Heritability

Last class we discussed heritability in the broad sense (H) and narrow sense heritability (h²).
Heritability is a term that refers to the degree to which a trait is determined by genetics. To estimate heritability we partition (divide) the phenotypic variance into genetic variance and environmental variance. We reviewed three general methods to estimate heritability.

What do I want you to take home from this discussion?

Understand heritability as a concept.
Learn some practical approaches to estimate heritability.
   - Wright’s method
   - Line-mean basis
   - Parent-offspring regression
Understand practical uses of heritability.
   - Context of a breeding program.
   - Context of measuring the contribution of a QTL.
Exposure to the theoretical basis of estimation techniques.
   - The science of breeding can make a contribution to science in the broader context.
   - Allows us to progress from tool users to tool developers.

Heritability in the broad sense (H) is:

\[ H = \frac{\sigma^2 G}{\sigma^2 P} \]

here is heredity determined by genotype.
degree of genetic determination.
Nature vs. Nurture.

Heritability in the narrow sense (h²) is:

\[ h^2 = \frac{\sigma^2 G_{\text{Additive}}}{\sigma^2 P} \]

heredity transmitted from parent to offspring.

Chapters in Lynch and Walsh that are relevant to heritability:

General treatment
   - Chapter 3 (especially pp 43-40, note the “breeders equation”)
   - Chapter 7 (especially pp 170-175, “the heritability concept”)
Specific methods and designs for estimating genetic variance
   - Chapter 17 (parent-offspring regression)
     Relevant to breeding programs for inbreeding and outcrossing species
   - Chapter 18 (sib analysis)
     More common to outcrossing species and long-lived perennials.
   - Chapter 19 (twins and clones)
     Clonally propagated species, outcrossing species, & long-lived perennials.
   - Chapter 20 (cross-classified designs)
     Relevant to breeding programs for inbreeding and outcrossing species
     Includes various “North Carolina” designs and diallele analysis
Why are there so many methods to estimate genetic variance and heritability?

Think about the concept of estimating a parameter. For a given population, we can calculate heritability, allele frequency, recombination frequency, or any number of genetic parameters from a data set. But, these calculations are based on a sample. The statistical tools used to estimate parameters allow us to extrapolate to larger populations and provide an estimate of our accuracy. In order to estimate parameters we need to make assumptions about the distribution of our data (the sampled population), and work within a theoretical framework for that distribution in order to calculate standard errors. We will often have more than one statistical tool to arrive at an estimation. These statistical tools will carry their own set of advantages and disadvantages.

Negative estimates of variance components.

For the exercise, we noted that some variance components were negative. This possibility is noted in Chapter 18 of Lynch and Walsh, and is basically a problem associated with small n and variance components that are close to zero. The estimates were derived using the proc varcomp default MIVQUE0 which works by solving the set of expected mean square equations. Thus arithmetic is used and may result in negative estimates when variances are very low (i.e. one low number is smaller than another…). Note that there are multiple ways to estimate variance components in proc varcomp by adding the type = method .

The SAS proc varcomp can use the following four estimation procedures for variance components:

type = type1 computes the Type 1 sum of squares for each effect, equates each MS involving only random effects to its expected value, and solves the set of equations.

MIVQUE0 is similar to type1, but is computationally simpler (and therefore is the default).

Maximum likelihood (ML) estimation uses a “W-transformation” of the expected mean squares equation and computes initial estimates using MIVQUE0. The program iterates until convergence.

Restricted Maximum likelihood (REML) similar to ML, but separates the likelihood into two parts (one with fixed effects, one without). Initial estimates are obtained using MIVQUE0, then iteration is performed until convergence for the equation that does not contain fixed effects. REML is emerging as the method of choice for genetic studies. The syntax would be proc varcomp type = REML.

The SAS procedure “proc mixed” can also be used for REML estimation. An advantage of using PROC MIXED over PROC VARCOMP is that the output will return estimates of error associated with the variance components. These estimates or error can then be used to place a standard error on our estimate of heritability. We will return to this point (below). PROC MIXED uses the following syntax:
proc mixed data=cdata covtest;
    class year rep gen;
    model hplc = gen year rep(year) gen*year gen*rep(year) / ddfm = satterth ;
    random year rep gen;
    title ‘Variance components using Proc mixed’;

cdata is the data file name following the data statement.
covtest option statement calculates standard errors.
a blank after model var = means that all affects are random. Fixed affects should be added here.
Degrees of freedom are estimated by Satterthwaite’s procedure.

Practical Considerations of Calculating Heritability from Variance Components of lines.
When using variance components from ANOVA to estimate broad sense heritability, the
practical application will determine the appropriate denominator. For example it is common to
exclude variance components due to blocks, years, and locations because it assumed that selection
will occur on means across replicate, location, and years. This practice is based on the assumption
that means will be corrected for differences between locations, blocks, and years (i.e. expressed as
deviation from the mean). The rationale for this approach is that heritability should be defined
based on the variance associated with the selection unit.

\[ BSH = \frac{\sigma^2(G)}{\sigma^2(x)} \text{ and } \sigma^2(x) = \sigma^2(P) \text{ when the selection unit is an individual.} \]

When the selection unit is not an individual, but a family, inbred line, or clone for which
replicated phenotypic data has been collected, the expression of phenotypic variation is adjusted to
represent the expected phenotypic variation among family (or clone, or inbred line) means.

Rules of thumb:
First, define selection unit.
Second, main effects of year, location, and rep are dropped from the denominator.
Third, variance estimates are adjusted for the selection unit.

For a one year, one location, randomized complete block design with r replications and n individuals
measured per plot and assuming we are selecting the best family or line, the heritability is:

\[ H_{\text{family}} = \frac{\sigma^2(G)}{\sigma^2(G) + \sigma^2(\text{error})/r + \sigma^2(\text{within family})/rn} \]

However if the goal is to select the best individual from each line

\[ H = \frac{\sigma^2(G)}{\sigma^2(G) + \sigma^2(\text{error}) + \sigma^2(\text{within family})} \]
Justification for ignoring main effects (variation due to location, year, or blocks) is based on the fundamental assumption that corrections will be made for the effects prior to using phenotypic measurement to select. This is an important point if significant main effects exist. If main effects for location, year, or blocks are small they will contribute little to the denominator.

So, the equation from lecture 2: \[
\frac{\sigma^2(G)}{\sigma^2(G) + \sigma^2(\text{error}) + \sigma^2(\text{GYL}) + \sigma^2(\text{GL}) + \sigma^2(\text{GY})}
\]

May be generalized as

\[
\frac{\sigma^2(G)}{\sigma^2(G) + \sigma^2(\text{error/rep*year*location}) + \sigma^2(\text{GYL/yr*location}) + \sigma^2(\text{GL/loc}) + \sigma^2(\text{GY/yr})}
\]

Hallauer and Miranda further generalize the equation as:

\[
H = \frac{\sigma^2(G)}{\sigma^2(G) + \sigma^2(\text{ge/loc}) + \sigma^2(\text{error/r*e})}
\]

Where \( r = \) number of reps and \( e = \) number of environments

The standard error for \( H \) is:

\[
\text{SE}(H) = \frac{\text{SE}\sigma^2(G)}{\sigma^2(G) + \sigma^2(\text{ge/loc}) + \sigma^2(\text{error/r*e})}
\]

For further discussion see:


**Uses of Variance Partitioning (Adapted from Falconer):**

<table>
<thead>
<tr>
<th>Data</th>
<th>Partition</th>
<th>Ratio Estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resemblance between relatives</td>
<td>Va:Vna + Veg + Ves</td>
<td>narrow sense heritability Va/Vp</td>
</tr>
<tr>
<td>Genetically uniform group</td>
<td>Va+Vna:Veg+Ves</td>
<td>degree of genetic determination</td>
</tr>
<tr>
<td></td>
<td>Va + Vna: Veg</td>
<td>broad sense heritability</td>
</tr>
<tr>
<td>Multiple measures</td>
<td>Vg+Veg:Ves</td>
<td>repeatability</td>
</tr>
</tbody>
</table>

\( a = \) additive \( \quad \) eg = general environment \( \quad \) es = special environment

\( na = \) non-additive

Take home message: Different calculations can be made depending on the practical application that is desired. It is important to consider the desired outcome (a relative magnitude of genetic and environmental variation in the germplasm pool vs a predictor of gain under selection).
Wright’s method:

Using some of the methods discussed in Lecture 2, we can derive estimates of the additive genetic variation. The equation $VF_2 = \frac{1}{2} A + \frac{1}{4} D + E$ was presented for the phenotypic variance of an F2 population. By adding further generations (e.g. reciprocal backcrosses) estimates of $Va$ and $Vd$ may be obtained (remember $\frac{1}{2} A = Va$ and $\frac{1}{4} D = Vd$).

\[

text{subtracting} \\
VF_2 = \frac{1}{2} A + \frac{1}{4} D + E \\
VBC = \frac{1}{4} A + \frac{1}{4} D + E \\
VF_2 - VBC = \frac{1}{4} A
\]

\[
E = (P1 + P2 + F1)/3
\]

An important take home message is that: **Estimation of narrow sense heritability requires appropriate mating (seggregating populations)**. Variance partitions for various selfing generations derived from two inbred parents are listed in Table 1. Note that as we progress beyond $F_2$ and $BC_1$ populations, where each plant is a genetically distinct individual, we must begin to consider within and between family variation as well as the total variation of the generation.

**Table 1. Variance components for selfing generations derived from two inbred parents**
(notation of Mather and Jinks)

<table>
<thead>
<tr>
<th>VarF2</th>
<th>VarBC</th>
<th>VarBC1.1+VarBC1.2</th>
<th>Var1F3 (F3 families)</th>
<th>Var2F3 (Ind w/in Fam)</th>
<th>Var3F3 (Families)</th>
<th>VarF3-tot</th>
<th>Var1F4 (bet. groups of fam)</th>
<th>Var2F4 (w/in groups)</th>
<th>Var3F4 (families)</th>
<th>VarF4-tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1/2</td>
<td>1/4</td>
<td>1/8</td>
<td>1/8</td>
<td>1/8</td>
<td>3/4</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
</tr>
<tr>
<td>D</td>
<td>1/4</td>
<td>1/4</td>
<td>1/16</td>
<td>1/16</td>
<td>1/16</td>
<td>3/16</td>
<td>1/64</td>
<td>1/32</td>
<td>1/16</td>
<td>1/16</td>
</tr>
<tr>
<td>w/in Fam</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>b</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Var-sample</td>
<td>0</td>
<td>0</td>
<td>1/n*Var2F3</td>
<td>1/n*Var2F4</td>
<td>1/n*Var3F4</td>
<td>1/n*Var3F4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Remember the difference between English quantitative geneticists and American quantitative geneticists that I hinted at in lecture 4:

English quantitative geneticists define the total phenotypic variance is $VF_2 = \frac{1}{2} A + \frac{1}{4} D + E$.

Note that additive genetic variance $Va = \sigma^2_a = \frac{1}{2} A$ and dominant genetic variance $Vd = \sigma^2_d = \frac{1}{4} D$ (the $Va$ or $\sigma^2_a$ notation is used more commonly by American quantitative geneticists)

<table>
<thead>
<tr>
<th>Expression</th>
<th>English</th>
<th>U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Genetic VarF3</td>
<td>$= 3/4 A + 3/16 D = 3/2 \sigma^2_a + 3/4 \sigma^2_d$</td>
<td>$= 3/4 A + 3/16 D = 3/2 \sigma^2_a + 3/4 \sigma^2_d$</td>
</tr>
<tr>
<td>Bet. Families VarF3</td>
<td>$= 1/2 A + 1/16 D = \sigma^2_a + 1/4 \sigma^2_d$</td>
<td>$= 1/2 A + 1/16 D = \sigma^2_a + 1/4 \sigma^2_d$</td>
</tr>
<tr>
<td>W/n Families VarF3</td>
<td>$= 1/4 A + 1/8 D = 1/2 \sigma^2_a + 1/2 \sigma^2_d$</td>
<td>$= 1/4 A + 1/8 D = 1/2 \sigma^2_a + 1/2 \sigma^2_d$</td>
</tr>
</tbody>
</table>
Table 2. Variance components for selfing generations derived from two inbred parents using alternative notation.

<table>
<thead>
<tr>
<th></th>
<th>VarA</th>
<th>VarD</th>
<th>VarE(w/in)</th>
<th>VarE(b)</th>
<th>VarE(samp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VarF2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>VarBC</td>
<td>1/2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>VarBC1.1+VarBC1.2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Var1F3</td>
<td>1/2</td>
<td>1/4</td>
<td>0</td>
<td>1</td>
<td>1/n*Var2F3</td>
</tr>
<tr>
<td>Var2F3</td>
<td>1/2</td>
<td>1/2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>VarF3-tot</td>
<td>3/2</td>
<td>3/4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Var1F4</td>
<td>1</td>
<td>1/16</td>
<td>0</td>
<td>0</td>
<td>1/n*Var2F4</td>
</tr>
<tr>
<td>Var2F4</td>
<td>1/2</td>
<td>1/8</td>
<td>0</td>
<td>0</td>
<td>1/n*Var3F4</td>
</tr>
<tr>
<td>Var3F4</td>
<td>1/4</td>
<td>1/4</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>VarF4-tot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Simplifying assumption: \( E_2 = \frac{1}{n}E_1 \) where \( E_1 = E_w \) and \( E_2 = E_b + \frac{1}{n}E_w \)

**Question:**
Can you explain how the regression equations in Foolad and Jones (1992) were derived?

**Parent-offspring regression.** This method of estimating heritability is intuitive, we ask the question “How well does the performance of parents predict the performance of progeny?”

<table>
<thead>
<tr>
<th>Parents</th>
<th>Progeny</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

High heritability trait

<table>
<thead>
<tr>
<th>Parents</th>
<th>Progeny</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Low heritability trait
The slope of the regression line and the correlation coefficient tell us something about the genetic basis of the trait.

We can break the process down as follows:

Does the regression line have a slope that is equal to zero?

- no, there is significant genetic variation for the trait.

What is the slope?

- an estimate of heritability.

What is the correlation coefficient?

- also an estimate of heritability (depending on experimental design).

\[ Y' = mX + b \]

regression line, \( Y \) and \( X \) are variables and \( Y' \) is an estimate of \( Y \) based on a value of \( X \)

\[ m = \frac{\sum(X_i - X_{ave})(Y_i - Y_{ave})}{\sum(X_i - X_{ave})^2} = \text{change in } X \]  
\[ \text{change in } Y \]

\[ b = \text{the } Y \text{ intercept} \]

\[ r = \frac{\text{Cov}(X,Y)}{\text{Sdev}_x \times \text{Sdev}_y} \]

The correlation coefficient "r" measures how closely two sets of data are associated. It is without units and as the limits of –1.0 to +1.0. the regression coefficient, \( b \), and the correlation coefficient, \( r \), always have the same sign. The correlation coefficient of \( Y \) on \( X \) is defined as the linear change of \( Y \) in standard deviations, for each increase of one standard deviation in \( X \).

\[ \text{Cov}(X,Y) = \sum(X_i - X_{ave})(Y_i - Y_{ave}) \]

\[ \frac{n - 1}{n} \]

For PO regression  
\[ h^2 = r = \frac{\text{Cov}(P,O)}{(\text{Var } P \times \text{Var } O)^{1/2}} \]

Note that \( r \) and \( m \) are related by:

\[ m = r(\text{Sdev}_x / \text{Sdev}_y) \]

When the variances of \( X \) and \( Y \) are equal, \( m = r \) and either can be used to estimate heritability. If the variances are unequal, standardized variables (sample has a mean of zero and a standard deviation of 1) can be used to insure that \( m = r \).

**Corrections for inbreeding:**

Parent offspring regression is used to estimate narrow sense heritability, but as described in Foolad and Jones (1992) the numerator often contains some contribution from variation due to dominance. In practice, regression coefficients are corrected based on the relationship between relatives used for the analysis i.e. the inbreeding coefficient.
The coefficient of relationship between full sibs is \( \frac{1}{2} \) so \( h^2 = 2^s m \) and the coefficient of relationship between half sibs is \( \frac{1}{4} \) so \( h^2 = 4^s m \). There is considerable argument in the literature about how to correct for mating system (i.e. selfing populations). For example, Smith and Kinman suggest that \( 1/(2fpo) \) (where \( fpo = \) is the coefficient of coancestry = \( \theta \) in Cockerham’s notation) is the appropriate correction.

Suggested Corrections for narrow sense heritability (Smith and Kinman)

<table>
<thead>
<tr>
<th>P-O generation</th>
<th>fpo</th>
<th>( \frac{1}{2}fpo ) * m</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 to F2</td>
<td>1/2</td>
<td>m</td>
</tr>
<tr>
<td>F2 to F3</td>
<td>3/4</td>
<td>2/3 * m</td>
</tr>
<tr>
<td>F3 to F4</td>
<td>7/8</td>
<td>4/7 * m</td>
</tr>
<tr>
<td>F4 to F5</td>
<td>15/16</td>
<td>8/15 * m</td>
</tr>
<tr>
<td>F5 to F6</td>
<td>31/32</td>
<td>16/31 * m</td>
</tr>
</tbody>
</table>

From Smith, J.D. and M.L. Kinman.1965 The use of parent-offspring regression as an estimator of heritability. Crop Science 5:595-596.

There is not complete agreement in the literature on the subject of corrections for inbreeding!

Nyquist (1991, Crit. Reviews in Plant Sciences 10(3)235-322 CRC press) argues that the correction (above) should be \( h^2 = \frac{m}{1 + Fp(1 - m)} \) where \( Fp = \) the inbreeding coefficient of parent \( P \) (\( Fp = 1 - (1/2)^{n-1} \) for selfing). For a cross derived from two inbred lines, the \( F_2 \) generation is defined as the reference generation (H-W equilibrium, etc…) and \( Fp = Ft = F_0 = 0 \)

Suggested Corrections for narrow sense heritability (Nyquist)

<table>
<thead>
<tr>
<th>P-O generation</th>
<th>Fp</th>
<th>correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>F2 to F3</td>
<td>0</td>
<td>( h^2 = m )</td>
</tr>
<tr>
<td>F3 to F4</td>
<td>1/2</td>
<td>( h^2 = \frac{m}{1 + Fp(1 - m)} )</td>
</tr>
</tbody>
</table>

The relationship between the inbreeding coefficient \( Fp = 1 - (1/2)^{n-1} \) and coefficient of ancestry \( fpo = (1 + Fp)/2 \) under selfing:

\[ \begin{align*}
\text{fpo} = \frac{1}{2} & \quad Fp = 0 \\
\text{fpo} = \frac{3}{4} & \quad Fp = \frac{1}{2} \\
\text{fpo} = \frac{7}{8} & \quad Fp = \frac{3}{4}
\end{align*} \]

The reference population should be defined as last outbreeding pop (i.e. \( F_2 \)) and this lack of definition by Smith and Kinman may explain difference in correction.
Covariances of relatives and Coefficients of ancestry
The discussion of corrections for inbreeding in P-O regression introduced some new concepts: coefficients of inbreeding and coancestry. These coefficients give a mathematical description of relationships among relatives. I would like to introduce this topic more formally as it will also relate to our discussion of genotypic variation.

Definitions (note that we will see these again and again and …):

\[ \theta = \text{fpo} = \frac{1 + F_p}{2} \] (coefficient of ancestry also called coancestry coefficient): the probability that a random allele from X is identical by descent to a random allele from Y.

\[ F_p = (1 - \frac{1}{2^{n-1}}) \] (inbreeding coefficient): the probability that two alleles are identical by descent i.e. homozygous. For selfing \( F_p = (1 - \frac{1}{2^{n-1}}) \) where \( n \) = generation number.

\( F_t \) = probability that alleles are identical in a common ancestor (for an F2 with two inbred parents, this is \( = 0 \))

Equations for estimating the covariance between relatives:

Falconer:
\[ \text{Cov}_{pq} = r(\sigma^2(a)) + u(\sigma^2(d)) \]

\[ \text{Cov}_{pq} = 2f_{pq}(\sigma^2(a)) + (f_{ac}f_{bd} + f_{ad}f_{bc})(\sigma^2(d)) \]
where \( a \) & \( b \) are parents of \( p \) and \( c \) & \( d \) are parents of \( q \).

Cockerham developed the following more general equation that can be used for inbreeding (for two alleles of equal frequency):
\[ \text{Cov}_{gg'} = (1 + F_t)(\sigma^2(a)) + [(1 + F_t)(1 - F_g)(1 - F_{g'})(1 - F_t)](\sigma^2(d)) \]

\[ \text{Cov}_{F2:F3} F_t = 0 \] where \( F_g = (1 - \frac{1}{2})^{2-1} \) and \( g' = (1 - \frac{1}{2})^{3-1} \)

\[ \text{Cov}_{F3:F4} F_t = (1 - \frac{1}{2})^{2-1} \] where \( g = (1 - \frac{1}{2})^{3-1} \) and \( g' = (1 - \frac{1}{2})^{4-1} \)

etc…

Cockerham’s equations have been used to calculate the covariances listed in Table 3. Consider this table in the context of the Foolad and Jones (1992) paper. We arrive at the same relationships from Wright’s methods, from considering coancestry, and from P-O regression.

Table 3. Covariance among relatives.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>D</th>
<th>VarA</th>
<th>VarD</th>
</tr>
</thead>
<tbody>
<tr>
<td>CovF2:F3</td>
<td>1/2</td>
<td>1/8</td>
<td>1</td>
<td>1/2</td>
</tr>
<tr>
<td>CovF3:F4 (F3 fam:F4 fam)</td>
<td>1/2</td>
<td>1/32</td>
<td>1</td>
<td>1/8</td>
</tr>
<tr>
<td>CovF3:F4 (F3 ind:F4 fam)</td>
<td>1/4</td>
<td>1/16</td>
<td>1/2</td>
<td>1/4</td>
</tr>
</tbody>
</table>

“Inbreeding introduces difficulties in expressing the covariances of relatives for a quantitative character. The situation is unmanageable for a completely general genetic model. Even with only additive and dominance effects (no epistasis), procedures developed previously for the interpretation of genetic variance content of covariances of relatives from self-fertilization were restricted to gene frequencies of one-half.” C.C. Cockerham. 1983 Covariances of relatives from self-fertilization. Crop Science 23:1177-1180
Conclusions:

We have not exhausted methods to estimate heritability. For example, another way to estimate additive genetic variance is to estimate realized heritability in selection experiments (and we will come back to this).

I would like you to take home some messages from this discussion:

Mating systems, mating designs, reproductive capacity of plant, etc… must be considered in choosing approach for estimating heritability. Some examples?

There is not agreement in the literature about all aspects of estimations. This is still an area of investigation (both theoretical and practical).

Most estimates of heritability are biased upward.

It is important to understand the relationships among populations and individuals within populations. These relationships can be described numerically using the coancestry coefficient and inbreeding coefficient.

Practical “rules of thumb” for breeding applications

<table>
<thead>
<tr>
<th>Heritability</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>&gt; 0.5</td>
</tr>
<tr>
<td>Moderate</td>
<td>0.25 to 0.5</td>
</tr>
<tr>
<td>Low</td>
<td>&lt; 0.25</td>
</tr>
</tbody>
</table>

single plant selection in later generations
population or family based selection in early generations

Review of analytical exercise

Which methods could we use to estimate heritability from the data sets?
What in the SAS output do we need to use?
Conclusions?

Discussion of papers in the context of this discussion

Foolad and Jones (1992). The paper is more than an example of P-O regression to estimate heritability of a trait. There is a theoretical connection between variance partitioning, covariance among relatives, and the regression analysis.

Cotterill, P. P. (1987). This paper presents a pretty clear discussion of the importance of defining the selection unit and the appropriate estimates from line means. There is also a discussion of correcting estimates based on the genetic relationships.