Introduction to Quantitative Genetics

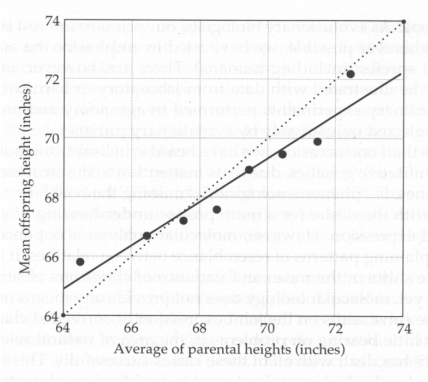
Historical Background

- Quantitative genetics is the study of continuous or quantitative traits and their underlying mechanisms.
- The main principals of quantitative genetics developed in the 20th century was largely in response to the rediscovery of Mendelian genetics.
- Mendelian genetics in 1900 centered attention on the inheritance of discrete characters, e.g., smooth vs. wrinkled peas, purple vs. white flowers.
- This focus was in stark contrast to the branch of genetic analysis by Sir Francis Galton in the 1870's and 1880's who focused on characteristics that were continuously variable and thus, not clearly separable into discrete classes.

Historical Background

- A contentious debate ensued between Mendelians and Biometricians regarding whether discrete characteristics have the same hereditary and evolutionary properties as continuously varying characteristics.
  - The Mendelians (led by William Bateson) believed that variation in discrete characters drove evolution through mutations with large effects
  - The Biometricians (led by Karl Pearson and W.F.R. Weldon) viewed evolution to be the result of natural selection acting on continuously distributed characteristics.
- This eventually led to a fusion of genetics and Charles Darwin's theory of evolution by natural selection: main principals of quantitative genetics, developed independently by Ronald Fisher (1918) and Sewall Wright (1921), arguable the two most prominent evolutionary biologists.

Historical Background



Biometricians

The height vs. pea debate (early 1900s)

Mendelians

Do quantitative traits have the same hereditary and evolutionary properties as discrete characters?

6

(Figure from Peter Visscher)

Historical Background

- Interestingly, Galton's methodological approaches to continuously distributed traits marked the founding of the Biometrical school, which is what many consider to be the birth of modern statistics (see Steve Stigler's book *The history of statistics*)
- Karl Pearson was inspired greatly by Galton, and he went on to develop a number of methods for the analysis of quantitative traits.
- We will talk more about Galton and Pearson later on in this course.

Introduction to Quantitative Genetics

- One of the central goals of quantitative genetics is the quantification of the correspondence between phenotypic and genotypic values
- It is well accepted that variation in quantitative traits can be attributable to many, possibly interacting, genes whose expression may be sensitive to the environment
- Quantitative geneticists are often focused on partitioning the phenotypic variance into genetic and nongenetic components.
- Classical quantitative genetics started with a simple model: Phenotype $=$ Genetic Value $+$ Environmental Effects.

Partitioning Phenotypic Variance

- A standard approach for partitioning the phenotypic variance is to model the phenotypic value of an individual Y to be the sum of the total effect of all genetic loci on the trait, which we will denote as the genotypic value G, and the total environmental effect or environmental deviation, which we will denote as E, such that:

$$Y = G + E$$

- From the properties of covariances between random variable, we can show that the covariance between phenotype and genotype values can be written as

$$cov(Y, G) = \sigma_{Y,G} = \sigma^2_G + \sigma_{G,E}$$

Broad Sense Heritability

- And the squared correlation coefficient is

$$\rho^2(Y,G) = \left(\frac{\sigma_{Y,G}}{\sigma_Y \sigma_G}\right)^2 = \frac{\left(\sigma_G^2 + \sigma_{G,E}\right)^2}{\sigma_Y^2 \sigma_G^2}$$

- Note that if we assume that there is no genotype-environmental covariance, i.e., if $\sigma_{G,E} = 0$, then $\rho^2(Y,G)$ reduces to $\frac{\sigma_G^2}{\sigma_Y^2}$
- The quantity $H^2 = \frac{\sigma_G^2}{\sigma_Y^2}$ is generally referred to as the **broad sense heritability** and it is the proportion of the total phenotypic variance that is genetic
- Will cover heritability and partitioning phenotypic variances into components in much greater depth in later lectures!

Broad Sense Heritability

- Note that that the covariance between genotypic values and environmental deviations causes the phenotype-genotype covariance to deviate from σ^2_G .
- What impact does a positive (or negative) covariance between genotype and environment have on the correlation between genotype and phenotype?

Characterizing Single Locus Influence on a Phenotype

- The genotype value G in the phenotype partition $Y = G + E$ can be viewed as the expected phenotype contribution (for a given set of genotypes) resulting from the joint expression of all of the genes underlying the trait.
- For a multilocus trait, G can potentially be a very complicated function.
- Let's consider for now the contribution of only a single autosomal locus to the phenotype. For simplicity, we will assume that the locus has two alleles.

Characterizing Single Locus Influence on a Phenotype

- The genotypic values for a bi-allelic locus with alleles A_1 and A_2 can be represented as follows:

$$\text{Genotype Value} = \begin{cases} 0 & \text{if genotype is } A_2A_2 \\ (1+k)a & \text{if genotype is } A_1A_2 \\ 2a & \text{if genotype is } A_1A_1 \end{cases} \quad (1)$$

- Note that we can transform the genotype values so that the homozygous A_2A_2 haves a mean genotype value of 0 by subtracting the mean genotypic value for A_2A_2 from each measure.

Characterizing Single Locus Influence on a Phenotype

$$\text{Genotype Value} = \begin{cases} 0 & \text{if genotype is } A_2A_2 \\ (1+k)a & \text{if genotype is } A_1A_2 \\ 2a & \text{if genotype is } A_1A_1 \end{cases} \quad (2)$$

- $2a$ represents the difference in mean phenotype values for the A_2A_2 and A_1A_1 homozygotes, and k is a measure of **dominance**.
- If $k = 0$, then the alleles A_1 and A_2 behave in an **additive** fashion
- if $k > 0$ implies that A_1 exhibits **dominance** over A_2. If $k = 1$ then there is complete dominance. Similarly, $k < 0$ implies that A_2 exhibits dominance over A_1.
- The locus is said to exhibit overdominance if $k > 1$, since then mean phenotype expression of the heterozygotes exceeds that of both homozygotes. The locus exhibits underdominance when $k < -1$.

Example: Fecundity (Litter Size) in Merino Sheep

- Litter size in sheep is known to be polygenic
- In Booroola Merino sheep, however, litter size is largely determined by a single polymorphic locus in the Booroola gene (Piper and Bindon 1988).
- The Booroola fecundity gene (FecB) located on sheep chromosome 6 increases ovulation rate and litter size in sheep and is inherited as a single autosomal locus.
- For the three genotypes at this gene, the mean litter sizes for 685 recorded pregnancies are

Genotype	Mean Litter Size
bb	1.48 (s.e. =.09)
Bb	2.17 (s.e. =.08)
BB	2.66 (s.e. =.09)

- Calculate an estimate of the dominance coefficient k for this locus.
- Is there significant evidence of a dominance effect?

Example: Fecundity (Litter Size) in Merino Sheep

Genotype	Mean Litter Size
bb	1.48 (s.e. =.09)
Bb	2.17 (s.e. =.08)
BB	2.66 (s.e. =.09)

- To estimate a, we have that the average of the difference between the mean genotype values of BB and bb is: $2\hat{a} = 2.66 - 1.48 = 1.18$, so $\hat{a} = 0.59$
- Can use the difference between the genotype mean values for bb and Bb to obtain a crude estimate for k. We have that $(1 + \hat{k})\hat{a} = 2.17 - 1.48 = .69$. Substituting our estimate for $\hat{a}$ and solving for $\hat{k}$, we have that $\hat{k} = .1695$

Example: Fecundity (Litter Size) in Merino Sheep

- To determine if there is a significant dominance effect, i.e., if k is significantly different from 0, we can perform a simple hypothesis test.
- Observed Scaled Genotype Values

Genotype	Scaled Mean Litter Size
bb	0 (s.e. =.09)
Bb	0.69 (s.e. =.08)
BB	1.18 (s.e. =.09)

- Expected Scaled Genotype Values under null hypothesis $H_0 : k = 0$

Genotype	Expected Scaled Mean Litter Size
bb	0 (s.e. =.09)
Bb	0.59 (s.e. =.08)
BB	1.18 (s.e. =.09)

Example: Fecundity (Litter Size) in Merino Sheep

- Under the null hypothesis of no dominance effect, the (scaled) expected mean heterozygote genotype value is estimated to be .59, while the observed value is 0.69. . Using the standard error of the mean heterozygote genotype value, a z-statistic for the observed data under the null hypothesis is

$$z = \frac{0.69 - 0.59}{0.08} = \frac{0.1}{0.08} = 1.25$$

  ν

- Using a two-sided test, the p-value for the z-statistic is around 0.21. So there is no significant evidence of a dominance effect at the .05 level.

Another Representation of Genotypic Values

- We previously represented the genotypic values for a bi-allelic locus relative to the mean genotype value of the lower homozygote.
- Alternatively, the genotype values can be represented relative to the average of the mean genotype values for the two homogygotes as follows:

$$\text{Genotype Value} = \begin{cases} -a & \text{if genotype is } A_2A_2 \\ d & \text{if genotype is } A_1A_2 \\ a & \text{if genotype is } A_1A_1 \end{cases}$$

- What values of d would correspond to completely additive model? Dominance? Complete dominance? Overdominance?
- Note that the previous scale can be completely recovered by adding a to all three measures on this new scale and letting $d = ak$.