

Estimating the dominance variance (V_D) from the correlation between full siblings

Anth/Biol 5221, 4 December 2009

A single-locus model of overdominance was used in lecture on December 2nd to show that there can be genetic variance (V_G) that is not additive and does not cause offspring to resemble their parents. (This model is developed fully in the Quantitative Characters handout.) At the end of that lecture we noted that although the dominance variance (V_D) does not contribute to the resemblance of parents and offspring, it *does* contribute to the resemblance of full siblings: the covariance of parent and offspring phenotypes is $\frac{1}{2}V_A$, but the covariance of full-sib phenotypes is $\frac{1}{2}V_A + \frac{1}{4}V_D$ (Gillespie's equation 6.23 on page 164, where $r = \frac{1}{2}$ and $r_2 = \frac{1}{4}$). The second term arises from the fact that full siblings may carry *two* alleles identical by descent at a given locus, whereas an outbred parent and offspring may carry at most *one* allele IBD. This difference in the distribution of allele-sharing by the two kinds of first-degree relatives can be used to "see" and estimate the dominance variance under circumstances where otherwise it would be *invisible*!

Recall that under this model, the homozygotes **AA** and **aa** have phenotypic values of $x=0$, while heterozygotes **Aa** have $x=1$. Mating is random, so the genotypic frequencies are $f(\text{AA})=p^2$, $f(\text{Aa})=2pq$, and $f(\text{aa})=q^2$. The *family* frequencies are as set out in the mating table below. Each row shows one of the six mother/father combinations, ignoring whether the mom's or dad's genotype comes first (*i.e.*, "**AA** x **Aa**" combines those matings where mom is **AA** and dad is **Aa**, with those where dad is **AA** and mom is **Aa**, since both kinds of families produce half **AA** and half **Aa** offspring, as indicated in the columns headed "offspring genotypes & phenotypes").

family type	family frequency	mid-parent phenotype	offspring genotypes & phenotypes		
			AA	Aa	aa
AA x AA	p^4	0	1		all 0
AA x Aa	$4p^3q$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	mixed
AA x aa	$2p^2q^2$	0		1	all 1
Aa x Aa	$4p^2q^2$	1	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
Aa x aa	$4pq^3$	$\frac{1}{2}$		$\frac{1}{2}$	mixed
aa x aa	q^4	0		1	all 0

The three kinds of "mixed" families (where at least one parent is a heterozygote, so the midparent phenotype is $\frac{1}{2}$ or 1) all produce equal numbers of offspring with phenotypic values of 0 and 1. When $p = q = \frac{1}{2}$, the families where both parents are homozygotes (with phenotypic values of 0) also produce, collectively, equal numbers of 0 and 1 offspring. Thus when $p = \frac{1}{2}$, there is *no* statistical relationship between the phenotypes of parents and the phenotypes of their offspring. On average, parents with $x=0$ produce offspring with $x=0$ and $x=1$ equally often, as do parents with $x=1$. In other words, you can't predict the phenotypes of the offspring from those of their parents – not even a little! When $p = \frac{1}{2}$, the heritability of the trait x is 0. This is shown graphically in the "spot-o-gram" at the top of the next page, which was also shown in lecture.

The slope of the regression line describing the relationship between offspring phenotypes and parental phenotypes is zero, whether we represent the parents by their average (midparent) values, as shown here, or represent them individually. In the latter case, parents with $x=0$ have equal numbers of offspring with $x=0$ and $x=1$, and so do parents with $x=1$. So no matter how we look at it, there's *no heritability* of trait x .

This is also what we would see if there were no genetic variance of any kind – that is, if all of the phenotypic variance were *environmental* variance (V_E). In that case an offspring's phenotype would also be uncorrelated with its parent's phenotype. But in fact this model is one with *no* environmental variance. Could we tell, just from observing the population at the *phenotypic* level, that the variance is non-additive *genetic* variance (dominance variance, V_D), and not V_E ?

It turns out that we can see the dominance variance by looking at the relationships between the phenotypes of *full siblings*. If the variance were all environmental, then they too would be uncorrelated with each other. But if some of the variance is non-additive genetic variance, then they will tend to resemble each other and this resemblance can be used to estimate V_D .

From the mating table on the previous page, we can write down the probabilities that an offspring of phenotype 0 or 1 has a sibling of phenotype 0 or 1. The mixed families always

		first offspring	
		0	1
second offspring	1	$p^3q + p^2q^2 + pq^3$	$p^3q + p^2q^2 + pq^3 + 2p^2q^2$
	0	$p^3q + p^2q^2 + pq^3 + p^4 + q^4$	$p^3q + p^2q^2 + pq^3$

contribute *equally* to each of these four possibilities, but the all-0 and all-1 families do not. Thus (0,0) and (1,1) sibships are more common than (0,1) and (1,0) sibships, and as a consequence there is a positive covariance (hence correlation and regression) among siblings, even though there is *no* covariance (hence *no* correlation or regression) between parents and offspring!

The spot-o-gram on the right shows the distribution of full-sib pairs when $p = 0.5$. More than half of the sib pairs (62.5%) have matching phenotypes (0,0 or 1,1), and fewer than half (37.5%) have unlike phenotypes (0,1 or 1,0). The covariance of sibling phenotypes is 0.0625. There is no additive variance, because there's no offspring-parent regression (as shown above). So the expression in the first paragraph (previous page) implies that $\text{cov}(\text{sib}, \text{sib}) = \frac{1}{4}V_D = 0.0625$. In other words, $V_D = 0.25$, which is equal to the phenotypic variance V_P . Thus from simple phenotypic measurements, and knowing nothing about the underlying genotypes, we could infer that all of the variance is genetic and none of it is environmental, even though there is no correlation between offspring and parent phenotypes!

