Lecture 4

Heritability

Heritability: An Intuitive Approach
First Definition

Broad Sense:  Proportion of the phenotypic variation due to genetic causes

\[ H^2 = \frac{\sigma_G^2}{\sigma_Y^2} \]

Useful to determine to what extent genetics vs environment impact a trait

Narrow Sense:  Proportion of the phenotypic variation due to additive genetic effects

\[ h^2 = \frac{\sigma_A^2}{\sigma_Y^2} \]

Useful to determine to what extent directional selection can improve a trait

Alternatively, \( h^2 \) is the proportion of the total variation attributable to differences in breeding values.
Perfectly heritable trait

All of the superiority of parent is passed to the offspring

Estimation of Heritability: Full sibs

♂♀ x ♂♀ x ♂♀ x ♂♀
From Lecture 3 Covariance Full sibs

<table>
<thead>
<tr>
<th>Number of IBD alleles</th>
<th>Probability of Sharing</th>
<th>Contribution to variances</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 IBD alleles</td>
<td>4/16</td>
<td>0</td>
</tr>
<tr>
<td>1 IBD alleles</td>
<td>8/16</td>
<td>$\frac{1}{2} \sigma_A^2$</td>
</tr>
<tr>
<td>2 IBD alleles</td>
<td>4/16</td>
<td>$\sigma_A^2 + \sigma_D^2$</td>
</tr>
</tbody>
</table>

$Cov(G_{Oi}, G_{O2}) = \text{Total covariance} = \text{Sum Probability x contribution} =$

$$\sigma_{FS} = \left(\frac{8}{16}\right)\left(\frac{1}{2} \sigma_A^2\right) + \left(\frac{4}{16}\right)\left(\sigma_A^2 + \sigma_D^2\right) = \frac{1}{2} \sigma_A^2 + \frac{1}{4} \sigma_D^2$$

Partitioning Variances into Sources of Variation

ANOVA
- Least Squares
- Method of Moments
- Linear Model

$$Z_{ij} = \mu + f_i + w_{ij}$$

Phenotype of the jth offspring from the ith family

$$\sigma_z^2 = \sigma_f^2 + \sigma_w^2$$
Genetic Variance Equivalence

• The phenotypic covariance between members of the same group equals the variance among groups

\[ \text{Cov(within family)} = \sigma(z_{ij}, z_{ik}) \]
\[ = \sigma[(\mu + f_i + w_{ij}), (\mu + f_i + w_{ik})] \]
\[ = \sigma(f_i, f_i) + \sigma(f_i, w_{ik}) + \sigma(w_{ij}, f_i) + \sigma(w_{ij}, w_{ik}) \]
\[ = \sigma_f^2 \]

Note the i subscript is the same.

Within family observation correlated vs. uncorrelated

Correlated observations within a family indicates that, given the family mean, you know something about the performance of the individual.

Uncorrelated observations within family: knowledge of the family mean tells you nothing about performance of the individual.

Thus if family means vary, then individuals within families are correlated.
Relationship of Estimated Variance Component with Causal Underlying Variance Components

Covariance within Family = Variance Among Families

\[ \sigma_{FS} = \sigma_f^2 \]

\[ \sigma_{FS} = \frac{1}{2} \sigma_A^2 + \frac{1}{4} \sigma_D^2 \]

Estimate the Variance Among Families by ANOVA

\[ \text{ANOVA} \]

Computational Formulas

<table>
<thead>
<tr>
<th>Factor</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among-families</td>
<td>N - 1</td>
<td>Ss_f = n \sum_{i=1}^{N} (y_i - \bar{y}_f)^2</td>
<td>MS_f = Ss_f / (N - 1)</td>
<td>\sigma^2_{n(FS)} + \sigma_f^2</td>
</tr>
<tr>
<td>Within-families</td>
<td>T - N</td>
<td>SS_w = \sum_{i=1}^{N} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2</td>
<td>MS_w = SS_w / (T - N)</td>
<td>\sigma^2_{n(FS)}</td>
</tr>
</tbody>
</table>

*Note: The total sample size is \( T = Nn \). Degrees of freedom are denoted by df, observed sums of squares by SS, and expected mean squares by E(MS).*
Method of Moments

Set Expected mean squares equal to estimated mean squares and solve

\[ \hat{\sigma}_f^2 = \frac{MS_f - MS_w}{n} \quad \hat{\sigma}_w^2 = MS_w \]

\[ \sigma_f^2 = \sigma_{FS} = \frac{1}{2} \sigma_A^2 + \frac{1}{4} \sigma_D^2 \]

\[ h^2 = \frac{\sigma_A^2}{\sigma_P^2} \cong \frac{2\hat{\sigma}_f^2}{\hat{\sigma}_f^2 + \hat{\sigma}_w^2} \]

Variances

\[ \sigma^2_w = \sigma_f^2 + \sigma_{12w}^2 \]
Turkey Example

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>ss</th>
<th>ms</th>
<th>E(ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among Family</td>
<td>1</td>
<td>1.5</td>
<td>1.5</td>
<td>$\sigma_w^2 + 3\sigma_f^2$</td>
</tr>
<tr>
<td>Within Family</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>$\sigma_w^2$</td>
</tr>
</tbody>
</table>

$\hat{\sigma}_f^2 = \frac{(1.5 - 1)}{3} = .167$

$\hat{\sigma}_w^2 = 1$

$s.e. = \text{formulas given in notes}$

Typically very large

Nested Full-sib and Half Sib

$z_{ijk} = \mu + s_i + d_{ij} + w_{ijk}$

$\sigma_z^2 = \sigma_s^2 + \sigma_d^2 + \sigma_w^2$
### ANOVA Table

Table 4.3. Summary of a (balanced) nested analysis of variance involving \(N\) sites, \(M\) dams per site and \(n\) offspring per dam. \(T = MNn\) is the total number of sibs in the design.

<table>
<thead>
<tr>
<th>Factor</th>
<th>(df)</th>
<th>Sums of Squares</th>
<th>MS</th>
<th>F(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sites</td>
<td>(N - 1)</td>
<td>(M_n \sum_{i=1}^{N} \sum_{j=1}^{M_i} (\tau_i - \tau)^2)</td>
<td>(SS_n/d_k)</td>
<td>(\sigma^2 + n\sigma^2 + M_n\sigma^2)</td>
</tr>
<tr>
<td>Dams (sires)</td>
<td>(N(M - 1))</td>
<td>(n \sum_{i=1}^{N} \sum_{j=1}^{M_i} G_{ij} - (\tau_i)^2)</td>
<td>(SS_d/d_k)</td>
<td>(\sigma^2 + n\sigma^2)</td>
</tr>
<tr>
<td>Sibs (dams)</td>
<td>(T = NM)</td>
<td>(\sum_{i=1}^{N} \sum_{j=1}^{M_i} \sum_{k=1}^{n} (x_{ijk} - \tau_i)^2)</td>
<td>(SS_u/d_k)</td>
<td>(\sigma^2_u)</td>
</tr>
</tbody>
</table>

### Components of Variance

- Covariance among members of a group equals the variance between groups
- Total variance can be partitioned into that within and among full sib families

\[
\sigma^2_z = \sigma^2_{\text{among FS}} + \sigma^2_{\text{Within FS}}
\]

\[
\sigma^2_z = \sigma^2_{FS} + \sigma^2_{W}
\]

From Lecture 3

\[
\sigma^2_s = \sigma^2_{HS} = \frac{1}{4} \sigma^2_A
\]

\[
\sigma^2_d = \sigma^2_{FS} - \sigma^2_{HS}
\]

\[
\sigma^2_d = \frac{1}{2} \sigma^2_A + \frac{1}{4} \sigma^2_D - \frac{1}{4} \sigma^2_A
\]

\[
= \frac{1}{4} \sigma^2_A + \frac{1}{4} \sigma^2_D
\]
### Total Variance in terms of linear model and in terms of genetic components

\[
\sigma_z^2 = \sigma_s^2 + \sigma_d^2 + \sigma_w^2 \quad \text{Linear Model}
\]

\[
\sigma_z^2 = \sigma_A^2 + \sigma_D^2 + \sigma_E^2 \quad \text{Genetic Effects}
\]

- \( \sigma_s^2 = \frac{1}{4} \sigma_A^2 \)
- \( \sigma_d^2 = \frac{1}{4} \sigma_A^2 + \frac{1}{4} \sigma_D^2 + \sigma_{Ec}^2 \)
- \( \sigma_w^2 = \frac{1}{2} \sigma_A^2 + \frac{3}{4} \sigma_D^2 + \sigma_{Es}^2 \)

### Estimating Genetic Components

\[
\hat{\sigma}_s^2 = \left( MS_s - MS_d \right) / Mn
\]

\[
\hat{\sigma}_d^2 = \left( MS_d - MS_w \right) / n
\]

\[
\hat{\sigma}_w^2 = MS_w
\]
Example: Table 4.4
10 sires; 3 dams/sire; 10 offspring/dam

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>ms</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sires</td>
<td>9</td>
<td>470</td>
<td>(\sigma_w^2 + 10\sigma_d^2 + 30\sigma_s^2)</td>
</tr>
<tr>
<td>dams/(sires)</td>
<td>20</td>
<td>170</td>
<td>(\sigma_w^2 + 10\sigma_d^2)</td>
</tr>
<tr>
<td>within dams</td>
<td>270</td>
<td>20</td>
<td>(\sigma_w^2)</td>
</tr>
</tbody>
</table>

\[\begin{align*}
\sigma_w^2 &= 20 \\
\sigma_d^2 &= \frac{170 - 20}{10} = 15 \\
\sigma_s^2 &= \frac{470 - 170}{30} = 10 \\
\sigma_p^2 &= 20 + 15 + 10 = 45 \\
\end{align*}\]

\[h^2 = \frac{4(10)}{45} = .89\]

Estimating \(h^2\)
parent offspring regression

\[\begin{align*}
z_{o1} &= \mu + \beta_{o \mid p} (z_{pi} - u) + e_i \\
z_{oi} &= \beta_0 + \beta_{o \mid p} z_{pi} + e_i
\end{align*}\]

\[E(\beta_{o \mid p}) = \frac{\sigma(z_o, z_p)}{\sigma^2(z_p)} = \frac{1}{2} \frac{\sigma_A^2 + \sigma(e_p, e_o)}{\sigma_z^2}\]

When is the environmental covariance between parent offspring not zero?

\[h^2 = 2\beta_{o \mid p}\]
Regression on One Parent: Example Butterfat (kg)

<table>
<thead>
<tr>
<th>Parent Dam (X)</th>
<th>Offspring Daughter (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>150</td>
<td>132</td>
</tr>
<tr>
<td>102</td>
<td>122</td>
</tr>
<tr>
<td>129</td>
<td>104</td>
</tr>
<tr>
<td>127</td>
<td>103</td>
</tr>
<tr>
<td>149</td>
<td>112</td>
</tr>
<tr>
<td>133</td>
<td>130</td>
</tr>
<tr>
<td>164</td>
<td>140</td>
</tr>
<tr>
<td>150</td>
<td>148</td>
</tr>
<tr>
<td>124</td>
<td>120</td>
</tr>
<tr>
<td>141</td>
<td>168</td>
</tr>
</tbody>
</table>

\[
\sum X_i = 1,369 \quad \sum Y_i = 1,279
\]
\[
\sum X_i^2 = 190,217 \quad \sum X_i Y_i = 176,447
\]
\[
h^2 = 2 \left( \frac{176,447 - \frac{(1,369)(1,279)}{10}}{190,217 - \frac{(1,369)^2}{10}} \right) = .96
\]

Sex Limited Traits must be multiplied by 2
Because only one of the parents was measured

Estimating \( h^2 \)
mid parent offspring regression

\[
z_{o1} \quad \ldots \quad z_{on} \quad \ldots \quad z_{m1} \quad z_{f1} \quad \ldots \quad z_{mn} \quad z_{fn}
\]
\[
z_{oi} = \mu + \beta_{ojmp} \left( \frac{z_{mi} + z_{fi}}{2} - \mu \right) + e_i
\]

\[
E(\beta_{ojmp}) = \frac{\sigma(z_{o}, z_{mp})}{\sigma^2(z_{mp})} = \frac{\sigma\left(z_{o}, \frac{z_{m} + z_{f}}{2}\right)}{\sigma^2\left(z_{m} + z_{f}\right)} = \frac{1}{4} \sigma^2(z_{o}, z_{m}) + \sigma^2(z_{o}, z_{f})
\]

\[
E(\beta_{ojmp}) = \frac{1}{4} \left[ \sigma(z_{o}, z_{p}) + \sigma(z_{q}, z_{p}) \right] = \frac{2\sigma(z_{o}, z_{p})}{\sigma^2(z_{p} + \sigma^2 z_{p})} = \frac{\sigma^2(z_{o}, z_{p})}{\sigma^2 z_{p}} = h^2
\]
Example calf weight a day 208

<table>
<thead>
<tr>
<th>Sire</th>
<th>Dam</th>
<th>Parental Average (X)</th>
<th>Offspring (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>371</td>
<td>378</td>
<td>374.5</td>
<td>329</td>
</tr>
<tr>
<td>285</td>
<td>255</td>
<td>270.0</td>
<td>304</td>
</tr>
<tr>
<td>318</td>
<td>323</td>
<td>320.5</td>
<td>361</td>
</tr>
<tr>
<td>355</td>
<td>318</td>
<td>336.5</td>
<td>258</td>
</tr>
<tr>
<td>349</td>
<td>373</td>
<td>361.5</td>
<td>281</td>
</tr>
<tr>
<td>324</td>
<td>333</td>
<td>328.5</td>
<td>324</td>
</tr>
<tr>
<td>379</td>
<td>410</td>
<td>394.5</td>
<td>350</td>
</tr>
<tr>
<td>359</td>
<td>376</td>
<td>367.5</td>
<td>371</td>
</tr>
<tr>
<td>323</td>
<td>310</td>
<td>316.5</td>
<td>299</td>
</tr>
<tr>
<td>448</td>
<td>352</td>
<td>400.0</td>
<td>420</td>
</tr>
</tbody>
</table>

Computation of \( h^2 \)

\[ \sum X_i = 3,469.5 \quad \quad \sum Y_i = 3,197 \]

\[ \sum X_i^2 = 1,218,194.7 \quad \quad \sum X_i Y_i = 1,120,684 \]

\[ h^2 = \frac{1,120,684 - \frac{(3,469.5)(3,197)}{10}}{1,218,194.7 - \frac{(3,469.5)^2}{10}} = .7947 \]
Estimation as Response to Selection
(Realized Heritability)

• Example: Broilers
  • 1000 Birds weighted at 7 weeks of age with an average weight of 2kg
  • The best 100 birds were saved for breeding and weighted 2.4kg
  • These birds were randomly mated when mature and used to produce the next generation
  • 2,500 progeny were produced and weighted an average of 2.1kg
  • What was the realized heritability?

Computation of Realized Heritability

\[ h_R^2 = \frac{\bar{Y}_r - \bar{Y}_p}{\bar{Y}_s - \bar{Y}_p} \]
\[ h_R^2 = \frac{2.1 - 2.0}{2.4 - 2.0} \]
\[ h_R^2 = .25 \]
Examples of heritabilities

<table>
<thead>
<tr>
<th>Trait</th>
<th>H²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Back Fat Thickness Pigs</td>
<td>.7</td>
</tr>
<tr>
<td>Daily Gain</td>
<td>.3</td>
</tr>
<tr>
<td>Litter size</td>
<td>.05</td>
</tr>
</tbody>
</table>

h² generally range .1 to .6, in breeding programs .35 is considered high, .1 is considered low

General trend: traits more closely related to fitness have lower heritabilities

Why

Hint 1: what trait(s) does natural selection act on? What of the allele frequencies?

Hint 2: traits more closely related to fitness generally show heterosis and H² > h²

Problem set 4

1. Continuing from the previous problem set, Falconer (1981) reported a partially dominant gene in the mouse called pg “pygmy.” At six weeks of age, they produce the following average weight phenotypes in grams (the actual weight of the heterozygote was 12, but it was reduced to 10 for this example):

- +/+ : 14, +/pg : 10, pg/pg : 6

- If the population of mice is randomly mating with p+ = 0.3, q.pg = 0.7

- Assuming no Environmental Effects, what are the narrow and broad sense heritabilities for this trait?
Problems

2) If for a given trait the broad sense and narrow sense heritabilities are as follows, in each which would be more effective at improving the trait, a breeding program, improving management, neither or both?

- H₂=0.9, h₂=0.1
- H₂=0.1, h₂=0.1
- H₂=0.9, h₂=0.9