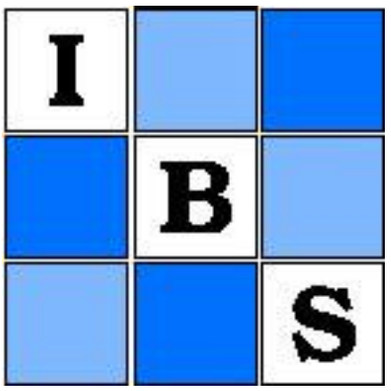

Query: Estimating "Heritability" of a Dichotomous Trait

It is common to conduct observations on some characteristic of laboratory mice, such as sleeping time under anaesthetic, in a number of different inbred strains. For a metric character such as sleeping time the analysis of variance can be used (with the usual assumptions) to test the significance of the differences among strains, and variance components attributable to within and among strain components can be calculated. The relative size of these components provides a point estimate such as the intraclass correlation or "heritability in the broad sense" with which the importance of genetic (among strain) and non-genetic (within strain) factors can be stated. The problem arises when the data are categorical. In this case, textbooks give methods of testing hypotheses, but I am unable to find a suitable estimate of the relative importance of the among strain and chance effect, except in the case of a 2×2 table.

As an example, I attach some data of Whitmire *et al.*, *Journal of the National Cancer Institute 47*, 1255–1265, who treated 12 strains of mice with a carcinogen and noted the number with and without tumours. The strains have been ranked in order of decreasing susceptibility. It is clear that the strain has an important influence on the incidence of tumours; the question is, how important?

Table 1

Strain	No. with Tumours	No. without Tumours	Total
1	26	1	27
2	27	3	30
3	35	14	49
4	18	9	27
5	33	20	53
6	11	11	22
7	11	11	22
8	13	15	28
9	13	22	35
10	5	19	24
11	5	30	35
12	2	24	26
TOTAL	199	179	378

Response

R. C. ELSTON

Department of Biostatistics and the Genetics Curriculum, University of North Carolina, Chapel Hill, North Carolina 27514, U.S.A.

A simple solution to this problem (Robertson and Lerner [1949]) is arbitrarily to assign the value 1 to each mouse that has tumours and 0 to each mouse that has no tumours, and then to calculate the intraclass correlation in the usual way from these "metric" values. If n_i is the number of mice tested from the ith strain, r_i of whom have tumours and $n_i - r_i$ of whom have no tumours, and there are c strains and a total of $N = \sum_i^c n_i$ mice, then the among-strain mean square is $\mathrm{MS}_A = \{\sum_i r_i^2/n_i - (\sum_i r_i)^2/N\}/(c - 1)$ and the within-strain mean square is $\mathrm{MS}_W = \{\sum_i r_i - \sum_i r_i^2/n_i\}/(N - c)$. MS_W is an unbiased estimate of the within-strain variance plus k times the among-strain variance, where $k = (N - \sum_i n_i^2/N)/(c - 1)$. (This was originally given, in another context, by Cochran [1939]). The intraclass correlation can thus be estimated by $r = (\mathrm{MS}_A - \mathrm{MS}_W)/\{\mathrm{MS}_A + (k - 1)\,\mathrm{MS}_W\}$. Using the data in Table 1, we find

$$\mathrm{MS}_A = (130.658 - 39601/378)/(12 - 1) = 2.354$$

$$\mathrm{MS}_W = (199 - 130.658)/(378 - 12) = 0.187$$

$$k = (378 - 13022/378)/(12 - 1) = 31.232$$

$$r = (2.354 - 0.187/\{2.354 + (31.232 - 1)0.187\} = 0.271$$

Several points should be noted about this estimate. In the first place heritability is a parameter that refers to a particular population as well as to a particular trait; what we have estimated is the heritability in that population of genotypes and environments of which the 12 strains are a random sample. Since the strains are inbred, the relevant population might be taken to be a conceptual one in which all individuals are homozygous. A more realistic relevant population would perhaps be a random mating one with the same gene frequencies as this conceptual population. Then, provided all the genetic variance is additive, what we have estimated is not the ratio of genetic variance (V_G) to total variance ($V_G + V_E$), but rather $2V_G/(2V_G + V_E)$. With that understanding the estimate is consistent, but not unbiased. It is not the maximum likelihood estimate; maximum likelihood estimation is possible, but would require an iterative procedure (except in the special case when all the n_i are equal). The estimate does not depend upon the values 0 and 1 that are chosen to quantify the dichotomous trait; any two distinct values a and b would result in exactly the same intraclass correlation.

If we let $p_i = r_i/n_i$, it can be seen that the numerator in MS_W is simply the sum of the variance estimates $n_i p_i(1 - p_i)$ of the binomial counts r_i. Thus, if we consider the within-strain variance, we notice that the environmental component (V_E) depends heavily on the genetic predisposition of the strain. One way of overcoming this theoretically undesirable feature is to use the arcsine square root transformation of p_i (Cochran [1940]). Let $x_i = \sin^{-1}\sqrt{p_i}$, with the result that the variance of $n_i x_i$ is approximately $n_i/4$ and the within groups sum of squares is simply $\sum n_i/4$. We then take

$$\mathrm{MS}_A = \{\sum_i n_i x_i^2 - (\sum_i n_i x_i)^2/N\}/(c - 1)$$

$$\mathrm{MS}_W = N/4(N - c)$$

and calculate k and r as before. For the example data we obtain

$$\begin{aligned} \mathrm{MS}_A &= (249.594 - 81253.714/378)/(12 - 1) = 3.149 \\ \mathrm{MS}_W &= 378/4(378 - 12) = 0.258 \\ k &= 31.232 \text{ as before,} \\ \text{and } r &= (3.149 - 0.258)/\{3.149 + (31.232 - 1)0.258\} = 0.264. \end{aligned}$$

The fact that these two estimates of the intraclass correlation are almost the same to two significant digits indicates that in practice we need not worry unduly about the theoretical difficulty previously mentioned.

Finally it should be noted that this estimate of "heritability" is relevant for the trait being analyzed, i.e., *presence or absence of tumours*. In the usual threshold model an underlying continuous "liability" to the presence of the trait is hypothesized, and the heritability *of this liability* is estimated (Falconer [1965]). Such an estimate of heritability is often utilized for rare diseases in man, and has the advantage of not being greatly influenced by the prevalence of the disease. It is, however, very dependent upon the form of the liability distribution assumed to be present in the population, which in turn depends upon the mode of inheritance. It is theoretically possible to assume any particular genetic model (such as monogenic or polygenic) for the data presented, and on the basis of that model estimate the heritability of liability to tumours. But, in the absence of any knowledge about the appropriate genetic mechanism, it would seem preferable to estimate the heritability of the trait itself. A numerical example showing how the underlying genetic assumptions can influence the heritability of liability has been presented by Kidd and Cavalli-Sforza [1973] and discussed by Elston [1973].

References

Cochran, W. G. [1939]. The use of the analysis of variance in enumeration by sampling. *Journal of the American Statistical Association 34,* 492–510.

Cochran, W. G. [1940]. The analysis of variance when experimental errors follow the Poisson or Binomial laws. *Annals of Mathematical Statistics 11,* 335–347.

Elston, R. C. [1973]. Methodologies in human behavior genetics. *Social Biology 20,* 276–279.

Falconer, D. S. [1965]. The inheritance of liability to certain diseases estimated from the incidence among relatives. *Annals of Human Genetics 29,* 51–76.

Kidd, K. K. and Cavalli-Sforza, L. L. [1973]. An analysis of the genetics of schizophrenia. *Social Biology 20,* 254–265.

Robertson, A. and Lerner, I. M. [1949]. The heritability of all-or-none traits: viability of poultry. *Genetics 34,* 395–411.

Alternative Response

WILLIAM G. HILL

Institute of Animal Genetics, Edinburgh EH 9 3JN, Scotland

CHARLES SMITH

ARC Animal Breeding Research Organisation, Edinburgh EH9 3JQ, Scotland.

In this note we consider some extensions to Elston's [1977] response and present an alternative form of analysis which is often used. Although the data are not ideal for such analysis, they serve to illustrate the methodology. The same data and notation are used.

Elston [1977] analysed the data by a standard analysis of variance of the data scored as 0 or 1. (This standard analysis seems preferable to the modification used by Robertson and Lerner [1949].) His intraclass correlation estimate ($\hat{r}_{01} = 0.271$) quantifies the ratio of among strain to total variation on the (0, 1) scale. The correlation on the (0, 1) scale depends, however, on the mean frequency of the 0 (or 1) class (i.e., proportion affected). With a different mean frequency, for example if lower or higher levels of carcinogen were given, different expected values of the intra-class correlation would be obtained. To remove this dependence on frequency a model of a continuous underlying distribution of liability (to induced tumours in this case) has often been used (e.g., Lush, Lamoreux and Hazel [1948], Robertson and Lerner [1949], Falconer [1965]). Individuals are affected if they exceed a certain threshold value of liability. Since there are probably several genetic and environmental factors affecting liability a continuous model is plausible (but rarely testable). Also, if the distribution is continuous it can be transformed (conceptually) to a normal distribution and use made of the properties of the normal curve.

Adjustment to continuous scale

Robertson and Lerner [1949] showed that the intra-class correlation (r_{01}) on the (0, 1) scale can be adjusted to a frequency independent correlation (r_c) on the continuous scale by

$$r_c = r_{01}\frac{p(1-p)}{z^2} \tag{1}$$

where $p = \sum_i r_i/N$ is the mean frequency of affected individuals and z is the ordinate on the standardised normal curve which corresponds to a probability p. With $p = 0.526$, $z = 0.398$ and $\hat{r}_{01} = 0.271$, $\hat{r}_c = (0.271)(0.526)(0.474)/(0.398)^2 = 0.427$. That is the intra-class correlation is much higher on the continuous scale than on the (0, 1) scale, and the change would have been greater had p been further from one-half.

An assumption made in (1) is that the variation in liability among groups is rather small. Our unpublished simulation results show, however, that the bias in the estimates is very small unless the correlation is high and the mean frequency (p) is near 0 or 1.

Standard errors

Equations giving exact, or even nearly exact formulae for standard errors of the intra-class correlation estimates are not available. As a first approximation we suggest the following, which has been found by simulation to give reasonably unbiassed estimates of the standard errors, except with frequencies near 0 or 1. From the extension of Fisher's formula given by Swiger, Harvey, Everson and Gregory [1964] the standard error of the estimate of

the intra-class correlation (t) in an unbalanced one-way classification analysis of normally distributed data is

$$\mathrm{SE}(\hat{t}) = (1 - t)[1 + (k - 1)t]\sqrt{\frac{2(N - 1)}{k^2(c - 1)(N - c)}}. \tag{2}$$

Following Robertson and Lerner [1949] we assume (2) holds for dichotomous traits, obtaining SE ($\hat{r}_{01}$) by inserting its estimate for t in (2). From Elston's analysis on the (0, 1) scale, SE ($\hat{r}_{01}$) = 0.0928 giving $\hat{r}_{01} = 0.27 \pm 0.09$. Now SE ($\hat{r}_c$) is obtained by scaling SE ($\hat{r}_{01}$) as for r_c itself, i.e.

$$\mathrm{SE}\ (\hat{r}_c) = \mathrm{SE}\ (\hat{r}_{01})[p(1 - p)/z^2], \tag{3}$$

giving SE ($\hat{r}_c$) = 0.146 or $\hat{r}_c = 0.43 \pm 0.15$ in the example.

Falconer's proband method

Another method of estimating the intra-class correlation of liability is directly from the frequency (p) in the population and the frequency (p_R) in the relatives of probands (Falconer [1965]). In this case all affected individuals are assumed independently ascertained and treated as probands. Then p_R is obtained as

$$\begin{aligned} p_R &= \sum r_i(r_i - 1)/\sum r_i(n_i - 1) \\ &= 4478/6913 = 0.648. \end{aligned} \tag{4}$$

Modifying the adjusted formula of Reich, James and Morris [1972, eq. (3)] to conform with our notation

$$\hat{r}_c = \frac{-x + x_R\sqrt{1 - (x^2 - x_R^2)(1 + x/a)}}{a + x_R^2(a + x)} \tag{5}$$

taking the root such that $0 \leq \hat{r}_c \leq 1$. In (5), x and x_R are the abscissae corresponding to p and p_R, respectively, on the standardised normal curve and $a = z/p$, where z is used as in (1). We have $p = 0.526$, $x = 0.065$, $a = 0.757$, $p_R = 0.648$, $x_R = 0.380$, giving $\hat{r}_c = 0.391$. This agrees well with the results obtained by the analysis of variance. Our simulation results show that in data such as in this example with multiple probands in each family, the method gives estimates which are biassed too low when the true intra-class correlation is high and Falconer's approximate formulae for standard errors of the estimates are not appropriate.

Genetic interpretation

So far we have made no genetic interpretation. If we can assume that there are no environmental effects common to all members of a strain (e.g., age, time) then the variance between strains can be interpreted genetically. The data analysed refer to well-known laboratory inbred strains, so we are not really entitled to consider them a random sample of any population. Had they been an unselected sample, the intra-class correlation would have estimated the variance among a conceptual population of inbred lines relative to the total variance. As discussed by Elston (see Falconer [1960], p. 267), if all genetic variance is additive and the lines are fully inbred the intra-class correlation among inbred line means measures $2V_G/(2V_G + V_E)$, where V_E is the environmental variance and equals unity by definition on the liability scale. With $r = 0.427$, this gives $V_G = 0.373$ and the heritability of liability in the foundation population as

$$h^2 = V_G/(V_G + V_E) = 0.27.$$

Inbred lines are usually available only for laboratory animals and in practice heritability of a dichotomous trait is assessed from sire families, dam families or twin pairs, which are sampled from a well defined population to which the estimated parameters will apply. After adjusting for any common environmental effects, the intra-class correlation on the continuous scale gives an estimate of the heritability of liability from $h^2 = r/R$ where R is the genetic relationship between family members (1/4 for half-sibs, 1/2 for full-sibs and 1 for identical twins).

To get a frequency independent statistic we have assumed there is a continuous underlying scale, and in the genetic interpretation we have assumed that there are no environmental effects common to all members of a strain and that there are many genetic factors involved. Even in the analysis of continuous traits we can never be sure of the latter. Also Kidd and Cavalli-Sforza [1973] have shown that an estimate of the heritability of liability may depend on the genetic model chosen.

References

Elston, R. C. [1977]. Estimating "heritability" of a dichotomous trait. *Biometrics 33*, 231–233.

Falconer, D. S. [1960]. *Introduction to Quantitative Genetics*. Oliver and Boyd, Edinburgh.

Falconer, D. S. [1965]. The inheritance of liability to certain diseases estimated from the incidence among relatives. *Annals of Human Genetics 29*, 51–76.

Kidd, K. K. and Cavalli-Sforza, L. L. [1973]. An analysis of the genetics of schizophrenia. *Social Biology 20*, 254–265.

Lush, J. L., Lamoreux, W. F. and Hazel, L. N. [1948]. The heritability of resistance to death in the fowl. *Poultry Science 27*, 375–388.

Reich, T., James, J. W. and Morris, C. A. [1972]. The use of multiple thresholds in determining the mode of transmission of semicontinuous traits. *Annals of Human Genetics 36*, 163–184.

Robertson, A. and Lerner, I. M. [1949]. The heritability of all-or-none traits: viability of poultry. *Genetics 34*, 395–411.

Swiger, L. A., Harvey, W. R., Everson, D. O. and Gregory, K. E. [1964]. The variance of intraclass correlation involving groups with one observation. *Biometrics 20*, 818–826.