Approximate Variance for Heritability Estimates

Dr. Bert Klei

Computational Genetics, Western Psychiatric Institute and Clinic, University of Pittsburgh Medical Center

Revised by S. Tsuruta

I am going to follow the following notation:

- a : estimate of the additive genetic variance.

- e : estimate of the environmental or residual.

- p : estimate of the phenotypic variance.

And the standard $p = a + e$ and the heritability

$$h^2 = \frac{a}{p} = \frac{a}{a + e}.$$

The negative inverse of the average information matrix gives us estimates (I_A^{-1}) for the variance of the estimators $var(a)$, $var(e)$ but also the covariance between the estimates a and e ($cov(a,e)$).

In general terms the variance of a function $f(\sigma)$ can be obtained as

$$var(f(\sigma)) = \frac{\partial f(\sigma)}{\partial \sigma} var(\sigma) \frac{\partial f(\sigma)}{\partial \sigma}.$$

In this case

$$f(\sigma) = h^2 = \frac{a}{a + e}$$

and

$$\sigma = \left(\begin{array}{c} a \\ e \end{array} \right).$$

The two partial derivatives are:

$$\frac{\partial h^2}{\partial a} = \frac{1}{a + e} - \frac{a}{(a + e)^2} = \frac{h^2(1 - h^2)}{a},$$

$$\frac{\partial h^2}{\partial e} = -\frac{a}{(a+e)^2} = -\frac{h^2(h^2)}{a}$$

and

$$var(\sigma) = \begin{pmatrix} var(a) & cov(a,e) \\ cov(e,a) & var(e) \end{pmatrix}$$

with appropriate entries taken from I_A^{-1}. Combining all of this information you get

$$var(f(\sigma)) = \frac{\partial f(\sigma)}{\partial \sigma} var(\sigma) \frac{\partial f(\sigma)}{\partial \sigma}$$

$$= \begin{pmatrix} \frac{h^2(1-h^2)}{a} & -\frac{h^2(h^2)}{a} \end{pmatrix} \begin{pmatrix} var(a) & cov(a,e) \\ cov(e,a) & var(e) \end{pmatrix} \begin{pmatrix} \frac{h^2(1-h^2)}{a} \\ -\frac{h^2(h^2)}{a} \end{pmatrix}$$

$$= \left(\frac{h^2}{a}\right) \left\{ (1-h^2)^2 var(a) - 2(1-h^2)h^2 cov(a,e) + (h^2)^2 var(e) \right\}.$$

This last quantity can easily be computed with the information obtained from the output of AIREMLF90. The question is how this is the same to what is written down in the ASREML manual. Here you have to realize that they present this in terms of a and p and not a and e. By using the fact that $p = a + e$ and as a result: $var(p) = var(a) + var(e) + 2cov(a,e)$ we can get the formula in the ASREML manual.

In the formula above are two pieces that involve $var(e)$ and $cov(a,e)$. It is easy to replace e with $p - a$. You get: $var(e) = var(p - e) = var(p) - var(a) - 2cov(a,e)$. In this you can replace $cov(a,e)$ with $cov(a, p - a) = cov(a,p) - var(a)$ and you get $var(e) = var(p) + var(a) - 2cov(a,p)$ You also need: $cov(a,e) = cov(a,p) - var(a)$ directly in the formula above.

Now replace the appropriate pieces in

$$var(h)^2 = \left(\frac{h^2}{a}\right)^2 \left\{ (1-h^2)^2 var(a) - 2(1-h^2)h^2 cov(a,e) + (h^2)^2 var(e) \right\}$$

$$= \left(\frac{h^2}{a}\right)^2 \left[(1-h^2)^2 var(a) - 2(1-h^2)h^2 \{cov(a,p) - var(a)\} + (h^2)^2 \{var(p) + var(a) - 2cov(a,p)\} \right]$$

$$= \left(\frac{h^2}{a}\right)^2 \left[\{1 - 2h^2 + (h^2)^2 + 2h^2 - 2(h^2)^2\} var(a) - 2\{h^2 - (h^2)^2 + (h^2)^2)\} \{cov(a,p) - var(a)\} + (h^2)^2 \{var(p)$$

$$= \left(\frac{h^2}{a}\right)^2 \{var(a) + (h^2)^2 var(p) - 2h^2 cov(a,p)\}$$

$$= (h^2)^2 \left\{ \frac{var(a)}{a^2} + \frac{(h^2)^2 var(p)}{a^2} - 2\frac{h^2 cov(a,p)}{a^2} \right\}$$

2

$$= \left(\frac{a}{p}\right)^2 \left\{ \frac{var(a)}{a^2} + \frac{var(p)}{p^2} - 2\frac{cov(a,p)}{ap} \right\}.$$

The last quantity is the one out of the ASREML manual, they replaced a with n for numerator and p with d for denominator. Their formula is correct for any ratio of the estimates, not only the heritability.

Approximate Variance for Correlation Estimates

Here I am going to use the following notation:

- x_i: estimate of the variance for trait i.

- x_j : estimate of the variance for trait j.

- x_{ij} : estimate of the covariance between trait i and j.

The variance and covariance can be additive, environmental, or phenotypic. The correlation will be denoted by r and is determined as

$$\frac{x_{12}}{\sqrt{x_1 x_2}}$$

. The negative inverse of the average information matrix gives us estimates for the variance of the estimators $var(x_1)$, $var(x_1)$, and $cov(x_{12})$ but also the covariance between the various estimates $(cov(x_1, x_2)$, $cov(x_1, x_{12})$, $cov(x_2, x_{12}))$. We can use the same general formula

$$var(f(\sigma)) = \frac{\partial f(\sigma)}{\partial \sigma} var(\sigma) \frac{\partial f(\sigma)}{\partial \sigma}$$

to obtain the variance of the correlation estimate. Now we have

$$f(\sigma) = r = \frac{x_{12}}{\sqrt{x_1 x_2}}$$

and

$$\sigma = \begin{pmatrix} x_1 \\ x_2 \\ x_{12} \end{pmatrix}.$$

The three partial derivatives are:

$$\frac{\partial r}{\partial x_1} = \frac{r}{2x_1},$$

$$\frac{\partial r}{\partial x_2} = \frac{r}{2x_2}$$

and

$$\frac{\partial r}{\partial x_{12}} = \frac{r}{x_{12}}.$$

Then

$$var(r) = \begin{pmatrix} -\dfrac{r}{2x_1} & \dfrac{r}{2x_2} & \dfrac{r}{x_{12}} \end{pmatrix} \begin{pmatrix} var(x_1) & cov(x_1, x_2) & cov(x_1, x_{12}) \\ cov(x_2, x_1) & var(x_2) & cov(x_2, x_{12}) \\ cov(x_{12}, x_1) & cov(x_{12}, x_2) & var(x_{12}) \end{pmatrix} \begin{pmatrix} -\dfrac{r}{2x_1} \\ -\dfrac{r}{2x_2} \\ \dfrac{r}{x_{12}} \end{pmatrix}.$$

Just multiplying this out gives

$$var(r) = r^2 \left[\frac{var(x_1)}{4x_1^2} + \frac{var(x_2)}{4x_2^2} + \frac{var(x_{12})}{x_{12}^2} + \frac{cov(x_1, x_2)}{2x_1 x_2} - \frac{cov(x_1, x_{12})}{x_1 x_{12}} - \frac{cov(x_2, x_{12})}{x_2 x_{12}} \right].$$