Heredity (1984), **53**(3), 527–543

© 1984. The Genetical Society of Great Britain

ON THE ESTIMATION OF VIABILITIES IN COMPETITION EXPERIMENTS

J. CONNOLLY AND C. GLIDDON

Statistics Dept., Agricultural Institute, 19 Sandymount Ave., Dublin 4, Ireland,
School of Plant Biology, University College of North Wales, Bangor

Revised 13.ii.84

SUMMARY

The biases in three currently used measures of relative viability are evaluated exactly and compared with previous approximate estimates of the bias. It is shown that the approximate estimate is often seriously misleading for one of the measures. The variances of the three estimators are also evaluated and it is shown that the bias is usually a lesser problem than the variability. The probabilities of the estimators being defined (*i.e.*, having non-zero denominators) are also presented. Exact confidence limits are calculated and attention is drawn to their size. An alternative approximate method based on the logit is discussed.

The experimental method of Anxolabehere *et al.* (1982) gives considerable estimation problems, particularly with low gene frequencies. The large variability of the estimators implies that only relatively large effects can be detected.

1. INTRODUCTION

The existence of problems associated with the estimation, from competition experiments, of relative viabilities has been known for some time. Haldane (1956) gave an approximate value for the bias in relative viability, due to sample size, and suggested an approach for its minimisation. More recently, attention was drawn to the bias found in two commonly used estimators and a new estimator was proposed (Anxolabehere *et al.*, 1982). They also suggested approximate estimates of the biases for two of the measures. In this paper, firstly we calculate the exact expectations and variances of the three estimators considered by Anxolabehere *et al.*, (1982), examine their relative merits and point out an error associated with the approximate approach. Secondly, confidence intervals are derived and considered, together with bias, in terms of their implications for experimental designs for the detection of differences in relative viabilities. Finally, an alternative approach, based on a method of Haldane (1955), is considered.

The experimental context of the present work supposes two independent samples of sizes N_1 and N_2 split into three classes (for example, three genotypes A_1A_1, A_1A_2 and A_2A_2) as in the following scheme:

		Classes (genotypes)		
		1	2	3
Sample	Size	$(A_1A_1$	A_1A_2	$A_2A_2)$
1	N_1	n_{11}	n_{12}	n_{13}
2	N_2	n_{21}	n_{22}	n_{23}

where the n_{ij}'s are outcomes from the two independent trinomial distributions

$$(P_{11}+P_{12}+P_{13})^{N_1}$$

and (1)

$$(P_{21}+P_{22}+P_{23})^{N_2}$$

Interest centres on the properties of a number of estimators of the ratios

$$(P_{12}/P_{11})/(P_{22}/P_{21})$$

and (2)

$$(P_{12}/P_{13})/(P_{22}/P_{23}).$$

Because of symmetry, attention will be confined to estimators of the first of these expressions. Three estimators that have been proposed in the genetic context, to be discussed in detail below, are:

$$W_1=\frac{n_{12}}{n_{11}}.\frac{n_{21}}{n_{22}},$$
$$X_1=\frac{2.(n_{12}+2n_{13})}{(2n_{11}+n_{12})}.\frac{n_{21}}{n_{22}} \tag{3}$$

for which Hardy–Weinberg equilibrium is assumed for the first sample, and

$$H_1=\frac{n_{12}}{(n_{11}+1)}.\frac{n_{21}}{(n_{22}+1)}.$$

These measures were originally proposed by Bundgaard and Christiansen (1972), Polivanov (1964) and Anxolabehere *et al.*, (1982) respectively.

2. Methods

(i) *Evaluation of bias and variance of the estimators*

The first two moments of all three estimators may be exactly evaluated as follows. Each may be written in the form XY where X is a function of the frequencies in the first sample and Y of those in the second (for example, for W_1, $X=n_{12}/n_{11}$ and $Y=n_{21}/n_{22}$). Hence, since the two samples are independent, the expectations and variances may be written (with an obvious notation) as:

$$E(XY)=E(X)E(Y)$$

and (4)

$$V(XY)=V(X)V(Y)+V(X)E(Y)^2+V(Y)E(X)^2.$$

Thus, the expectation and variance of each estimator (3) may be calculated if $E(X)$, $E(Y)$, $V(X)$ and $V(Y)$ are known. These latter quantities are calculated numerically, as described below, for given parameters of the trinomial distributions (1). A computer program was written to evaluate the expectation and variance of specified functions of the cell frequencies for a trinomial distribution of size N. To reduce the effects of rounding error in the calculation of probabilities for large values of N, a recursive formula was used, based on the logarithm of probability. As a check on accuracy, runs were done in single and double precision and the results

differed negligibly. As an additional accuracy check, the calculated expectation for H_1 was compared with the exact expression for the expectation given in Anxolabehere *et al.*, (1982) and again the difference was negligible.

A complication arose in that W_1 is undefined if n_{11} or n_{22} is zero and X_1 is undefined if n_{22} or both n_{11} and n_{12} are zero. This was allowed for by restricting the trinomial distribution to values for which this was not a problem and scaling the remaining probabilities to sum to unity. H_1 is defined for all possible outcomes. The probability of being undefined is examined in more detail below.

(ii) *Confidence intervals for* $(P_{12}/P_{11})/(P_{22}/P_{21})$

The distributions of all three estimators W_1, X_1 and H_1 are skewed to the right for many combinations of gene frequency and sample size. This has been pointed out by Anxolabehere *et al.*, (1982) for W_1. It is also suggested by the fact that all three are bounded below by 0 and above by $(N_1-1)(N_2-1)$, $2(2N_1-1)(N_2-1)$ and N_1N_2 for W_1, X_1 and H_1 respectively, while the mean value is expected to be somewhere about 1 under the null hypothesis of no difference between probabilities for the two trinomials. As will become clear from the results section below, this suggestion of skewness is also supported by the information on the large standard deviations of the estimators. All these pointers suggest that tests and confidence intervals based on the estimated value and its standard error may not be very good, frequently overrepresenting the lower tail, perhaps even giving a lower confidence bound less than 0, and underrepresenting the upper tail. An interval estimation approach based on somewhat different criteria may be more appropriate and one such is now suggested.

For a 2×2 contingency table formed by the classification of two independent samples into two classes, Gart (1962) considered a number of interval estimation approaches for the relative risk, which is defined as the first ratio in formulae (2), except that the row probabilities $P_{11}+P_{12}$ for such a table would sum to unity, and similarly for $P_{21}+P_{22}$. However, the ratios in (2) may be considered as relative risks for 2×2 subtables of the 2×3 array shown earlier (Goodman, 1964*a*). If attention is confined to the first two columns of this 2×3 array then the appropriate 2×2 subtable is

		Classes	
Sample	Size	1	2
1	$n_{11}+n_{12}$	n_{11}	n_{12}
2	$n_{21}+n_{22}$	n_{21}	n_{22}

The corresponding relative risk is the first formula in (2) and the analysis in Gart (1962) may be applied to it. He discusses a method due to Fisher (1935) and Cornfield (1956) for deriving exact confidence limits by the iterative solution of equations. These limits are calculated in the next section for a number of cases, using an algorithm of Thomas (1971), which gives one or two sided intervals as required, with equal probabilities in the tails of the two sided intervals. Baptista and Pike (1977) give an alternative algorithm which produces two sided intervals which differ from those of Thomas (1971) in that the tail area probabilities are not constrained to equal each other, leading to a shorter confidence interval in many cases.

The relative merits of the two procedures are discussed in Mantel and Halperin (1981) and Barnard (1982). For further discussion of confidence limits on the relative risk, see Gart (1966), and Goodman (1964*a*) who discusses the placing of simultaneous confidence limits on more than one relative risk in $R \times C$ (R rows and C columns) two way tables. This may be of relevance if both homozygotes are subject to differential selection and simultaneous confidence limits on both relative risks in (2) are required.

3. Results and discussions

The theory in the previous section was developed in a fairly general context but for the purpose of calculating some specific results a number of simplifying assumptions were made. The competition experiment as described in Anxolabehere *et al.*, (1982) was assumed in the following. The three classes in the 2×3 layout are the numbers of A_1A_1, A_1A_2 and A_2A_2 genotypes in the two samples where P is the A_1 gene frequency under the null hypothesis of no competitive effect in either sample. Examination was restricted to equal sample sizes and Hardy-Weinberg segregation was assumed under the null hypothesis of no competitive effects. The exact mean and standard deviation of W_1, X_1 and H_1 were numerically evaluated for a range of gene frequencies and sample sizes under the null hypothesis. The approximate expectations of W_1 and X_1 as given in Anxolabehere *et al.*, (1982) were also calculated.

(i) *Expectation of estimators*

The expectations for the three estimators are shown in Figs. 1 to 3. There was considerable bias for all three estimators, particularly at low gene frequencies and small sample sizes. The bias was always positive for X_1 and negative for H_1, while for W_1 the bias was negative for lower and positive for higher gene frequencies. For X_1 and H_1, increased sample size reduced the bias. The result with W_1 was more complex, the effect of sample size being to push the peak of the bias profile towards lower gene frequencies, without greatly reducing the maximum positive bias (Fig. 1). Thus, for certain gene frequencies, the bias in W_1 could be greater or of opposite direction for larger, compared with smaller, sample sizes.

Of the three estimators, H_1 was least biased overall, but by no means uniformly so, particularly at low gene frequencies where X_1 and sometimes W_1 were less biased. The negative biases in W_1 and H_1 with low gene frequencies are somewhat worrying in that selection pressure from competition favouring heterozygotes would tend to push the ratio in the same direction, *i.e.*, towards zero.

Also shown in figs. 1 and 2 are the approximate estimates of bias derived from Taylor's series expansions by Anxolbehere *et al.*, (1982). While the approximation was fairly good for X_1, the approximation for W_1 was quite incorrect and even in the wrong direction for low gene frequencies. With greater sample size, the approximation was defective for a smaller range of low gene frequencies. The poor results of the approximate estimates of W_1 are due to the fact that the conditions for expansion of W_1 in a Taylor's series are not met for low gene frequencies and small sample sizes, cf. Appendix I for details.

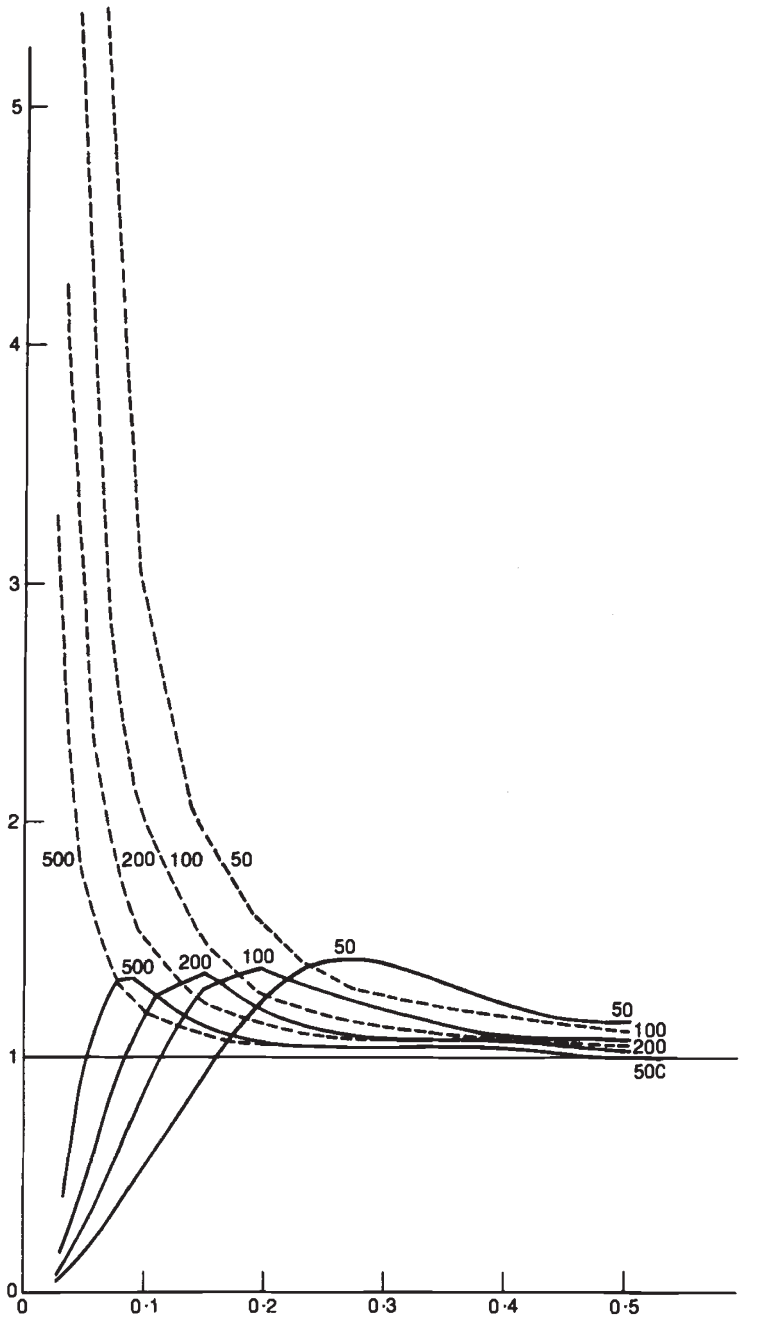

FIG. 1. Exact (——) and approximate (- - - -) expectation of W_1, for various sample sizes and gene frequencies.

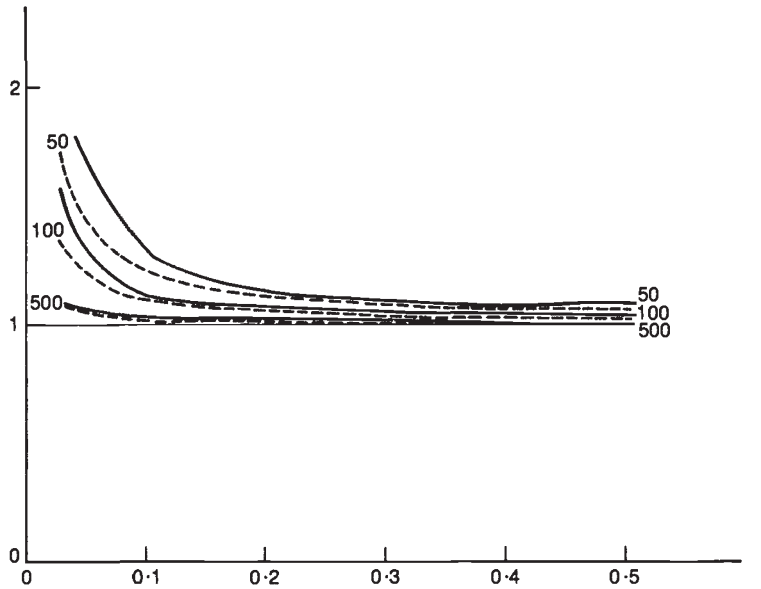

FIG. 2. Exact (——) and approximate (- - - -) expectation of X_1, for various sample sizes and gene frequencies.

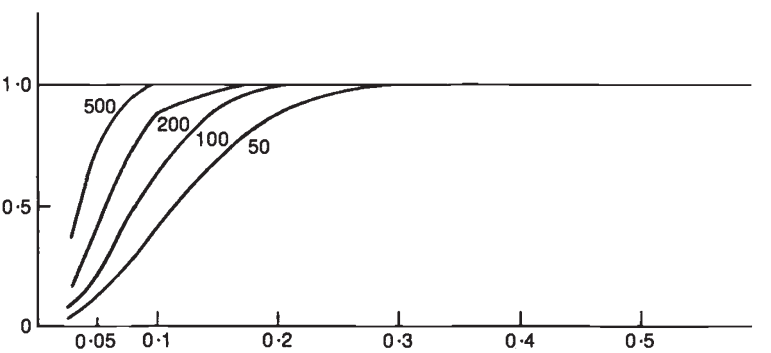

FIG. 3. Expectation of H_1, for various sample sizes and gene frequencies.

(ii) *Standard deviation of estimators*

Perhaps of more concern than the biases in these estimators are their large standard deviations which are shown in figs. 4 to 6 for W_1, X_1 and H_1 respectively. The standard deviation was greater than the bias (and usually much greater) for all three estimators, with gene frequencies in excess of 0·1 in all cases examined. Increasing sample size tended to shift the cases with large standard deviations towards the lower end of the gene frequency range, without greatly reducing this maximum. Neither did increasing sample size consistently reduce the standard deviation of W_1 and H_1, which is a somewhat surprising result.

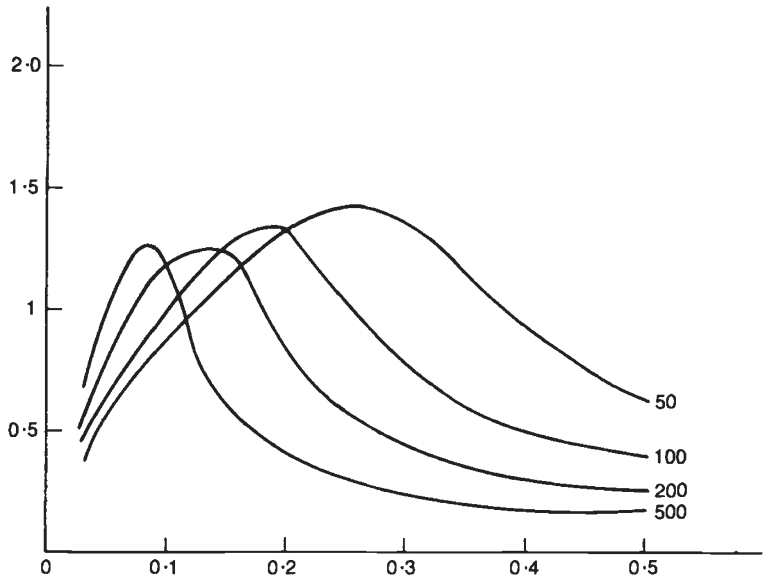

FIG. 4. Standard deviation of W_1, for various sample sizes and gene frequencies.

Again no estimator was uniformly the best. The standard deviation of H_1 was lower than that of W_1 for all cases. That of X_1 was very large for low gene frequency, but less than that of H_1 for higher gene frequencies.

A further difficulty with the estimator W_1 and to a much smaller extent with X_1 is illustrated in fig. 7 which shows the probability of the estimators being defined (*i.e.*, having a non-zero denominator) for a range of sample sizes and gene frequencies. W_1 is undefined and hence of no use in a high proportion of cases with low gene frequency and small sample size. For X_1 the probabilities of being undefined are much lower, but again are greatest for low gene frequency and small sample size. It should be realised that calculating the moments of W_1 conditional on the denominator not being zero, in situations where the probability that it is zero is quite high, is somewhat artificial and may not convey the true inadequacy of the estimator.

(iii) *Confidence limits for the relative risk*

Table 1 gives exact two sided 95 per cent confidence limits with a range of sample sizes and gene frequencies, for several sets of data. For each set of parameters, a number of possible outcomes are given for the two samples, the first of which is the expected number of A_1A_1 and A_1A_2 genotypes under the null hypothesis and assuming Hardy-Weinberg segregation. The other examples were constructed to illustrate the severity of competition required to produce a significant effect, *i.e.*, confidence interval not including unity. This was approached in two ways. Initially, the numbers in the first sample were taken to equal those expected under no competition, while the A_1A_1 genotype numbers in the second sample were reduced until a confidence interval, not including unity, was produced. This did not produce the required interval for the first, second and fourth parameter sets, even

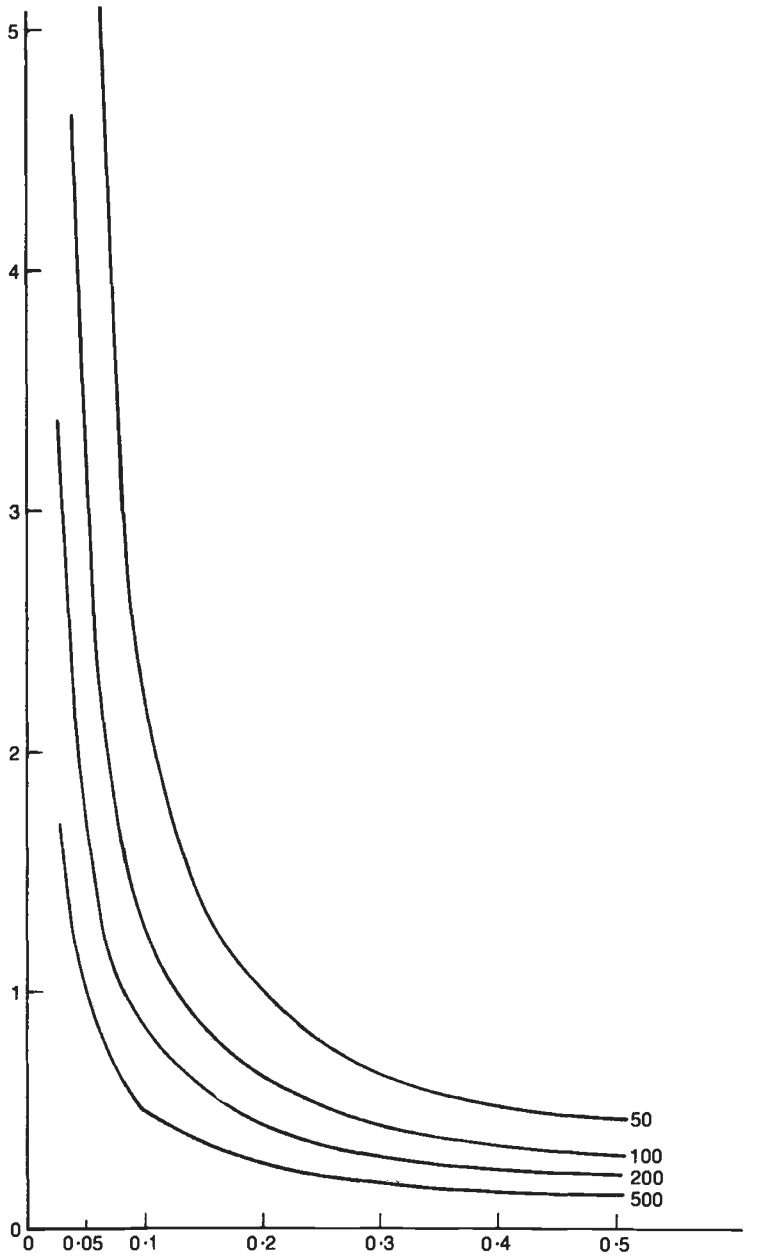

FIG. 5. Standard deviation of X_1, for various sample sizes and gene frequencies.

with the number of the A_1A_1 genotype reduced to 1, the confidence interval still including unity, as is shown in table 1. In these cases the A_1A_1 number was held at 1 in the second sample and the A_1A_1 number in the first sample increased until the confidence interval excluded unity. Results for these cases are also shown in table 1. (The number of the A_1A_1 could have been

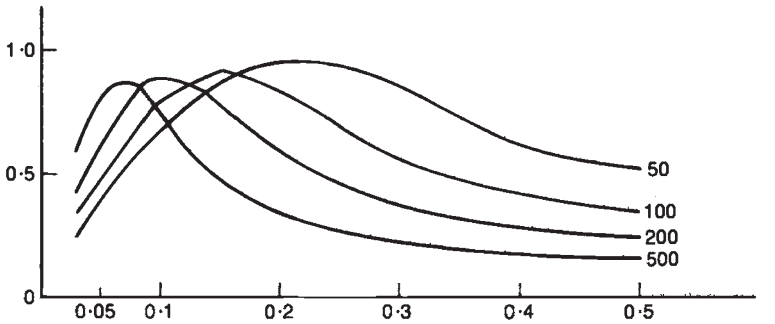

FIG. 6. Standard deviation of H_1, for various sample sizes and gene frequencies.

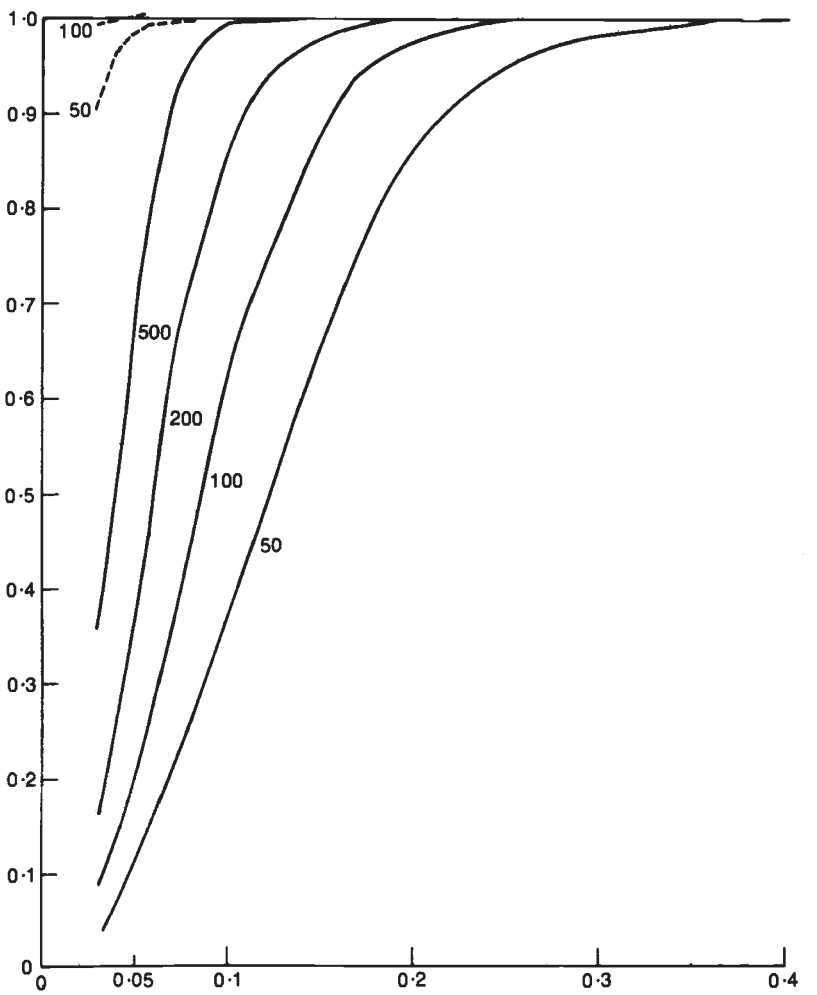

FIG. 7. Probability of W_1 (——) and X_1 (- - - -) being defined, for various sample sizes and gene frequencies.

TABLE 1

Exact and approximate confidence limits (95 percent) on the relative risk for a range of sample sizes and gene frequencies and a selection of genotype frequencies

Sample size	Gene frequency	Genotype frequency Sample 1 A_1A_1 n_{11}	Sample 1 A_1A_2 n_{12}	Sample 2 A_1A_1 n_{21}	Sample 2 A_1A_2 n_{22}	Confidence limits* Exact Lower	Exact Upper	Approx (9) Lower	Approx (9) Upper	Approx (8) Lower	Approx (8) Upper
100	0·1	1	18	1	18	0·012	82·81	0·095	10·52	0·121	8·28
		7	12	1	18	0·002	0·95	0·020	0·90	0·024	0·76
100	0·2	4	32	4	32	0·170	5·87	0·248	4·03	0·259	3·85
		4	32	1	35	0·005	2·52	0·045	2·06	0·054	1·72
		8	28	1	35	0·002	0·84	0·023	0·86	0·028	0·72
100	0·5	25	50	25	50	0·479	2·09	0·510	1·96	0·511	1·96
		25	50	13	62	0·179	0·96	0·201	0·91	0·202	0·91
500	0·1	5	90	5	90	0·222	4·51	0·296	3·38	0·309	3·24
		5	90	1	94	0·004	1·77	0·042	1·62	0·051	1·34
		8	87	1	94	0·003	0·90	0·028	0·95	0·034	0·78
500	0·2	20	160	20	160	0·490	2·04	0·522	1·92	0·525	1·91
		20	160	8	172	0·138	0·92	0·169	0·88	0·171	0·87
500	0·5	125	250	125	250	0·730	1·37	0·739	1·35	0·739	1·35
		125	250	99	276	0·517	0·99	0·525	0·98	0·525	0·98

* These approximate limits are based on the estimates of the variance given in equations (8) and (9).

reduced to 0 rather than 1, as this method allows zero genotype frequencies. Perhaps the approach used gives a clearer idea of the level of compeititon required, in cases of "soft selection", for significance to be established).

These confidence intervals were disappointingly wide, particularly for low gene frequencies, and further emphasise the problem with variability in this type of data. They do show, and allow for, the skewness in the distribution. In fact, in cases with equal genotype frequencies in the samples, the upper limit was the inverse of the lower. The width of these intervals for low gene frequencies was partly a reflection of the low expected numbers of A_1A_1 and A_1A_2 genotypes in this situation, even for quite large sample sizes.

The table also shows that competition would need to be quite severe, particularly with low gene frequencies and sample sizes, to be detected using this method. For example, for a sample size of 500 and gene frequency of 0·2 the expected number in the A_1A_1 class was 20, which had to be reduced to 8 under competition, for the effect to be considered significant. With lower gene frequencies or smaller sample sizes, the degree of competition required for significance was considerably greater. As an example of this, consider the same gene frequency with a sample size of 100. In this case, reducing the A_1A_1 numbers to 1, under competition, was not enough to achieve significance, and it was necessary to increase the number of homozygotes in the control sample to 8, which was twice its expected value.

The results in all this section have been derived under restrictive assumptions on sample size and expected genotype probabilities. However, the formulae given in (4) show that the biases and variances of all three estimators are functions of the biases and variances of their independent components. Hence, even if the trinomials have different parameters from each other (*i.e.*, different sample sizes or different genotype probabilities for the two samples), the estimators will reflect the biases and variability of their components and will have similar properties to those discussed above.

(iv) *An alternative approach*

An alternative approach to the estimation of relative viabilities has been suggested by Haldane (1955). The logarithm of the first relative risk in equation (2) is

$$\log(P_{12}/P_{11}) - \log(P_{22}/P_{21}) \tag{5}$$

which is the difference between two logits. Haldane's method is based on estimation of the difference between the logits and constructing confidence limits for this difference, assuming the estimate to be normally distributed. The exponential of these limits should compare with the exact limits calculated above. Four estimators for the logits in equation (5) were considered by Gart and Zweifel (1967), all of which may be written, for nonzero numerator and denominator, for the first sample, as

$$L_1 = \log \frac{n_{12}+h}{n_{11}+h} \tag{6}$$

and similarly L_2, for the second sample. The estimator with $h=0$ was proposed by Berkson (1953), with modifications for zero values, for $h=0{\cdot}5$

by Haldane (1955) and Anscombe (1956) and $h = 0{\cdot}25$ by Hitchcock (1962). All these estimators are biased to some extent, (Gart and Zweifel, 1967), and none was uniformly best. They concluded that $h = 0{\cdot}5$ was best, on average, for their examples, which differed greatly, however, from those considered here. While the difference $L_1 - L_2$, is unbiased when the null hypothesis of no competition is true, this is not necessarily so otherwise.

The bias discussed here is bias on the log scale, whereas earlier we considered bias, in estimators of the relative risk, on the untransformed scale. A whole family of estimators of the relative risk, on the untransformed scale, may be generated from (6) as

$$R = \frac{n_{12}+h}{n_{11}+h} \cdot \frac{n_{21}+h}{n_{22}+h} \tag{7}$$

for all values of h, and the properties of these examined to give an improved point estimator. One point worth noting is that these estimators are defined for all possible outcomes, zeroes included, for $h > 0$, and even for $h = 0$, with a suitable redefinition, (cf. Gart and Zweifel, 1967). Table 2 gives the

TABLE 2

Expected value and standard deviation of r in equation (7), under the null hypothesis, for three values of h and a range of gene frequencies and sample sizes

Sample size	Gene frequency	Expected values h			Standard deviation h		
		0·25	0·50	1·00	0·25	0·50	1·00
50	0·06	1·90	1·46	1·25	4·03	1·81	0·99
	0·08	2·08	1·53	1·27	3·64	1·80	1·01
	0·10	2·30	1·62	1·30	3·62	1·89	1·06
	0·15	2·57	1·77	1·37	4·08	2·16	1·21
	0·20	2·32	1·71	1·38	4·13	2·20	1·26
	0·30	1·50	1·37	1·26	2·53	1·52	1·03
	0·50	1·14	1·13	1·12	0·62	0·60	0·57
100	0·06	2·02	1·49	1·24	2·88	1·55	0·90
	0·08	2·33	1·62	1·29	3·34	1·78	1·00
	0·10	2·48	1·71	1·33	3·71	1·97	1·10
	0·15	2·17	1·63	1·34	3·82	2·04	1·17
	0·20	1·59	1·40	1·26	2·76	1·60	1·02
	0·30	1·17	1·15	1·14	0·83	0·71	0·63
	0·50	1·06	1·06	1·06	0·37	0·38	0·37
200	0·06	2·33	1·62	1·28	3·26	1·73	0·97
	0·08	2·44	1·70	1·33	3·71	1·96	1·09
	0·10	2·23	1·65	1·33	3·77	2·00	1·13
	0·15	1·47	1·34	1·23	2·39	1·42	0·94
	0·20	1·19	1·17	1·14	0·99	0·78	0·66
	0·30	1·07	1·07	1·07	0·42	0·42	0·40
	0·50	1·03	1·03	1·03	0·26	0·26	0·25
500	0·06	2·28	1·65	1·32	3·69	1·95	1·09
	0·08	1·78	1·47	1·28	3·14	1·73	1·03
	0·10	1·39	1·29	1·20	2·07	1·26	0·86
	0·15	1·11	1·11	1·10	0·59	0·55	0·51
	0·20	1·06	1·06	1·06	0·38	0·37	0·36
	0·30	1·03	1·03	1·03	0·24	0·24	0·24
	0·50	1·01	1·01	1·01	0·16	0·16	0·16

expected values and standard deviations of these estimators, evaluated under the null hypothesis, for three values of h, 0·25 (Hitchcock, 1962), 0·5 (Haldane, 1955 and Anscombe, 1956) and 1·0, and a range of sample sizes and gene frequencies. Both bias and standard deviation were uniformly lower for $h = 1{\cdot}0$ than for the other values. However, there is a danger in examining just the null hypothesis case, since as h gets increasingly greater, the expected value of the estimator automatically tends to unity, and the standard deviation to zero, because of the form of the estimator. Again, as with the other estimators considered earlier, the variability was much greater than the bias for all cases examined, which reduces the importance of bias. The large standard deviations indicate the skewness of the distributions. By comparison with estimator H_1, the bias is smaller at the lower gene frequencies and smaller sample sizes, but not at the higher ones, and the standard deviation is considerably larger at the lower gene frequencies. For a full examination, these estimators would have to be evaluated, with and without competition, and, perhaps (Hitchcock, 1962), for different numerator and denominator values of h, and this is beyond the scope of the current work.

Turning to the variance of the logit estimators, fully nine different variance estimators were considered by Gart and Zweifel. These are primarily of interest in deriving confidence limits for the logit difference (5) above, and in this context, their adequacy in estimating the appropriate variance is important, but so also is the normality assumption. The confidence limits depend on the logit estimator and its estimated variance. With a large number of possible logit estimators from (6), and nine or more variance estimators, the number of possible combinations that could be assessed is very large, and beyond the scope of this work. In any case, since a computer algorithm exists, (Thomas, 1971), for the calculation of exact confidence intervals for the relative risk, the desirability of extensive work on approximations is somewhat debatable.

To get a flavour of the kind of results that arise from these approximate methods, we considered two combinations, the commonly used Haldane estimator of the logit ($h = 0{\cdot}5$), with either the variance estimator suggested by Goodman (1964*b*), as a modification of that proposed by Haldane (1955),

$$V_3 = \frac{(n_{11}+n_{12}+1)(n_{11}+n_{12}+2)}{(n_{11}+n_{12})(n_{11}+1)(n_{12}+1)}. \tag{8}$$

or, alternatively, a variance estimator proposed by Gart (1966),

$$V_4 = \frac{1}{n_{11}+0{\cdot}5} + \frac{1}{n_{12}+0{\cdot}5}. \tag{9}$$

Table 1 gives approximate confidence limits for the relative risk, calculated by using the logit estimator in equation (6), with $h = 0{\cdot}5$, and both of the variance estimators in equations (8) and (9), deriving Normal approximation limits for the logit difference in (5), and taking the exponential of the resulting limits. Gart (1962) using the same estimator of the logit, but yet another variance estimator, concluded that the logit based approach gave limits that were too narrow for a number of examples considered by him. The results here are in agreement with this conclusion, the approximate intervals being narrower than the exact for virtually all cases, with equation

(9) giving somewhat better results than (8). The approximations were worst for small gene frequencies and sample sizes. There may, however, be estimators of the logit, and associated variance estimators, which would lead to a better approximation, on average, to the exact limits.

4. Conclusions

(i) *Bias in estimators*

There are serious biases with all three estimators, particularly for low gene frequency. Of the estimators examined, overall, H_1 was best but not uniformly so, particularly at low gene frequency (p), where X_1 and sometimes W_1 were less biased. The approximate methods of Anxolabehere *et al.*, (1982) for estimating the bias are good for X_1, but are often completely misleading for W_1, giving large corrections that are in the wrong direction for low p. The probability of being able to define W_1 is quite low for low p. The negative bias in W_1 and H_1 for low p values is particularly worrying in that selection pressures would tend to push the ratio in the same direction, *i.e.*, towards zero.

(ii) *Variability of estimators*

Of more concern than bias is the large variability of all three estimators. The standard deviation was greater than the bias (and usually much greater) for all three estimators with gene frequencies in excess of 0·1 in all cases examined. Increasing the sample size shifted the cases with large standard deviations towards smaller gene frequencies. None of the estimators had the uniformly lowest standard deviation. The standard deviation of X_1 was very large for low gene frequencies. The width of the confidence intervals emphasises the problem of variability. It also reflects the considerable positive skewness of the distributions in most cases of interest and hence the inappropriateness of using the standard deviation for significance testing or interval estimation, particularly with low gene frequencies and small sample sizes. A confidence interval approach seems a more appropriate method. Nevertheless, the confidence intervals are disappointingly wide. This partly reflects the low numbers of A_1A_1 and A_1A_2 genotypes from even quite large sample sizes and low gene frequencies. Competition would need to be quite severe, particularly with low gene frequencies and small sample sizes, to be detected.

An alternative approximate approach based on the logit was discussed. Point estimates derived from this would also usually be biased. A large range of possibilities exist for the calculation of approximate confidence limits using this approach, but more detailed study of this is not considered of great importance in the light of the exact method described above.

While previous work (Haldane, 1956; Anxolabehere *et al.*, 1982) rightly called attention to the bias in estimators of differential selection, the problem of bias is generally of less importance than that of the variability of the estimators, and the skewness of their distributions. This bias problem can be partly avoided by emphasising interval rather than point estimation. The experimental method of Anxolabehere *et al.*, (1982) gives considerable estimation problems, particularly with low gene frequencies. The large

variability of the estimators implies that only relatively large effects can be detected. This suggests that, in some cases, experiments might be made with homozygote genotype frequencies higher than might be implied by the rarity of the gene, to ensure a reasonably large number of homozygote individuals on whom competition can act.

Appendix I

Approximate methods for calculating the expected values of W_1 and X_1 have been proposed by Anxolabehere *et al.*, (1982). These are based on Taylor's series expansions (cf. Kendall and Stuart 1958 p 231 sqq) of the expressions in (3). Expectations are taken over the terms of the expansion ignoring third and higher order terms. A simpler method is given below from which it is easier to evaluate where such approximate methods are valid and why in certain cases they are not applicable.

As in the case of the exact evaluation of moments above, it is clear from the independence of the trinomial distributions that rather than consider the expansion of the whole expression at (3), the two components X and Y may be independently expanded and approximate expectations combined as in (4). Both X and Y are ratios of random variables for all three estimators in (3). In general let $X = A/B$ where A and B are random variables. Let $A = E(A) + d_A$ where d_A is the deviation of A from its expected value. Similarly for B. Then

$$X = A/B = (E(A)/E(B))(1 + d_A/E(A))(1 + d_B/E(B))^{-1} \qquad \text{(A1)}$$

This may be expanded in a Taylor's series provided $d_B/E(B) < 1$. This will be almost always the case when the coefficient of variation (CV) of B is small. Further the smaller the CV, the faster the convergence. As a rough guide the following condition may be used as a criterion to indicate whether there is likely to be a problem with the expansion. It is desirable to have

$$2CV(B) = 2SD(B)/E(B) < 1 \qquad \text{(A2)}$$

where $SD(B)$ is the standard deviation of B.

To give some feel for where difficulties may arise in practice, minimum sample sizes for which the denominator variable in the ratios in formulae (3) satisfy this condition were calculated for a range of underlying probabilities of the trinomial. For the sake of placing this in the appropriate genetic context the probabilities are assumed to be those for the three genotypes A_1A_1, A_1A_2 and A_2A_2 with a gene frequency p for A_1 and with segregation following the Hardy-Weinberg rule. A range of gene frequencies is assumed and the results are shown in table 1.

A related calculation is presented in table 2 which shows the minimum value of p, the gene frequency, for which (A2) is satisfied for various sample sizes.

It is clear from these two tables that there may be problems in expanding n_{12}/n_{11} (and consequently W_1) in a Taylor's series for small values of p, the gene frequency. The difficulties increase as the sample size decreases. For X_1 there does not appear to be much of a problem as both it's

TABLE 1

Minimum sample size for which the denominator in the ratios in formulae (3) satisfies condition (A2) for a range of gene frequencies and assuming Hardy-Weinberg segregation

Gene frequency (p)	Denominator				
	n_{11}	n_{22}	$2n_{11}+n_{12}$	$n_{11}+1$	$n_{22}+1$
0·1	396	18	18	1	1
0·2	96	9	8	1	1
0·3	40	6	5	1	1

TABLE 2

Minimum gene frequency for which the denominator in the ratios in formulae (3) satisfies condition (A2) for a range of sample sizes and assuming Hardy-Weinberg segregation

	Denominator				
Sample size (N)	n_{11}	n_{22}	$2n_{11}+n_{12}$	$n_{11}+1$	$n_{22}+1$
50	0·272	0·039	0·039	0	0
100	0·196	0·020	0·020	0	0
200	0·140	0·010	0·010	0	0
500	0·089	0·004	0·004	0	0

components have denominators for which condition (A2) is readily satisfied even at sample sizes as low as 50. Note that this does not mean that convergence will be very fast at such low levels of sample size, and higher order terms in the expansion may not be negligible. If they are ignored this may lead to some bias in the approximation. Both components of H_1 satisfy condition (A2) for all sample sizes and non zero gene frequencies. (This would not be true for a more stringent condition *e.g.*, $3CV<1$). Thus expansions of H_1 are likely to cause no problems and also convergence might be expected to be fairly rapid.

Acknowledgements. We wish to thank Ms Lynne Jones for drawing the diagrams. This work was done while one of the authors, J. Connolly, was on study leave at the University of North Wales, Bangor. Thanks are due to the Agricultural Institute for this leave and to the School of Plant Biology of the University College of North Wales for facilities.

5. REFERENCES

ANSCOMBE, F. J. 1956. On estimating binomial response relations. *Biometrika*, *43*, 461-464

ANXOLABEHERE, D., GOUX, J. M. AND PERIQUET, G 1982. A bias in estimation of viabilities from competition experiments. *Heredity*, *48*, 271-282.

BAPTISTA JOYCE AND PIKE, M. C. 1977. Algorithm AS 115. Exact two sided confidence limits for the odds ratio in a 2×2 table. *Appl. Statist.*, *26*, 214-220.

BARNARD, G. A. 1982. 2×2 tables. *Appl. Statist.*, *31*, 304-305.

BERKSON, J. 1953. A statistically precise and relatively simple method of estimating the bio-assay with quantal response, based on the logistic function. *J. Am. Statist. Ass.*, *48*, 565-599.

BUNDGAARD, J. AND CHRISTIANSEN, F. B. 1972. Dynamics of polymorphisms: I. Selection components in an experimental population of *Drosophila melanogaster*. *Genetics*, *71*, 439-460.

CORNFIELD, J. 1956. A statistical problem arising from retrospective studies. *Proc. Third Berkeley Symp.*, *4*, 135-148.

FISHER, R. A. 1935. The logic of inductive inference. *J. R. Statist. Soc. A*, *98*, 39-54.

GART, J. J. 1962. Approximate confidence limits for the relative risk. *J. R. Statist. Soc. B*, *24*, 454-463.

GART, J. J. 1966. Alternative analysis of contingency tables. *J. R. Statist. Soc B*, *28*, 164-179.

GART, J. J. AND ZWEIFEL, J. R. 1967. On the bias of various estimators of the logit and its variance with application to quantal bioassay. *Biometrika*, *54*, 181-187.

GOODMAN, L. A. 1964a. Simultaneous confidence limits for cross product ratios in contingency tables. *J. R. Statist. Soc. B*, *26*, 86-102.

GOODMAN, L. A. 1964*b*. Interactions in multidimensional contingency tables. *Ann. Math. Statist.*, *35*, 632-646.

HALDANE, J. B. S. 1955. The estimation and significance of the logarithm of a ratio of frequencies. *Ann. Hum. Genet.*, *20*, 309-311

HALDANE, J. B. S. 1956. The estimation of viabilities. *J. Genet.*, *54*, 294-296.

HITCHCOCK, S. E. 1962. A note on the estimation of the parameters of the logistic function using the minimum logit χ^2 method. *Biometrika*. *49*, 250-252.

KENDALL, M. G. AND STUART, A. 1958. *The Advanced Theory of Statistics.*, *Vol. 1* p 433. Charles Griffin and Co. Ltd., London.

MANTEL, N. AND HALPERIN, M. 1981. The Baptista-Pike algorithm (AS 115). *Appl. Statist.*, *30*, 73-74.

POLIVANOV, S. 1964. Selection in experimental populations of *Drosophila melanogaster* with different genetic backgrounds. *Genetics*, *50*, 81-100.

THOMAS, D. G. 1971. Algorithm AS 36. Exact confidence limits for the odds ratio in a 2×2 table. *Appl. Statist.*, *20*, 105-110.