Theoretical and Applied Genetics 42, 275–276 (1972)

A Note on Jana's Paper: Simulation of Quantitative Characters from Qualitatively Acting Genes

J. LEE and P. J. KALTSIKES

Plant Science Department, University of Manitoba, Winnipeg, Manitoba (Canada)

Jana (1971) proposed a matrix method for estimating 3^n genetic parameters in a quantitative character governed by n loci in the case where the individual genotypes can be identified. The author employed a system of exact solution

$$\hat{X} = A^{-1} y$$

where $\hat{X}$ is a vector of the 3^n parameter estimates, A is a matrix of functional coefficients derived from the genotypic expectations given by Seyffert (1966), and y is an observed vector of the 3^n phenotypic means each with a unique variance.

The method of exact solutions in Jana's context can be shown to be a special case [when (1) the number of parameters to be estimated is equal to the number of equations and (2) there is no linear dependency in that system of equations] of the least squares method of estimation of genetic parameters used by Hayman (1958 and later) and summarized in Mather and Jinks (1971). The identity of the two methods is shown as follows:

$$\begin{aligned}\hat{X} &= (A'\,A)^{-1}\,A'\,Y \quad \text{(least squares method)}\\ &= A^{-1}\,(A')^{-1}\,A'\,Y\\ &= A^{-1}\,I\,Y\\ &= A^{-1}\,Y \qquad\qquad \text{(exact solution)}\end{aligned}$$

Of course, the variance of the elements in vector X is zero if and only if the elements in vector Y is not subjected to error; that is, $var\,(y_i) = var\,(y_j) = 0$.

In a replicated genetic experiment, it is well-known that a portion of the phenotypic variance is due to the varying effects of the environmental complex exerted on a particular genotype. Similarly, the variance of the phenotypic means for the 3^n genotypes, each replicated a number of times, can never be equal due to the varying environmental effects. Consequently, the variance of the 3^n phenotypic means are invariably heterogeneous. Jana (1971) did not consider this factor in his estimate of the genetic parameters; therefore, his method is equivalent to the unweighted least squares analysis. If the variances of the 3^n phenotypic means are heterogeneous and if the unweighted least squares analysis were used, then the estimates of the genetic parameters, although unbiased, would not have minimum variance (Kendall and Stuart, 1961; Draper and Smith, 1966).

We therefore proposed that the weighted least squares analysis should be employed for the estimation of genetic parameters. According to this procedure, $\hat{X} = (A'\,E^{-1}\,A)^{-1}\,(A'\,E^{-1}\,Y)$, where E is the variance-covariance matrix of Y. When there is no correlation between genotypes, E reduces to a diagonal matrix whose elements are the variances of the 3^n phenotypic means. The variances of the elements in $\hat{X}$ is $(A'\,E^{-1}\,A)^{-1}$. The chi-square test for fitting the elements in X is $(Y'\,E^{-1}\,Y) - (Y'\,E^{-1}\,A\,\hat{X})$.

Estimation of genetic parameters using the weighted least squares approach has the following advantages:

(1) The estimators of the genetic parameters so obtained would be unbiased and also have minimum variance;

(2) The method applies whether the number of parameters to be fitted is equal to or less than the number of equations in the system. Practically, this is advantageous since one general procedure can be programmed for computer analysis;

(3) An investigator may not want to fit all of the 3^n genetic parameters in the model, thus leaving some degrees of freedom for the chi-square test for the adequacy of the model (Hayman, 1958; Mather and Jinks, 1971). Parenthetically, if all 3^n genetic parameters were fitted and if some of the genetic parameters were in fact not significant from zero, then the estimates of the 3^n genetic parameters will again not have minimum variance.

We have written a general purpose weighted/unweighted least squares analysis computer program (Lee and Kaltsikes, 1971) to perform all of the aforementioned procedures, including the chi-square testing of the goodness-of-fit of the model. The program can be used to estimate genetic parameters from (1) identifiable genotypes such as the 3^n genotypes derived from isogenic lines, (2) a set of generation means derived from crossing two inbred lines, or (3) a set of generation variances and covariances. The

program, which is written in FORTRAN IV, can be obtained from us upon request.

References

1. Draper, N. R., Smith, H.: Applied Regression Analysis. New York: Wiley 1966. — 2. Hayman, B. I.: The separation of epistatic from additive and dominance variation in generation means. Heredity **12**, 371—390 (1958). — 3. Jana, S.: Simulation of quantitative characters from qualitatively acting genes. I. Nonallelic gene interaction involving two or three loci. Theor. Appl. Gen. **41**, 216—226 (1971). — 4. Kendall, M. G., Stuart, A.: The Advanced Theory of Statistics, Volume 2, Inference and Relationship. London: Griffin 1961. — 5. Lee, J., Kaltsikes, P. J.: A computer program for the weighted/unweighted least squares analysis. Crop Science **11**, 314—315 (1971). — 6. Mather, K., Jinks, J. L.: Biometrical Genetics, 2nd Ed. London: Chapman and Hall 1971. — 7. Seyffert, W.: Die Simulation quantitativer Merkmale durch Gene mit biochemisch definierbarer Wirkung. I. Ein einfaches Modell. Züchter **36**, 159—163 (1966).

Received June 19, 1972

Communicated by W. Seyffert

J. Lee
Dr. P. J. Kaltsikes
Plant Science Department
University of Manitoba
Winnipeg, Manitoba R3T 2N2 (Canada)