Flexible heritability estimation from linear mixed models

Australian National University

The `heritable` R package brings the commonly used heritability estimators together in one interface, so you can estimate, compare and report heritability directly from models fitted with asreml, lme4 or glmmTMB, or from any other fitted model, given extracted estimates of the relevant model parameters.

Available on CRAN



Background

In a plant breeding trial, heritability measures how much of the variation in a trait is due to genetics. It tells breeders whether a trial holds enough genetic signal to select on. It is widely reported, yet there are several ways to calculate it.

Challenges

1. Several estimators are in common use (Standard, Cullis, Oakey, Piepho and Delta), and they agree only in simple, balanced trials.
2. The calculations depend on the software that fitted the model, so they are often re-coded by hand.
3. Multi-environment trials, spatial residuals and genomic relationship matrices go beyond textbook formulas.

Aims

1. Bring the commonly used heritability estimators together in one interface.
2. Work directly with asreml, lme4 and glmmTMB models, including complex random-effect and residual structures.
3. Make heritability easy to report reproducibly, with clear definitions and confidence intervals.

Statistical framework

Estimating heritability means partitioning variance under a linear mixed model (LMM):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_g\mathbf{u}_g + \mathbf{Z}_{ge}\mathbf{u}_{ge} + \boldsymbol{\varepsilon}$$

with genetic effects $\mathbf{u}_g$, genotype-by-environment (G×E) effects $\mathbf{u}_{ge}$ and residual errors $\boldsymbol{\varepsilon}$.

Phenotypic variance partition

$$H^2 = \frac{V_G}{V_P}$$

Based on genotype means or best linear unbiased estimates (BLUEs).

Standard · Delta (BLUE) · Piepho

Genetic variance partition

$$H^2 = 1 - \frac{V_{\mathrm{PEV}}}{V_G}$$

Based on the prediction error variance (PEV) of BLUPs.

Cullis · Delta (BLUP) · Oakey

Each estimator is built from contrasts between genotypes. Contrasts within one environment give stratified heritability, and contrasts averaged across environments give marginal heritability. A known additive relationship matrix gives narrow-sense heritability, h^2.

Impact

Reliable, reusable heritability calculations from fitted linear mixed models for transparent reporting.

Looking ahead

1. Support more of the variance structures available in asreml.
2. Expand the package to support other models.

Usage

Name the genetic term with `target`; all estimators are returned by default.

```
library(heritable)
lettuce_subset <- subset(lettuce_phenotypes, loc == "L2")
fit <- lme4::lmer(y ~ rep + (1 | gen), data = lettuce_subset)
H2(fit, target = "gen")
```

```
   Cullis     Oakey    Piepho     Delta  Standard
0.8294971 0.8294971 0.8294971 0.8294971 0.8294971
```

In a balanced single trial, all estimators agree.

One model, many questions

One asreml G×E fit gives heritability for the main genetic effect, averaged over locations (marginal) or within one location (stratified).

```
fit_gxe <- asreml::asreml(y ~ rep, random = ~ gen + loc + gen:loc,
                          data = lettuce_phenotypes)
H2(fit_gxe, target = "gen", marginal = FALSE)
H2(fit_gxe, target = "gen", marginal = TRUE)
H2(fit_gxe, target = "gen", stratification = data.frame(loc = "L1"))
```

Heritability	Standard	Piepho	Cullis	Delta	Oakey
Main effect	0.775	0.781	0.782	0.782	0.782
Marginal (all locations)	0.890	0.890	0.892	0.892	0.882
Stratified: location L1	0.725	0.725	0.787	0.787	0.778
Stratified: location L2	0.798	0.798	0.831	0.831	0.822
Stratified: location L3	0.795	0.795	0.830	0.830	0.821

Downy mildew resistance of 89 lettuce genotypes at three locations (Hadasch et al., 2016); darker cells are higher. The question asked moves the estimate more than the choice of estimator.

Narrow-sense and uncertainty

`h2()` works the same way for models with a genomic relationship matrix (GRM), and `confint()` gives parametric bootstrap confidence intervals.

```
fit_grm <- asreml::asreml(y ~ rep,
  random = ~ vm(gen, source = lettuce_GRM) * idv(loc),
  data = lettuce_phenotypes)
h2(fit_grm, target = "gen", source = list(lettuce_GRM = lettuce_GRM))
```

```
   Cullis     Oakey    Piepho     Delta  Standard
0.9233324 0.6867032 0.8980665 0.9228154 0.8974431
```

```
estimate <- H2(fit, target = "gen", method = c("Cullis", "Standard"))
confint(estimate, B = 100, seed = 2026)
```

```
              2.5 %    97.5 %
Cullis    0.7792096 0.9395049
Standard  0.7792096 0.9395049
```

Thank you, contributors

Thank you to everyone who has contributed code, testing and bug reports, especially:

Fonti Kar (FK)
Built the foundations and the first CRAN release.

Yidi Deng (YD)
Added narrow-sense, marginal and stratified heritability, and confidence intervals.

Weihao (Patrick) Li (WL)
Laid the groundwork for glmmTMB support.

Paul Schmidt (PS)
Added the glmmTMB backend and the reliability metric.

Get involved

Contributions of all shapes and sizes are welcome!

- CRAN: `install.packages("heritable")`
- Latest features: `pak::pak("anu-aagi/heritable")`
- Docs: anu-aagi.github.io/heritable
- Contribute or report issues: github.com/anu-aagi/heritable
- Contact: Emi Tanaka emi.tanaka@anu.edu.au

References. Standard: Falconer & Mackay (1996) *Introduction to Quantitative Genetics*. Cullis: Cullis, Smith & Coombes (2006) *JABES*. Oakey: Oakey et al. (2006) *Theoretical and Applied Genetics*. Piepho: Piepho & Möhring (2007) *Genetics*. Delta: Schmidt et al. (2019) *Crop Science*; Schmidt et al. (2019) *Genetics*. Lettuce data: Hadasch et al. (2016) *The Plant Genome*.

AAGI Analytics for the Australian Grains Industry · GRDC Grains Research & Development Corporation · Adelaide University · Curtin University · The University of Queensland Australia

The support of the Grains Research and Development Corporation (GRDC) provided through the Analytics for the Australian Grains Industry (AAGI) project is acknowledged.