0309
100			0.0990	0.1055	0.0098	0.0086	0.0096
200			0.0995	0.1008	0.0049	0.0037	0.0049
10	0.5	0.2	0.1808	0.2164	0.0922	0.1021	0.0754
30			0.1936	0.2045	0.0307	0.0257	0.0292
100			0.1981	0.1956	0.0092	0.0087	0.0091
200			0.1990	0.2092	0.0046	0.0046	0.0045
10	0.5	0.3	0.2727	0.2314	0.0828	0.0995	0.0753
30			0.2909	0.3198	0.0276	0.0291	0.0256
100			0.2973	0.2870	0.0083	0.0067	0.0083
200			0.2986	0.3148	0.0041	0.0033	0.0040
10	0.5	0.4	0.3664	0.4114	0.0706	0.0837	0.0614
30			0.3888	0.4130	0.0235	0.0276	0.0221
100			0.3966	0.3860	0.0071	0.0117	0.0071
200			0.3983	0.3949	0.0035	0.0042	0.0031
10	0.5	0.5	0.4625	0.4935	0.0563	0.0758	0.0536
30			0.4875	0.5142	0.0188	0.0226	0.0177
100			0.4963	0.4962	0.0056	0.0073	0.0056
200			0.4981	0.4900	0.0028	0.0036	0.0029
10	0.7	0.1	0.0901	0.0836	0.0980	0.1084	0.0803
30			0.0967	0.0887	0.0327	0.0452	0.0301
100			0.0990	0.0983	0.0098	0.0100	0.0096
200			0.0995	0.1144	0.0049	0.0058	0.0048
10	0.7	0.2	0.1808	0.1904	0.0922	0.1142	0.0758
30			0.1936	0.2060	0.0307	0.0312	0.0288
100			0.1981	0.1952	0.0092	0.0123	0.0090
200			0.1990	0.1891	0.0046	0.0053	0.0046
10	0.7	0.3	0.2727	0.2364	0.0828	0.1220	0.0719
30			0.2909	0.2809	0.0276	0.0396	0.0265
100			0.2973	0.2828	0.0083	0.0092	0.0083
200			0.2986	0.2962	0.0041	0.0048	0.0041
10	0.7	0.4	0.3664	0.3287	0.0706	0.1081	0.0654
30			0.3888	0.4052	0.0235	0.0353	0.0221
100			0.3966	0.4078	0.0071	0.0096	0.0068
200			0.3983	0.3975	0.0035	0.0041	0.0035
10	0.7	0.5	0.4625	0.4798	0.0563	0.0959	0.0529
30			0.4875	0.4868	0.0188	0.0252	0.0189
100			0.4963	0.4771	0.0056	0.0091	0.0059
200			0.4981	0.4956	0.0028	0.0035	0.0028
10	0.9	0.1	0.0901	0.0578	0.0980	0.0756	0.0785
30			0.0967	0.0966	0.0327	0.0669	0.0294
100			0.0990	0.0744	0.0098	0.0167	0.0096
200			0.0995	0.0785	0.0049	0.0062	0.0049
10	0.9	0.2	0.1808	0.1009	0.0922	0.0940	0.0704
30			0.1936	0.1688	0.0307	0.0614	0.0277
100			0.1981	0.1540	0.0092	0.0180	0.0092
200			0.1990	0.1983	0.0046	0.0087	0.0045
10	0.9	0.3	0.2727	0.1298	0.0828	0.1004	0.0683
30			0.2909	0.2989	0.0276	0.0758	0.0241
100			0.2973	0.2603	0.0083	0.0225	0.0083
200			0.2986	0.3035	0.0041	0.0135	0.0040
10	0.9	0.4	0.3664	0.2055	0.0706	0.1385	0.0617
30			0.3888	0.3337	0.0235	0.0900	0.0233
100			0.3966	0.3720	0.0071	0.0285	0.0071
200			0.3983	0.4026	0.0035	0.0135	0.0034
10	0.9	0.5	0.4625	0.2475	0.0563	0.1626	0.0489
30			0.4875	0.4444	0.0188	0.0844	0.0194
100			0.4963	0.4755	0.0056	0.0244	0.0058
200			0.4981	0.5069	0.0028	0.0092	0.0027

(*) $E(\hat{F})$: Expected $\hat{F}$ value; $\bar{\hat{F}}$: mean of $\hat{F}$ in the 100 experiments; $Var_p(\hat{F})$: parametric variance of $\hat{F}$; $Var(\hat{F})$: variance of $\hat{F}$ in the 100 experiments; $\bar{V\hat{a}r}(\hat{F})$: mean of the 100 estimates of $Var(\hat{F})$ obtained from $V\hat{a}r(\hat{F})$.

Table III - Mean and variance of the **t** estimator in N = 100 experiments with different size samples, taken from a population with two alleles with a diverse combination of gene frequencies and inbreeding coefficients(*).

N	p	F	t	$E(\hat{t})$	$\bar{\hat{t}}$	$Var_p(\hat{t})$	$Var(\hat{t})$	$\bar{Var}(\hat{t})$
10	0.1	0.1	0.8182	0.9860	0.9940	0.2678	0.1168	0.4414
30				0.8732	0.9061	0.0893	0.1153	0.1247
100				0.8420	0.8250	0.0268	0.0413	0.0289
200				0.8264	0.8546	0.0134	0.0248	0.0151
10	0.1	0.2	0.6667	0.8058	0.8736	0.1778	0.1320	0.3575
30				0.7117	0.7746	0.0593	0.1150	0.0952
100				0.6800	0.7070	0.0178	0.0450	0.0218
200				0.6734	0.7025	0.0089	0.0247	0.0104
10	0.1	0.3	0.5385	0.6518	0.8109	0.1160	0.1840	0.3359
30				0.5750	0.6885	0.0387	0.1421	0.0820
100				0.5493	0.5904	0.0116	0.0387	0.0155
200				0.5739	0.5513	0.0058	0.0127	0.0063
10	0.1	0.4	0.4286	0.5190	0.7042	0.0735	0.2056	0.2798
30				0.4576	0.5699	0.0245	0.1417	0.0620
100				0.4373	0.4668	0.0073	0.0402	0.0103
200				0.4329	0.4355	0.0037	0.0145	0.0041
10	0.1	0.5	0.3333	0.4035	0.7479	0.0444	0.1817	0.2957
30				0.3560	0.4491	0.0148	0.1179	0.0424
100				0.3400	0.3429	0.0044	0.0150	0.0053
200				0.3367	0.3495	0.0022	0.0095	0.0026
10	0.3	0.1	0.8182	0.9860	1.0026	0.2678	0.5781	0.6310
30				0.8732	0.8831	0.0893	0.1121	0.1188
100				0.8420	0.8329	0.0268	0.0394	0.0293
200				0.8264	0.8503	0.0134	0.0136	0.0147
10	0.3	0.2	0.6667	0.8058	0.8005	0.1778	0.2741	0.3649
30				0.7117	0.7038	0.0593	0.1127	0.0809
100				0.6800	0.6781	0.0178	0.0199	0.0192
200				0.6734	0.6786	0.0089	0.0119	0.0092
10	0.3	0.3	0.5385	0.6518	0.7029	0.1160	0.2821	0.3094
30				0.5750	0.5745	0.0387	0.0468	0.0502
100				0.5493	0.5604	0.0116	0.0147	0.0131
200				0.5739	0.5444	0.0058	0.0072	0.0061
10	0.3	0.4	0.4286	0.5190	0.5992	0.0735	0.2792	0.2541
30				0.4576	0.4762	0.0245	0.0537	0.0373
100				0.4373	0.4295	0.0073	0.0095	0.0078
200				0.4329	0.4443	0.0037	0.0056	0.0041
10	0.3	0.5	0.3333	0.4035	0.4121	0.0444	0.1368	0.1221
30				0.3560	0.3816	0.0148	0.0412	0.0248
100				0.3400	0.3432	0.0044	0.0075	0.0050
200				0.3367	0.3557	0.0022	0.0036	0.0026
10	0.5	0.1	0.8182	0.9860	1.2567	0.2678	1.4662	1.2124
30				0.8732	0.8738	0.0893	0.0962	0.1145
100				0.8420	0.8221	0.0268	0.0244	0.0280
200				0.8264	0.8226	0.0134	0.0107	0.0137
10	0.5	0.2	0.6667	0.8058	0.8047	0.1778	0.4549	0.4392
30				0.7117	0.6936	0.0593	0.0668	0.0830
100				0.6800	0.6833	0.0178	0.0188	0.0194
200				0.6734	0.6593	0.0089	0.0091	0.0089
10	0.5	0.3	0.5385	0.6518	0.7768	0.1160	0.4680	0.4267
30				0.5750	0.5422	0.0387	0.0452	0.0452
100				0.5493	0.5602	0.0116	0.0100	0.0129
200				0.5739	0.5241	0.0058	0.0045	0.0056

Continued

Table III - Continuation.

N	p	F	t	$E(\hat{t})$	$\bar{\hat{t}}$	$Var_p(\hat{t})$	$Var(\hat{t})$	$\bar{V\hat{a}r}(\hat{t})$
10	0.5	0.4	0.4286	0.5190	0.4946	0.0735	0.1674	0.1642
30				0.4576	0.4357	0.0245	0.0310	0.0294
100				0.4373	0.4521	0.0073	0.0141	0.0087
200				0.4329	0.4368	0.0037	0.0044	0.0039
10	0.5	0.5	0.3333	0.4035	0.3926	0.0444	0.0931	0.0985
30				0.3560	0.3344	0.0148	0.0192	0.0174
100				0.3400	0.3414	0.0044	0.0069	0.0049
200				0.3367	0.3445	0.0022	0.0030	0.0024
10	0.7	0.1	0.8182	0.9860	1.0438	0.2678	0.5082	0.6370
30				0.8732	0.9100	0.0893	0.1535	0.1307
100				0.8420	0.8359	0.0268	0.0279	0.0290
200				0.8264	0.8034	0.0134	0.0164	0.0132
10	0.7	0.2	0.6667	0.8058	0.8499	0.1778	0.4504	0.4673
30				0.7117	0.6967	0.0593	0.0726	0.0743
100				0.6800	0.6879	0.0178	0.0254	0.0199
200				0.6734	0.6884	0.0089	0.0112	0.0097
10	0.7	0.3	0.5385	0.6518	0.7666	0.1160	0.3068	0.3566
30				0.5750	0.5993	0.0387	0.0628	0.0562
100				0.5493	0.5679	0.0116	0.0142	0.0135
200				0.5739	0.5474	0.0058	0.0070	0.0061
10	0.7	0.4	0.4286	0.5190	0.6220	0.0735	0.2513	0.2543
30				0.4576	0.4516	0.0245	0.0476	0.0335
100				0.4373	0.4278	0.0073	0.0109	0.0078
200				0.4329	0.4340	0.0037	0.0043	0.0039
10	0.7	0.5	0.3333	0.4035	0.4236	0.0444	0.1311	0.1237
30				0.3560	0.3618	0.0148	0.0258	0.0209
100				0.3400	0.3598	0.0044	0.0084	0.0055
200				0.3367	0.3394	0.0022	0.0028	0.0024
10	0.9	0.1	0.8182	0.9860	0.9790	0.2678	0.1256	0.4431
30				0.8732	0.9060	0.0893	0.1298	0.1266
100				0.8420	0.8873	0.0268	0.0476	0.0334
200				0.8264	0.8644	0.0134	0.0187	0.0153
10	0.9	0.2	0.6667	0.8058	0.9174	0.1778	0.1381	0.3913
30				0.7117	0.7830	0.0593	0.1217	0.0978
100				0.6800	0.7561	0.0178	0.0408	0.0245
200				0.6734	0.6792	0.0089	0.0179	0.0096
10	0.9	0.3	0.5385	0.6518	0.8716	0.1160	0.1388	0.3589
30				0.5750	0.7830	0.0387	0.1217	0.0978
100				0.5493	0.7561	0.0116	0.0408	0.0245
200				0.5739	0.6792	0.0058	0.0179	0.0096
10	0.9	0.4	0.4286	0.5190	0.7882	0.0735	0.1856	0.3220
30				0.4576	0.5805	0.0245	0.1366	0.0630
100				0.4373	0.4821	0.0073	0.0413	0.0109
200				0.4329	0.4363	0.0037	0.0164	0.0041
10	0.9	0.5	0.3333	0.4035	0.7433	0.0444	0.1964	0.2988
30				0.3560	0.4491	0.0148	0.1098	0.0414
100				0.3400	0.3720	0.0044	0.0257	0.0066
200				0.3367	0.3327	0.0022	0.0076	0.0024

(*) $E(\hat{t})$: Expected value of $\hat{t}$; $\bar{\hat{t}}$: mean of $\hat{t}$ in the 100 experiments; $Var_p(\hat{t})$: parametric variance of $\hat{t}$; $Var(\hat{t})$: variance of $\hat{t}$ in the 100 experiments; $\bar{V\hat{a}r}(\hat{t})$: mean of the 100 estimates of $Var(\hat{t})$ obtained from $V\hat{a}r(\hat{t})$.

Publication supported by FAPESP.

RESUMO

O trabalho objetivou verificar as propriedades de estimadores do coeficiente de endogamia e da taxa de fertilização cruzada de uma população diplóide. As estimativas foram baseadas na análise de variância de freqüências gênicas de amostras de indivíduos e obtidas pelo método dos momentos. Por se tratar de estimadores que são quocientes entre variáveis aleatórias, utilizaram-se aproximações pelo desenvolvimento de uma função em série de Taylor. Para verificar as expressões deduzidas, analisaram-se dados obtidos de simulação. Os resultados mostraram que os estimadores são tendenciosos, sendo positivo o viés do coeficiente de endogamia e negativo o da taxa de fertilização cruzada. As expressões dessas tendências são função de 1/n, indicando que elas tornam-se desprezíveis com o aumento do tamanho da amostra (n). A simulação confirmou as tendências e mostrou a validade das expressões obtidas para quantificar o erro das estimativas.

REFERENCES

Cockerham, C.C. (1969). Variance of gene frequencies. *Evolution 23*: 72-84.

Fisher, R.A. (1970). *Statistical Methods for Research Workers*. 14th edn. New York, Hafner Press, pp. 362.

Johnson, N.L. and Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions*. 2. John Wiley & Sons, New York, pp. 360.

Kendal, G.M. and Stuart, A. (1963). *The Advanced Theory of Statistics*. Vol. 1. 2nd edn. Charles Griffin & Company Limited, London, pp. 433.

Mood, A.M., Graybill, F.A. and Boes, D. (1974). *Introduction to the Theory of Statistics*. 3rd edn. McGraw-Hill Kogakusha Ltd., Tokyo, pp. 564.

Nei, M. and Chakravarti, A. (1977). Drift variances of F_{st} and G_{st} statistics obtained from a finite number of isolated populations. *Theor. Pop. Biol. 11*: 307-325.

Robertson, A. and Hill, W.G. (1984). Deviations from Hardy-Weinberg proportions: sampling variances and use in estimation of inbreeding coefficients. *Genetics 107*: 703-718.

SAS/GRAPH (1990). *Software: Reference*. Version 6. 1st edn. Vol. 1. NC: SAS Institute Inc., Cary, pp. 794.

Searle, S.R. (1971). *Linear Models*. John Wiley & Sons, New York, pp. 532.

Vencovsky, R. (1992). Análise de variância de freqüências alélicas. In: Congresso Latino Americano de Genética, 10, Rio de Janeiro. *Proceedings, Rev. Bras. Genét. 15* (Suppl. 1): 53-60.

Weir, B.S. (1990). *Genetic Data Analysis: Methods for Discrete Population Genetic*. Sinauer Associates Inc. Publishers, Sunderland, pp. 377.

Weir, B.S. and Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution 38*: 1358-1370.

(Received February 22, 1996)