A CRITICAL EXAMINATION OF THE USE OF THE JOLICOEUR-BRUNEL DIAGRAMS TO DEMONSTRATE SELECTION IN BIOLOGICAL POPULATIONS

DAVID WOOL

Department of Zoology, Tel-Aviv University, Tel-Aviv, Israel

Received September 21, 1971

ABSTRACT

The use of the JOLICOEUR-BRUNEL hexagonal diagrams to demonstrate selection in biological populations was tested in a computer-simulation study.—The results show convincingly that the probabilities associated with the six deviation patterns illustrated in the hexagonal diagrams are not equal, but depend on the sample size and in particular on the genetic composition of the sampled population. The diagrams thus cannot be used to demonstrate selection, but only to illustrate trends of deviation patterns, which may be real biologically.

Jolicoeur and Brunel (1966) have suggested a graphic technique to illustrate deviation patterns in trinominal populations. A deviation between two points plotted in the triangular coordinates of a de Finetti diagram (LI, 1955) is represented by a single point, plotted in a hexagon. The center of the hexagon represents zero deviations in all directions. The six component equilateral triangles represent the six possible deviation patterns, and the perpendiculars from the plotted point to the three diagonals of the hexagon represent the direction and magnitude of the deviation with respect to each of the three components of the population (Figure 1).

SOKAL and SONLEITNER (1968) have used this method to demonstrate selection in hybrid populations of *Tribolium castaneum* Herbst by plotting changes, for successive generations or life history stages, in the zygotic frequencies of $++$, $+b$ and bb (where b is the semidominant locus, black). They applied a test of significance, based on the multinominal distribution of the plotted points in the hexagons, the null hypothesis being that in the absence of selection, the six deviation patterns were equally likely. The test was applied by the present author to his own set of data (1969). However, a reviewer of the manuscript of that study questioned the null hypothesis and demonstrated that for samples of size 3, the six deviation patterns are *not* equally likely. He showed that the probability of each depends on the genetic composition of the population and on the sample size.

The probability of each deviation-pattern is in fact $1/6$ if one tests only the null hypothesis that any *points* are randomly distributed in the hexagon. However, when the points represent differences in zygotic frequencies, the probabilities may not be equal. This question was tested by a computer simulation study

FIGURE 1.—An example of the JOLICOEUR-BRUNEL diagram, fitted to plot the deviations (in proportions) of three genotypes: $++$, $+b$, and bb , in samples of *Tribolium castaneum*, from the Hardy-Weinberg expectations. A sample may deviate from expectation with respect to each of the three genotypes represented by the diagonals, by containing more (Pos.= Positive) or less (Neg.= Negative) of that genotype than expected. The six possible deviation patterns are represented by the six component equilateral triangles in the hexagon. The sequence of $+$ and $-$ symbols identifies a deviation pattern and should be read from left to right (Thus $+-+$ means more $++$, less $+b$, and more bb than expected).

employing sample sizes of $N = 10$, $N = 100$ and $N = 1000$.

METHODS

The simulation program was written for the CDC-3400 computer at the Tel-Aviv University computing center.

A subprogram generates a sample of N random numbers ranging from 0 to 1, each representing one individual. The genotype of each individual is decided upon by comparing the random number with the given proportions of the three genotypes in the sampled population. The main line then compares the frequency of each genotype in the sample with the theoretical expectation (assuming no selection), and identifies the pattern of deviations from expectation. After a specified number of samples is obtained in this manner, the frequency of each deviation pattern is recorded and printed. A special counter is set up for the cases in which the observed frequency of one of the genotypes exactly equals the expected one (i.e., the point lies on one of the diagonals). These cases were omitted from the calculations.

The observed frequencies of the six patterns of deviation were tested for a significant departure from the 1:1:1:1:1:1 expectation by chi-square tests.

The proportion of cases in each class was compared, within each population composition—among sample sizes, and within each sample size, among the different population compositions. Departures from equality were tested for significance by the method given in SNEDECOR (1956, p. 227).

FIGURE 2.—The frequency of the six deviation patterns (in percent of total) in random samples of $N =$ (A) 1000, (B) 100, (C) 10 organisms, taken from populations of the genetic composition $++ : +b : bb =$ (1) $1:1:1$, (2) $9:6:1$, (3) $9:1:6$. The dashed line across each histogram indicates the expected proportions on H_0 ($P = 16.67\%$ in each class). Roman numerals correspond to markings in Figure 1 and identify the deviation patterns.

RESULTS

Some of the results are illustrated in Figure 2, arranged to demonstrate the effects of both the genotypic composition of the sampled population and the sample size. In each figure, the proportion of samples in the six deviation-pattern classes is shown for $N = 1000$ (A), $N = 100$ (B), and $N = 10$ (C).

The distribution of 2000 samples of size 1000 taken from a population in which the proportions of the three genotypes were equal ($1:1:1$) did not differ significantly from the $1:1:1:1:1:1$ expectation (Figure 2—1A).

Reducing the sample size to $N = 100$ increased the difference in frequency between the classes somewhat, but the departure from the all-equal expectation was still not significant when 6000 such samples were generated (Figure 2—1B). But with $N = 10$ (Figure 2—1C), highly significant differences from the all-equal expectations were found ($P < 0.001$).

The situation was entirely different when the genetic composition of the population was other than $1:1:1$. The distribution of 2731 samples of size 1000 from a population in which the genetic composition was $++ : +b : bb = 9:6:1$

is illustrated in Figure 2—2A, together with 2900 samples of size 100 (Figure 2—2B) and 3000 of size 10 (Figure 2—2C). The differences from the all-equal expectation was very highly significant using all sample sizes ($P < 0.001$) by the chi-square test. It is clear from the Figures that two classes (patterns), namely $++-$ and $-+-$, are relatively rare.

A population with the composition 6 : 9 : 1 yielded, when sampled, a very similar distribution. The "mirror image" population, 1 : 6 : 9, yielded again two classes less frequent than the rest, but these were $+-+$ and $-++$, which were very frequent in the previous example. Distributions in these cases were all significantly different from the all-equal hypothesis ($P < 0.001$).

Figure 2—3(a, b, c) illustrates the distribution of samples from a population in which the composition was 9 : 1 : 6. The two least frequent classes, using any of the three sample sizes, were $+-+$ and $-+-$. A similar distribution was obtained from a mirror image population with the composition 6 : 1 : 9.

With the exception of $N = 100$ and $N = 1000$ with the composition 1 : 1 : 1, the proportions of cases in each class were always not the same when different sample sizes were used ($P < 0.05$ by the SNEDECOR test), within any population composition.

DISCUSSION

The results convincingly illustrate that the relative frequency of each of the six deviation patterns is a function both of the genetic composition of the sampled population and of the sample size.

The two least-frequent deviation patterns, with sample sizes employed here, appear to be those in which both most-frequent genotypes deviate in one direction, and the least-frequent genotype in the opposite direction. But this does not have to be true for other sample sizes. Only in the case of a 1 : 1 : 1 population, from which equal and large (> 100) samples are taken, may the all-equal null hypothesis perhaps be considered valid.

Since generally one does not know the true genetic composition of the population he samples, the test of significance in its present form cannot be used to demonstrate selection in biological populations. It was duly omitted from the published version of the present author's study (WOOL 1970).

However, the hexagonal graphs may still be effectively used to illustrate trends in deviation patterns, as in SOKAL and SONLEITNER (1968), since these trends may be real biologically.

First and foremost I must thank the reviewer of my paper in *Genetics*, who pointed out the error to me. I am grateful to Dr. M. BROWN of this University, and to Drs. R. R. SOKAL and F. J. ROHLF of the State University of New York at Stony Brook for their criticism of the original manuscript.

LITERATURE CITED

- JOLICOEUR, P. and P. BRUNEL, 1966 Application de diagramme hexagonal à l'étude de la sélection de ses proies par la morue. *Vie et Milieu* 17: 419-433.

- LI, C. C., 1955 *Population Genetics*. Univ. of Chicago Press.
- SNEDECOR, G. W., 1956 *Statistical Methods*. Iowa State College Press.
- SOKAL, R. R. and F. J. SONLEITNER, 1968 The ecology of selection in hybrid populations of *Tribolium castaneum*. *Ecol. Monogr.* **38**: 345-379.
- WOOL, D., 1969 Deviations of zygotic frequencies from expectations in eggs of *Tribolium castaneum*. Ph.D. Thesis, Univ. of Kansas, Lawrence, Kansas. —, 1970 Deviations of zygotic frequencies from expectations in eggs of *Tribolium castaneum*. *Genetics* **66**: 115-132.