Faculty Papers and Publications in Animal Science

Animal Science Department

January 1960

THE NUMBER OF DAUGHTER-DAM PAIRS NEEDED FOR ESTIMATING HERITABILITY

L. Dale Van Vleck
University of Nebraska-Lincoln, dvan-vleck1@unl.edu

S. R. Searle
New Zealand Dairy Board, Wellington, N. Z.

C. R. Henderson
Cornell University

Follow this and additional works at: http://digitalcommons.unl.edu/animalscifacpub

Part of the Animal Sciences Commons

Van Vleck, L. D., S. R. Searle, and C. R. Henderson. 1960. The number of
     daughter-dam pairs needed for estimating heritability. *Journal of Animal
     Science* 19:916–920.

Abstract: A formula for estimating the number of daughter-dam pairs needed
     when estimating heritability by daughter on dam regression in the one-
     way classification analysis for fixed values of the average number of pairs
     per class, of the ratio of daughter to dam variance, of the anticipated
     heritability estimate, and of the desired variance of the estimate has
     been presented. The effect on N of changes in these four variables has
     been illustrated. This procedure was extended to multiple classification
     analysis and was also used to determine the number of pairs needed to
     estimate repeatability by regression of second records on first records.

THE NUMBER OF DAUGHTER-DAM PAIRS NEEDED FOR ESTIMATING HERITABILITY

L. D. VanVleck, S. R. Searle [1] and C. R. Henderson

Cornell University,[2] Ithaca, New York

THE number of observations required to obtain reliable estimates of heritability has been discussed recently for the method which utilizes the intra-class correlation coefficient. Wearden (1959) has made use of the power function to determine the number of progeny per sire needed to estimate heritability. Tallis (1959), utilizing the formula for the variance of the intra-class correlation coefficient as given by Fisher (1952), has discussed the minimum number of offspring necessary for estimating heritability with a given accuracy.

Another method of estimating heritability is from the regression of daughter records on dam records. These estimates are subject to sampling variances and many daughter-dam pairs must be used in order to obtain accurate estimates of heritability. It is, therefore, useful to know the number of pairs needed to estimate heritability with a pre-set standard error of the estimate.

In the one-way classification analysis (where classes are often herds or sires) the estimated variance of the regression coefficient (b) is

$$V(b) = \frac{\Sigma y^2 - \frac{(\Sigma xy)^2}{\Sigma x^2}}{(N-S-1)\Sigma x^2}, \tag{1}$$

where $\Sigma y^2 =$ the within class sum of squares of the dependent variable (daughter records),

$\Sigma x^2 =$ the within class sum of squares of the independent variable (dam records),

$\Sigma xy =$ the within class sum of crossproducts of daughter and dam records,

$N =$ total number of daughter-dam pairs, and

$S =$ total number of classes.

Dividing both the numerator and denominator by Σx^2 we have:

$$V(b) = \frac{\frac{\Sigma y^2}{\Sigma x^2} - \frac{(\Sigma xy)^2}{(\Sigma x^2)^2}}{N-S-1}.$$

Since heritability, G, is estimated as $2b = \frac{2(\Sigma xy)}{(\Sigma x^2)}$ when assuming no

relationship between daughters within classes, $V(2b)$ can be expressed as

$$4V(b) = V(G) = 4\left(\frac{\frac{\Sigma y^2}{\Sigma x^2} - \frac{G^2}{4}}{N-S-1} \right).$$

Dividing both numerator and denominator of $\frac{\Sigma y^2}{\Sigma x^2}$ by $N-S$, we have

$$V(G) = 4\left(\frac{\frac{\hat{\sigma}_y^2}{\hat{\sigma}_x^2} - \frac{G^2}{4}}{N-S-1} \right), \tag{2}$$

where $\hat{\sigma}_y^2 =$ the estimated error variance of daughter records, and
$\hat{\sigma}_x^2 =$ the estimated error variance of dam records.

For convenience let us set $X = \frac{\hat{\sigma}_y^2}{\hat{\sigma}_x^2}$. Now if we let $R = \frac{N}{S} =$ average number
of pairs per class and solve equation (2) for N we obtain

$$N = \frac{(R)}{(R-1)} \frac{[4X - G^2 + V(G)]}{V(G)}. \tag{3}$$

When heritability is estimated, the ratio of variances, X, usually approximates to 1; R is known or estimated; and $V(G)$ is fixed in order to obtain a certain confidence interval. Then, for an anticipated value of G, the number of daughter-dam pairs required in order to estimate G with a predetermined standard error may be estimated. It can be seen that N increases with an increase in X, but decreases as R, G, and $V(G)$ increase.

Assuming $X=1$, a quick approximation to N can be given as $N = \frac{(R)}{(R-1)} \frac{4}{V(G)}$. Setting $G=0$ will tend to overestimate the number required.

For $R \geqslant 12$ and $X=1$, N can be approximated by $\frac{4}{V(G)}$ for anticipated values of $G \leqslant .50$.

Figure 1 can be used to estimate the number of daughter-dam pairs needed if it can be assumed that the ratio of daughter to dam variances is 1. The values plotted are for a desired standard error of the heritability estimate of 0.04. Values for multiples (p) of 0.04 can be obtained by multiplying the value taken from the graph by $\frac{1}{p^2}$. Since $G \pm 2[V(G)]^{\frac{1}{2}}$ is the approximate 95 percent confidence interval of estimated G, the number of pairs required for obtaining estimates of $G \pm 2[V(G)]^{\frac{1}{2}}$ can also be obtained from figure 1. It should be observed that doubling the confidence interval desired reduces the number of pairs required by a factor of 4.

As an example, suppose that the ratio of variances is 1, the average number of pairs per herd is 4, the heritability of the trait is thought to be 0.4, and that a confidence interval of $\pm 2(.12)$ is desired. Now $0.12 = 3(.04)$ and therefore $p=3$. From the graph we find for $R=4$ and heritability
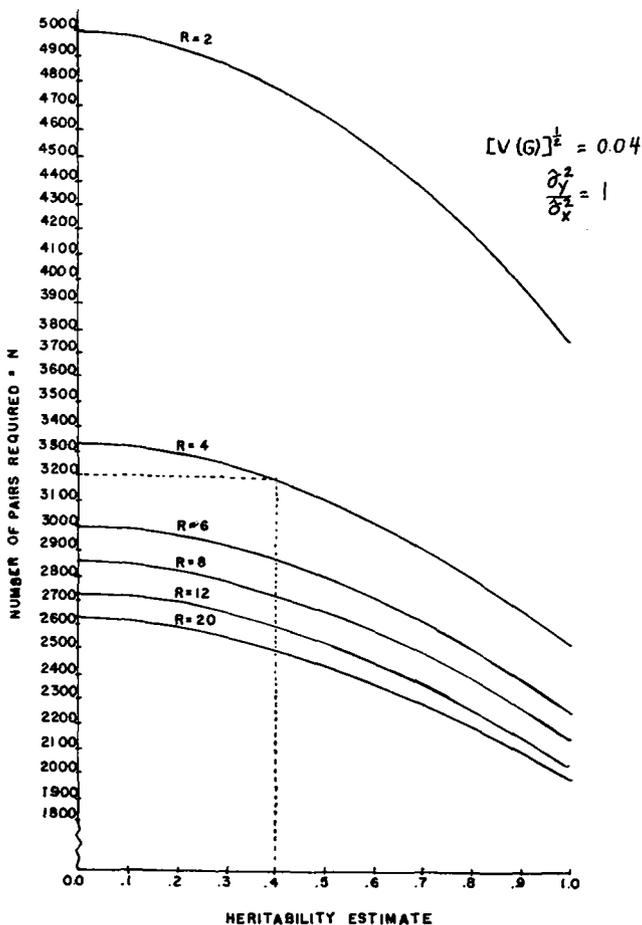
Figure 1. The required number of daughter-dam pairs, N, needed to estimate heritability for various values of the expected heritability estimate, G, and for several values of average number of pairs per class, R, when the standard error of the estimate is set at .04 and the ratio of daughter to dam variances=1.

=0.4 that 3200 pairs are required when $[V(G)]^{1/2}=0.04$. Multiplying 3200 by $\frac{1}{3^2}$ gives 356 which is the number of pairs needed. This result may also be obtained by application of equation (3).

This procedure may be easily extended to a many way classification since only within subclass variance and covariance components are used. In this analysis the sums of squares and crossproducts in equation (1) are within subclass sums of squares and crossproducts.

The number of pairs of records needed to estimate repeatability by regression of second records on first records may be determined similarly. Let the estimate of repeatability $=$ b. Then the estimate of $V(b)$ is that given in equation (1). Since $b = \dfrac{\Sigma xy}{\Sigma x^2}$ this variance can be written as

$$V(b) = \frac{\dfrac{\hat{\sigma}_y^2}{\hat{\sigma}_x^2} - b^2}{N - S - 1},$$

where $\hat{\sigma}_y^2 =$ the estimated error variance of second records,
  $\hat{\sigma}_x^2 =$ the estimated error variance of first records, and
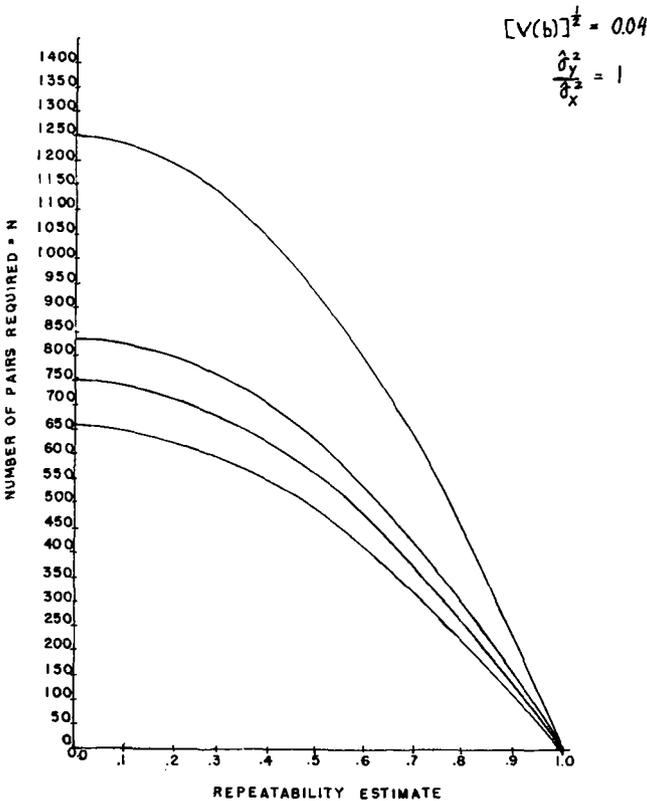 $S =$ the number of subclasses.

Figure 2. The required number of pairs of records, N, needed to estimate repeatability for various values of the expected repeatability estimate, b, and for several values of average number of pairs per class when the standard error of the estimate is set at .04 and the ratio of variances$=1$.

Substituting $X = \dfrac{\hat{\sigma}_y^2}{\hat{\sigma}_x^2}$ and $R = \dfrac{N}{S}$ and solving for N we obtain

$$N = \frac{(R)}{(R-1)} \frac{[X - b^2 + V(b)]}{V(b)}.$$

The number of pairs of records needed when $X = 1$ for various values of R, b, and V(b) may be obtained from figure 2.

Summary

A formula for estimating the number of daughter-dam pairs needed when estimating heritability by daughter on dam regression in the one-way classification analysis for fixed values of the average number of pairs per class, of the ratio of daughter to dam variance, of the anticipated heritability estimate, and of the desired variance of the estimate has been presented. The effect on N of changes in these four variables has been illustrated. This procedure was extended to multiple classification analysis and was also used to determine the number of pairs needed to estimate repeatability by regression of second records on first records.

Literature Cited

Fisher, R. A. 1952. Statistical Methods for Research Workers. Oliver and Boyd, London.

Tallis, G. M. and E. W. Klosterman. 1959. Efficient estimates of heritability from paternal half-sib correlations. J. Animal Sci. 18:622.

Wearden, Stanley. 1959. The use of the power function to determine an adequate number of progeny per sire in a genetic experiment. Biometrics 15:417.