The Goodness-of-Fit Test for Detecting Natural Selection in Random Mating Populations

R. C. Lewontin; C. Clark Cockerham

Evolution, Vol. 13, No. 4. (Dec., 1959), pp. 561-564.

Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28195912%2913%3A4%3C561%3ATGTFDN%3E2.0.CO%3B2-T

Evolution is currently published by Society for the Study of Evolution.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ssevol.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org
Wed Feb 20 17:53:26 2008

NOTES AND COMMENTS

THE GOODNESS–OF–FIT TEST FOR DETECTING NATURAL SELECTION IN RANDOM MATING POPULATIONS

R. C. Lewontin and C. Clark Cockerham

Departments of Genetics and Experimental Statistics, North Carolina State College, Raleigh, N. C.

The issue has recently been raised in the pages of this journal and elsewhere as to the possibility of judging the relative fitnesses of genotypes from the frequencies of these genotypes in populations. Wallace (1958) has provided several numerical examples where severe selection pressures do not result in appreciable deviations of the observed zygotic frequencies from expectation under Hardy-Weinberg equilibrium. Novitski and Dempster (1958) attempted by means of a digital computer to estimate the adaptive values of genotypes in *Drosophila melanogaster* using only the observed genotypic frequencies. They found that the computer produced an infinite variety of best fit values for the fitnesses. We are indebted to them for first calling our attention to this problem and to Dr. Bruce Wallace for urging us to publish our findings.

If the frequency of an allele B in a population is p and that of its alternate allele b is $q = 1 - p$, then following random mating but *before* any natural selection has occurred the zygotes will be in the relative frequencies

$$\begin{array}{lccc} & \mathrm{BB} & \mathrm{Bb} & \mathrm{bb} \\ (1) & p^2 & 2pq & q^2 \end{array}$$

Assume now that natural selection operates in such a way that the three genotypes have the fitnesses W_1, W_2, and W_3. Then after selection the genotypes will be in the relative frequencies

$$\begin{array}{lccc} & \mathrm{BB} & \mathrm{Bb} & \mathrm{bb} \\ (2) & \dfrac{W_1p^2}{\overline{W}} & \dfrac{2W_2pq}{\overline{W}} & \dfrac{W_3q^2}{\overline{W}} \end{array}$$

where $\overline{W} = W_1p^2 + 2W_2pq + W_2p^2$. The frequency of B is no longer p but p′ where

$$(3) \qquad p' = \frac{W_1p^2 + W_2pq}{W_1p^2 + 2W_2pq + W_3q^2}.$$

Suppose, now, that a sample is taken from such a population and that the three genotypes are distinguishable. There will be three observed frequencies, n_1, n_2 and n_3, but only two *independent* observations since there is the linear restriction

$$n_1 + n_2 + n_3 = N_{\mathrm{TOTAL}}.$$

There are, however, four independent parameters, W_1, W_2, W_3 and p. Since the geneticist is usually not interested in the absolute values of the W's but only their relative magnitudes (partly because he does not know how to measure their absolute values), it is possible to reduce the number of independent parameters to be estimated to three, by making W_2 (or W_1 or W_3) equal to unity. However, there is still one more parameter to be estimated than there are available observations, clearly an impossibility.

If estimation of W_1 and W_3 is impossible, is it nevertheless possible to test some hypothesis about these two selection coefficients, or at least to make some judgment about their relative sizes from the observations? Specifically, the geneticist would like to test the hypothesis that $W_1 = W_3 = 1$, that is, that no natural selection is operating. Moreover, if such a hypothesis is rejected, he would, at least, like to know whether W_1 and W_3 are greater or less than unity so that some judgment about the order of fitness of the three genotypes can be made.

The standard procedure presently in use by geneticists is to estimate p by the usual method

$$\hat{p} = \frac{2n_1 + n_2}{2N}$$

and then to test the observed frequencies against the binomial proportions given by (1) with $\hat{p}$ and $\hat{q}$ in place of p and q. If the χ^2 with one degree of freedom is significant it is assumed that natural selection is operating (non-random mating having been ruled out by ancillary information). Moreover, the *directions* of the deviations from expectation are used as a basis for judging the type of selection operating. If, for example, there is an excess of the Bb class with a concomitant deficiency of BB and bb classes, it is assumed that heterozygotes have a greater fitness than homozygotes. A non-significant χ^2 is taken as evidence that $W_1 = W_2 = 1$, that is, that no natural selection is taking place.

What this procedure really amounts to is as follows. First, $\hat{p}$ is an estimate of p′ given by (3) and not of p before selection since no independent estimate of that quantity exists. Second, the expectations calculated are of the form

$$\begin{array}{lccc} & \mathrm{BB} & \mathrm{Bb} & \mathrm{bb} \\ (4) & p'^2 & 2p'q' & q'^2 \end{array}$$

Finally, the χ^2 calculated is a test of whether the sampled distribution (4) is the same as distribution (2). These distributions are the same if simultaneously

$$(5) \qquad p'^2 = \frac{W_1p^2}{\overline{W}}, \quad p'q' = \frac{W_2pq}{\overline{W}},$$

and

$$q'^2 = \frac{W_3q^2}{\overline{W}}.$$

The solution to these three equations is simply

$$(6) \qquad W_2^2 = W_1W_3.$$

Thus, what is really tested is that the product of the adaptive values of the homozygotes is equal to the square of the heterozygote fitness. This is a much less restrictive condition than the requirement that $W_1 = W_2 = W_3$. Adaptive values like

BB	Bb	bb
2	1	.50

give a theoretically perfect fit so that the χ^2 test will be significant only α of the time where α is the probability level chosen for the test. Of course if χ^2 *is* significant, there is evidence for natural selection, a point that should not be lost sight of.

The second question of importance is whether the direction of deviations in the goodness-of-fit test gives some hint as to the sort of selection operating. Specifically, does an excess of the heterozygous class mean that the heterozygote has a superior fitness?

The deviation of the heterozygotes from expected is an estimate of

$$(7) \qquad \frac{2pqW_2}{\overline{W}} - 2p'q'$$

which may be greater than, equal to, or less than zero. These three conditions correspond respectively to an excess of heterozygotes, no deviation, and a deficiency of heterozygotes. Substitution of (3) into (7) yields the relationships:

$W_1W_3 < W_2^2$: excess of heterozygotes
$W_1W_3 = W_2^2$: expected heterozygotes
$W_1W_3 > W_2^2$: deficiency of heterozygotes

Thus, an excess of heterozygotes will appear *whenever the geometric mean of the fitnesses of the homozygotes is less than the fitness of the heterozygote.* This obviously *includes* the case of heterosis but is *not restricted* to it. For example, a recessive lethal will cause a considerable excess of heterozygotes since W_1W_3 will be zero. Some numerical examples of this fact have been presented by Wallace (1958).

An extremely important qualification of these results is that they apply only to a population in which gene frequency is changing. *If the gene frequencies are at equilibrium in the population,* $p' = p$, *and the deviations from the Hardy-Weinberg equilibrium are perfectly adequate for estimating* W_1 *and* W_3 *independently.*

Our previous discussion may give the impression that any attempt to determine the sort of selection operating in nature by comparing observed with theoretical frequencies is hopeless. This is not entirely ture. A consideration of the power of the χ^2 test show that it is relatively insensitive to some sorts of selection while being much more sensitive to others. Specifically, the test is much more powerful in detecting heterotic situations than cases of partial or complete dominance.

The power of a test is the probability that the null hypothesis will be rejected (that the test will be significant), when the true parameters of the population differ from the hypothesized parameters by a given amount. Clearly if the true parameters differ from the hypothesized ones by a very large amount, the test will have great power, while for small deviations the power will be small. In addition, for any given deviation, the power of the test will increase with increasing sample size. For a χ^2 test, the power is determined from the distribution of "non-central χ^2" using a "non-centrality" parameter, λ. The larger λ is, the greater the power of the test. For the case in hand the parameter of non-centrality turns out to be:

$$\lambda = \left[\frac{p^2q^2(W_1W_3 - 1)^2}{(pW_1 + 1)^4(qW_3 + 1)}\right]N$$

where W_2 has been fixed at unity. As expected, λ will increase not only with increasing sample size, N, but also as the product W_1W_3 deviates more and more from unity. Note that when W_1W_3 equals unity, λ is zero and the power of the test is zero.

To illustrate the relationship between the kinds of selection and the power of the χ^2 test to detect them, table 1 has been calculated. It is assumed in this table that $p = q = \frac{1}{2}$ and a-l though the numerical values change for other values of p, the qualitative results are the same. What the table shows is the size of sample necessary to be 90% or 50% sure of detecting different combinations of adaptive values of the two homozygotes ($W_2 = 1$).

It is assumed that the 5% level will be used as a significance criterion in the χ^2 test. Across the top of the table are possible values of W_1, the fitness of BB, and along the left hand side are the values of W_3, the fitness of bb. In the body of the table there are two numbers corresponding to each combination of W_1 and W_3. The roman (lower) numbers are the minimum sample sizes necessary to be 90% sure of detecting selection when the true values of W_1 and W_3 are those given at the borders of the table. The italic (upper) numbers give the minimum sample size for 50% power and obviously will in every case be smaller than their corresponding 90% values.

TABLE 1. *Sample sizes necessary to be 50% sure (italic numbers) and 90% sure (roman numbers) of detecting selection when the true fitnesses are those given at the borders of the table*

	W_1 = fitness of BB										
W_3 = fitness of bb	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.0	*4* 11										
0.1	*5* 13	*6* 16									
0.2	*5* 15	*7* 19	*8* 23								
0.3	*7* 18	*8* 23	*11* 29	*13* 36							
0.4	*8* 21	*10* 27	*13* 35	*17* 45	*21* 57						
0.5	*9* 24	*12* 32	*15* 42	*20* 55	*27* 73	*35* 95					
0.6	*10* 27	*13* 37	*18* 50	*25* 67	*33* 91	*45* 124	*61* 168				
0.7	*11* 30	*15* 42	*22* 59	*30* 82	*42* 116	*59* 162	*85* 231	*123* 337			
0.8	*12* 34	*18* 48	*25* 69	*36* 100	*53* 145	*78* 212	*118* 323	*186* 508	*311* 851		
0.9	*14* 38	*20* 56	*30* 81	*44* 120	*66* 182	*103* 282	*168* 459	*293* 801	573 1567	*1387* 3794	
1.0	*15* 42	*23* 63	*35* 95	*53* 145	*84* 229	*138* 378	*246* 673	*493* 1349	*1244* 3405	*5546* 15,175	— —
1.2	*18* 50	*29* 80	*46* 126	*77* 210	*134* 368	*262* 716	*607* 1660	*2099* 5742	*37,646* 102,999	*10,486* 28,691	*1859* 5086
1.4	*22* 61	*36* 99	*61* 168	*111* 304	*224* 614	*553* 1513	*2212* 6053	*159,847* 437,343	*4978* 13,620	*1181* 3233	*553* 151
1.6	*26* 71	*45* 122	*81* 222	*162* 443	*393* 1074	*1460* 3996	*41,544* 113,665	*5210* 14,254	*1072* 2934	*484* 1325	*2893* 790
1.8	*30* 82	*54* 149	*106* 290	*240* 658	*753* 2060	*6776* 18,538	*12,045* 32,956	*1288* 3523	*504* 1378	*283* 773	*188* 515
2.0	*35* 95	*65* 179	*138* 378	*365* 999	*1694* 4634	— —	*2212* 6053	*625* 1709	*311* 851	*195* 534	*138* 378
2.2	*39* 108	*78* 214	*181* 494	*575* 1573	*5353* 14,647	*8850* 24,213	*983* 2690	*443* 1212	*221* 604	*148* 404	*109* 299
2.4	*44* 121	*93* 255	*236* 647	*957* 2619	*54,910* 150,234	*2497* 6833	*587* 1606	*277* 759	*170* 465	*119* 326	*91* 248
2.6	*50* 136	*110* 301	*311* 851	*1738* 4756	*60,980* 166,841	*1245* 3405	*406* 1112	*214* 585	*138* 378	*100* 274	*78* 213
2.8	*56* 152	*129* 354	*412* 1129	*3662* 10,018	*7549* 20,655	*780* 2134	*307* 840	*174* 476	*117* 320	*87* 237	*68* 187
3.0	*61* 168	*152* 416	*553* 1513	*10,386* 28,416	*3011* 8239	*553* 1513	*246* 673	*147* 401	*101* 277	*77* 210	*61* 168
4.0	*96* 263	*323* 884	*3457* 9458	*4057* 11,100	*523* 2431	*216* 591	*125* 343	*86* 234	*64* 176	*51* 140	*43* 117
10.0	*465* 1272	— —	*669* 1831	*196* 537	*101* 276	*65* 179	*48* 130	*37* 102	*31* 84	*26* 71	*23* 63

The table is divided into three sections by heavy lines. The diagonal stepped line running from the lower left hand corner to the right middle margin separates combinations of W_1 and W_3 which give *excesses* of heterozygotes (above the line) from those combinations giving *deficiencies* of heterozygotes (below the line). The upper area is then subdivided by a horizontal line into an upper region where heterosis obtains, and a lower one in which there is no heterosis. It is this second region which is so troublesome from the standpoint of interpretation of the observations. The worst situation occurs in the three cells containing dashes because here, despite intense selection, there is no observable deviation from expected proportions so that no sample size of whatever magnitude can detect the operation of selection. These are the cases where W_1 is the reciprocal of W_3. The important point about the table is the way in which the sample sizes change with fitness. The upper left hand corner of the table, representing a balanced lethal situation, requires the smallest sample size for detection. Progressing downward and to the right in the table the necessary sample sizes increase until the diagonal is reached, whereupon N begins to decrease. In the case of heterozygote excess it is easier to detect cases of heterosis (above the horizontal line) than those of partial dominance. The χ^2 test is more sensitive to heterotic selection than to other types and is completely incapable of detecting some extreme cases of partial or complete dominance.

It should not be supposed that if the χ^2 test is significant with a sample size of 10, say, that heterosis is demonstrated simply because $N = 10$ happens to fall above the horizontal line in the table. *Any* type of selection may give a significant χ^2 with *any* sample size. As a matter of fact even when no selection is operating χ^2 will be significant 5% of the time if this is the significance level chosen. All that the table shows is that for a given sample size heterotic situations will be more often detected than non-heterotic ones.

LITERATURE CITED

Novitski, E., and E. R. Dempster. 1958. An analysis of data from laboratory populations of *D. melanogaster*. Genetics, **43**: 470–479.

Wallace, Bruce. 1958. The comparison of observed and calculated zygotic distributions. Evolution, **12**: 113–115.

NOTES ON RELATIVE FITNESS OF GENOTYPES THAT FORMS A GEOMETRIC PROGRESSION.

C. C. Li

Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania

Let the three positive numbers W_1, W_2, W_3 denote the relative selective value (or fitness) of the three genotypes BB, Bb, bb, respectively. In a random mating population the frequencies of the three genotypes are in simple binomial proportions (p^2, $2pq$, q^2) at birth or, more generally, before selection takes place. In the preceding paper Lewontin and Cockerham (1959) have shown that if $W_2^2 = W_1W_3$, the genotypic frequencies *after* the operation of selection in the population will again be in simple binomial proportions (p'^2, $2p'q'$, q'^2) where p' and q' are the new gene frequencies after selection. This fact yields the important conclusion that under such a selective scheme, the observed adult genotype and gene frequencies—presumably those after the operation of selection—will tell us nothing about the value of the W's. The relation $W_2^2 = W_1W_3$ simply means that the three numbers W_1, W_2, W_3 constitute a geometric series. To the author's knowledge, Dempster (1955) is the first who considers such a set of fitness values, taking the W's to be $1, (1\text{-}s), (1\text{-}s)^2$ in studying the problem of balance between mutation and selection. The purposes of this note are to supply a model for the genesis of the geometric selective scheme; to point out some simple properties of such a type of selection; and to extend the results to more general cases.

1. First of all, let us review briefly the main result of Lewontin and Cockerham. For their purposes, they find it convenient to make $W_2 = 1$. For our present purpose, however, it is more convenient to take $W_1 = 1$, $W_2 = r$, $W_3 = r^2$, where r may be greater or smaller than unity. When $r = 1$, there is no selection operating. The procedure of analysis and the selection result may be summarized in a tabular form (see table 1). The columns of table 1 are numbered to correspond with the serial number of the expressions in the preceding paper. The only discrepancy is that the common denominator $\overline{W}$ of their expression (2) has been put at the bottom of our column (2). The imposed condition (5) is to

LINKED CITATIONS

The Goodness-of-Fit Test for Detecting Natural Selection in Random Mating Populations

R. C. Lewontin; C. Clark Cockerham

Evolution, Vol. 13, No. 4. (Dec., 1959), pp. 561-564.

Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28195912%2913%3A4%3C561%3ATGTFDN%3E2.0.CO%3B2-T

Literature Cited

The Comparison of Observed and Calculated Zygotic Distributions
Bruce Wallace
Evolution, Vol. 12, No. 1. (Mar., 1958), pp. 113-115.
Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28195803%2912%3A1%3C113%3ATCOOAC%3E2.0.CO%3B2-3