Standard errors of heritability estimates

Citation for published version:

Link:
Link to publication record in Edinburgh Research Explorer

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Standard errors of heritability estimates

Author: Stefan McKinnon Edwards
Date: September 2017

PDF version with equations correctly formatted. dmu/heritabilities.pdf

Cite as:

Derivation

The heritability, and similar genetic parameters, are non-linear transformations of the vector of variance components \( \boldsymbol{\Theta} = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2) \). i.e. we can write it more generally as \( H = f(\boldsymbol{\Theta}) = \frac{\sigma_1^2}{\sigma_2^2} \), for a subset of elements of \( \boldsymbol{\Theta} \). If \( f(\boldsymbol{\Theta}) \) is differentiable and continuous around the estimated \( \hat{\boldsymbol{\Theta}} \), then we can perform Taylor series expansion of the transformation.

This provides a tangible expression of the pairwise, conditional co-variances of the transformation.

The standard error of the heritability estimate is then:

\[
\text{s.e.}(H^2) = \sqrt{\left( \frac{\partial^2 f}{\partial \sigma_1^2} \right)^2 (se_{\sigma_1})^2 + \left( \frac{\partial^2 f}{\partial \sigma_1^2 \partial \sigma_2^2} \right)^2 (se_{\sigma_1 \sigma_2})^2 + 2 \left( \frac{\partial^2 f}{\partial \sigma_1^2 \partial \sigma_2^2} \right) \rho_{\sigma_1 \sigma_2} se_{\sigma_1} se_{\sigma_2}}
\]

The partial derivatives are given as

\[
\frac{\partial f}{\partial \sigma_1^2} = \frac{\sigma_1^2}{\sigma_2^2}, \quad \frac{\partial f}{\partial \sigma_2^2} = -\frac{\sigma_1^2}{\sigma_2^2}
\]

\[
\rho_{\sigma_1 \sigma_2} se_{\sigma_1} se_{\sigma_2} = \text{cov}(\sigma_1^2, \sigma_2^2)
\]

i.e. the variation of the non-linear transformation. The left-hand side of (3) becomes a tangible expression:

\[
f(\boldsymbol{\Theta}) = f(\hat{\boldsymbol{\Theta}}) + \sum_{i=1}^{t} \frac{\partial f}{\partial \sigma_i^2} |_{\sigma_i = \sigma_i^*} (\sigma_i^2 - \sigma_i^*)
\]

Taking expectations on both sides we get:

\[
E[f(\boldsymbol{\Theta})] = f(\hat{\boldsymbol{\Theta}})
\]

Subtracting (2) from (1), the right hand side of (2) cancels the first term in (1), and the left hand side is the deviation from the expectation, as given in the following expression:

\[
f(\boldsymbol{\Theta}) - E[f(\boldsymbol{\Theta})] = \sum_{i=1}^{t} \frac{\partial f}{\partial \sigma_i^2} |_{\sigma_i = \sigma_i^*} (\sigma_i^2 - \sigma_i^*)
\]

Final step is squaring and taking the expectation. The left-hand side of (3) becomes

\[
E[f(\boldsymbol{\Theta}) - E[f(\boldsymbol{\Theta})]^2] = \text{Var}(f(\boldsymbol{\Theta}))
\]

i.e. the variation of the non-linear transformation. The left-hand side of (3) becomes a tangible expression:

\[
E[f(\boldsymbol{\Theta}) - E[f(\boldsymbol{\Theta})]^2] = \sum_{i=1}^{t} \left( \frac{\partial f}{\partial \sigma_i^2} |_{\sigma_i = \sigma_i^*} \right)^2 (\sigma_i^2 - \sigma_i^*)^2
\]

Remembering that \( E(\sigma_i^2) = \sigma_i^2 \) (when \( \sigma_i^2 \) maximizes the likelihood), we can recognize \( \text{Var}(\sigma_i^2) = (\sigma_i^2 - \sigma_i^*)^2 = \sigma_i^2 \), i.e. the variance of the variance component.

Similarly, \( \text{Cov}(\sigma_i^2, \sigma_j^2) = (\sigma_i^2 - \sigma_i^*) (\sigma_j^2 - \sigma_j^*) = \sigma_i^2 \sigma_j^2 \). Thus the Taylor series expanded estimate of the variance of a non-linear transformation of the variance component becomes

\[
\text{Var}(f(\boldsymbol{\Theta})) = \sum_{i=1}^{t} \left( \frac{\partial f}{\partial \sigma_i^2} |_{\sigma_i = \sigma_i^*} \right)^2 \sigma_i^2 + 2 \sum_{i=1}^{t} \sum_{j=i+1}^{p} \frac{\partial f}{\partial \sigma_i^2} \frac{\partial f}{\partial \sigma_j^2} \sigma_i \sigma_j
\]
The genetic correlation can be estimated using a bivariate model, such as:

\[ \text{Cor}(\sigma_i^2, \sigma_j^2) = \frac{\text{Cov}(\sigma_i^2, \sigma_j^2)}{\sqrt{\text{Var}(\sigma_i^2) \cdot \text{Var}(\sigma_j^2)}} = \frac{\sigma_{i,j}}{\text{se}_{i} \cdot \text{se}_{j}} \]

with the product in the denominator (\text{se}_{i} and \text{se}_{j}) simply being the asymptotic standard errors. Hence, using the output from DMU, the values needed can be calculated as:

\[ \sigma_{i,j} = \text{Cor}(\sigma_i^2, \sigma_j^2) \cdot \text{se}_{i} \cdot \text{se}_{j} \quad \text{and} \quad \sigma_{i}^2 = \left(\frac{\text{se}_{i}}{\text{se}_{i}}\right)^2 \]

Some genetic parameters

Heritability (narrow-sense)

For

\[ h^2 = \frac{\sigma_e^2}{\sigma_e^2 + \sigma_g^2}, \]

then

\[ \frac{\partial h^2}{\partial \sigma_e} = -\frac{\sigma_e^2}{(\sigma_e^2 + \sigma_g^2)^2} \]

and

\[ \frac{\partial h^2}{\partial \sigma_g} = \frac{\sigma_g^2}{(\sigma_e^2 + \sigma_g^2)^2} \]

Note: There really is a negative sign in the latter. Note: Use the quotient rule for deriving partial derivations.

\[ \text{Var}(h^2) = \left(\frac{\partial h^2}{\partial \sigma_e} \right)^2 \sigma_e^2 + \left(\frac{\partial h^2}{\partial \sigma_g} \right)^2 \sigma_g^2 + 2 \left(\frac{\partial h^2}{\partial \sigma_e} \right) \left(\frac{\partial h^2}{\partial \sigma_g} \right) \sigma_{e,g} \]

Heritability (total)

For

\[ h^2 = \frac{\sigma_e^2}{\sigma_e^2 + \sigma_m^2 + \sigma_g^2}, \]

then

\[ \frac{\partial h^2}{\partial \sigma_e} = \frac{\sigma_e^2 + \sigma_m^2}{(\sigma_e^2 + \sigma_m^2 + \sigma_g^2)^2} \]

and

\[ \frac{\partial h^2}{\partial \sigma_m} = \frac{\partial h^2}{\partial \sigma_g} = \frac{-\sigma_g^2}{(\sigma_e^2 + \sigma_m^2 + \sigma_g^2)^2} \]

Therefore,

\[ \text{Var}(h^2) = \left(\frac{\partial h^2}{\partial \sigma_e} \right)^2 \sigma_e^2 + \left(\frac{\partial h^2}{\partial \sigma_m} \right)^2 \sigma_m^2 + \left(\frac{\partial h^2}{\partial \sigma_g} \right)^2 \sigma_g^2 + 2 \left(\frac{\partial h^2}{\partial \sigma_e} \right) \left(\frac{\partial h^2}{\partial \sigma_m} \right) \sigma_{e,m} + 2 \left(\frac{\partial h^2}{\partial \sigma_e} \right) \left(\frac{\partial h^2}{\partial \sigma_g} \right) \sigma_{e,g} + 2 \left(\frac{\partial h^2}{\partial \sigma_m} \right) \left(\frac{\partial h^2}{\partial \sigma_g} \right) \sigma_{m,g} \]

Genetic correlation

The genetic correlation can be estimated using a bivariate model, such as:

\[
\begin{pmatrix}
  y_1 \\
  y_2
\end{pmatrix} =
\begin{pmatrix}
  \mu_1 \\
  \mu_2
\end{pmatrix} +
\begin{pmatrix}
  g_1 \\
  g_2
\end{pmatrix} +
\begin{pmatrix}
  e_1 \\
  e_2
\end{pmatrix}
\]
where \( \text{Var}(g_i) = G \sigma^2_i \) and \( \text{Var}(g_j) = G \sigma^2_j \) with variance-covariance matrix \( G \) for both random effects, and \( \text{Cov}(g_i, g_j) = \sigma_{1,2} \) as the genetic covariance between two traits \( g_i \) and \( g_j \). In this context, \( \sigma_{1,2} \) is no different from \( \sigma_i^2 \), as they are both variance components that are estimated by e.g. AI-REML.

For genetic correlation

\[
\rho = \frac{\sigma_{1,2}}{\sqrt{\sigma_1^2 \cdot \sigma_2^2}}
\]

then

\[
\frac{\partial \rho}{\partial \sigma_{1,2}} = \frac{1}{\sqrt{\sigma_1^2 \cdot \sigma_2^2}}
\]

\[
\frac{\partial \rho}{\partial \sigma_1^2} = \frac{-\sigma_{1,2}}{2 \cdot \sqrt{\sigma_1^2 \cdot (\sigma_2^2)^3}}
\]

and

\[
\frac{\partial \rho}{\partial \sigma_2^2} = \frac{-\sigma_{1,2}}{2 \cdot \sqrt{\sigma_2^2 \cdot (\sigma_1^2)^3}}
\]

Therefore,

\[
\text{Var}(\rho) = 2 \left( \frac{\partial \rho}{\partial \sigma_1^2} \right)^2 \sigma_1^4 + 2 \left( \frac{\partial \rho}{\partial \sigma_2^2} \right)^2 \sigma_2^4 + \left( \frac{\partial \rho}{\partial \sigma_{1,2}} \right)^2 \sigma_{1,2}^4 + 2 \left( \frac{\partial \rho}{\partial \sigma_1^2} \right) \left( \frac{\partial \rho}{\partial \sigma_{1,2}} \right) \sigma_1 \sigma_{1,2} \sigma_{1,3} + 2 \left( \frac{\partial \rho}{\partial \sigma_2^2} \right) \left( \frac{\partial \rho}{\partial \sigma_{1,2}} \right) \sigma_2 \sigma_{1,2} \sigma_{1,3}.
\]

References

To cite the software DMU, refer to Madsen et al., 1994 and Madsen & Jensen, 2000. When using the AI-REML algorithm for estimating variance components (with dmub) also refer to Gilmeur et al., 1995 and Johnson & Thompson, 1995, and Jensen et al., 1997.

6. @incollection{Madsen1994, author = {Madsen, Per and Jensen, Just and Thompson, Robin}, booktitle = {Proceedings of 5th WCGALP}, pages = {455–462}, title = {Estimation of (Co)variance Components in Multivariate Mixed Linear Models using Average of Observed and Expected Information}, year = {1994} }
8. @inproceedings{Jensen1997a, address = {Guelph, Canada}, author = {Jensen, Just and Mantsysaari, Esa A. and Madsen, Per and Thompson, Robin}, booktitle = {Proceedings of 5th WCGALP}, isbn = {9076496518}, title = {Residual Maximum Likelihood Estimation of (Co) Variance Components in Multivariate Mixed Linear Models using Average Information}, year = {1997} }
9. @inproceedings{Johnson2001, address = {Guelph, Canada}, author = {Johnson, D.L. and Thompson, R.}, booktitle = {Proceedings of 5th WCGALP}, isbn = {9076496518}, title = {Restricted Maximum Likelihood Estimation of Variance Components for Mixed Linear Models Using Average Information}, year = {1997} }

http://www.iysik.com/dmu:heritabilities