Introduction to Evolutionary Quantitative Genetics

SBM

February 23, 2018

This week we will discuss evolutionary applications of quantitative genetics. These models are distinct from the population genetic models we've looked at so far in that they focus on the evolution of 'quantitative' rather than categorical traits. Things like body mass, height, limb length, fecundity, etc. rarely fall into discrete categories and are typically thought to be controlled by genes at many loci. Rather than explicitly model changes in the frequencies of particular genes, we focus on partitioning phenotypic variation into more generic "genetic" and "envrionmental" components. This branch of modeling has its origins in breeding studies attempting to improve yields from plants and livestock. In the 1980's and 1990's a number of researchers realized that the methods that breeders had developed for predicting the results of artificial selection could also be used to measure selection in wild populations and predict evolutionary trajectories, without knowing anything about the genes involved (at least for a few generations).

To give you an overview of this week's short tour of quantitative genetics, we will start by building the explicit genetic model for a quantitative character and then use this framework for understanding and predicting evolution in these traits. Along the way we will talk about the 'infinitessimal model,' 'heritability', 'selection differentials,' the 'breeder's equation,' and the response to selection. We will also talk about the Price equation, Fisher's fundamental theorem of natural selection, and what happens when we have multiple correlated traits.

The starting point for this sort of thinking is to assume that there are a large number of unlinked loci each contributing some small amount to the phenotype. That is, we will assume that there are N loci and that each locus contributes additively to the phenotype. Ultimately, we will write the phenotype as $z = g + \epsilon$ where z is the phenotype, g is the genetic contribution to phenotype, and ϵ is the 'environmental' contibution. We assume throughout that the mean and variance of the envrionmental effect are 0 and V_e respectively and that the genetice and environmental effects are independent. Extensions of this framework allow for epistasis, maternal effects, and genotype x environment interactions.

Single locus

Before we get into applications of quantitative genetics, I think it is a good idea to see how these models are built up from the single-locus models we've worked with all quarter. To make the connection, we start by saying that the genetic contribution to phenotype of locus i is g_i which takes on values

$$g_i = \begin{cases} +\gamma_i & \text{AA} \\ \delta_i & \text{Aa} \\ -\gamma_i & \text{aa} \end{cases}$$

So the homozygotes at locus i get $\pm\gamma_i$ added to their phenotype. Heterozygotes at locus i get δ_i, which is 0 if there is no dominance, between 0 and γ_i if there is dominance and is greater than γ_i if there is overdominance. For simplicty, let's assume that there is no dominance, i.e. in all of the following $\delta_i = 0$.

Let p_i be the frequency of allele A in the population and $q_i = 1 - p_i$ is the frequency of allele a. Let's assume that the population is at Hardy-Weinberg equilibrium so the genotype frequencies are $p_i^2, 2p_iq_i, q_i^2$ for AA, Aa, and aa respectively. Now we'll calculate the mean and variance in phenotype in the population contributed by that locus. To make the notation a little less cumbersome, I am (temporarily) dropping the subscript i which we will use again when we get to more than 1 locus.

The mean genotype (and phenotype, since $E(\epsilon) = 0$) is given by

(1)

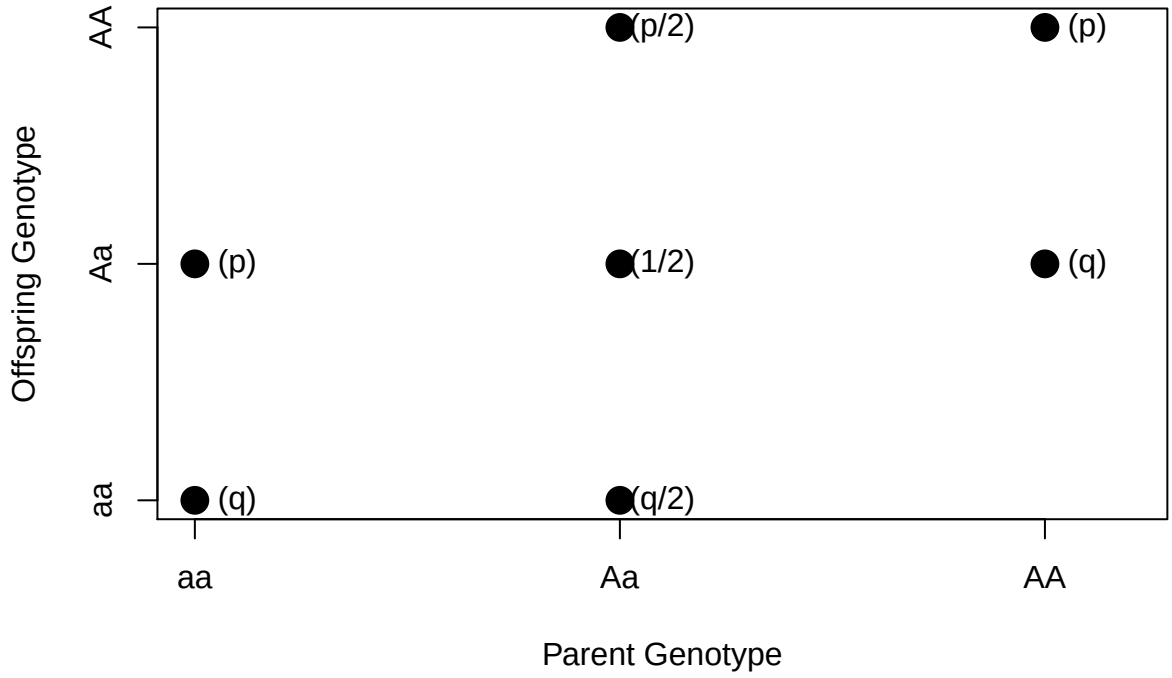
$$m = E(g) = p^2(\gamma) + 2pq(0) + q^2(-\gamma) = (p^2 - q^2)\gamma = (p - q)\gamma$$

The variance is most easily calculated using the handy fact that $V(x) = E(x^2) - E(x)^2$. Here, $E(g_i^2) = (p_i^2 + q_i^2)\gamma^2$. Subtracting $E(g_i)^2$ and simplifying, we get

(2)

$$V(g_i) = 2p_i q_i \gamma^2$$

Next, let's think about the relationship between parent and offspring phenotypes in this single locus model. The points in the plot below show the possible offspring genotypes for each parent genotype. The things in parentheses indicate the frequency with which these offspring occur, when mating is random and we have averaged over all other parent genotypes. To make this more concrete, consider the frequency with which AA parents make Aa offspring: with probability p^2 they mate with an AA parent and all offspring are AA and with probability $2pq$ they mate with a Aa parent and half of their offspring are AA. So, overall, given that one parent is AA, the frequency of AA offspring is $p^2 + 2pq/2 = p^2 + pq = p(p + q) = p$. It is not hard (though sort of tedious) to work out the rest.



The upshot of this figure is that it is clear that parent and offspring genotypes are not independent; clearly knowing one parent's genotype tells you something about the likely genotypes of offspring which is the definition of non-independence. And, of course, this is biologically obvious. In previous problems where we had just one variable, we summarized them in terms of the mean and variance. Now, we have two variables - the parent and offspring genotypes (or phenotypes). To summarize the relationship between them, we will introduce a new idea - the covariance.

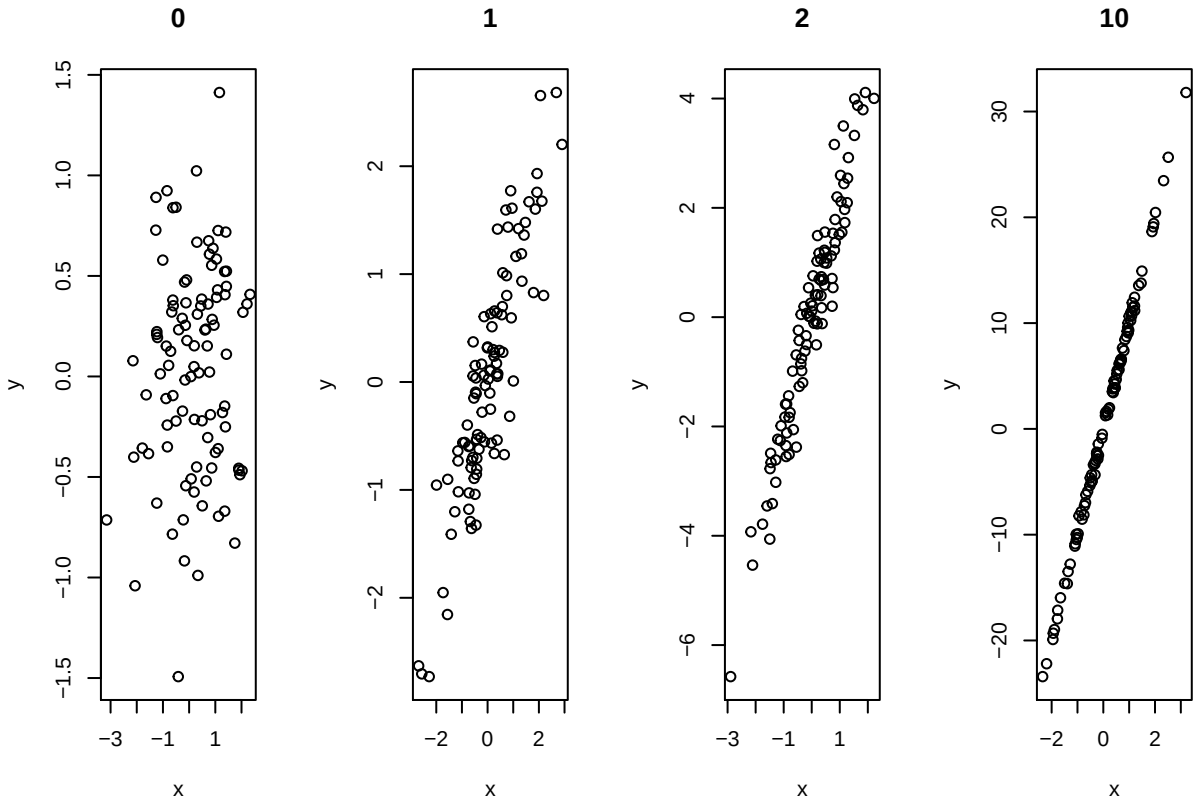
A mathematical aside

Covariance measures the linear dependence between two variables. Specifically, the covariance between two random variables is defined as

(3)
$$Cov(X,Y) = E[(X-\bar{x})(Y-\bar{y})]$$

where $\bar{x}$ and $\bar{y}$ are the mean of X and Y respectively. If you substitute X in for Y in this definition, you end up with the variance of X - hence the name, covariance. To build some intuition for covariance, here's some code that simulates two random variables with different amounts of covariance

```
n=100 #number to sample
C<-c(0,1,2,10)

par(mfrow=c(1,4))
for (i in 1:4){
  x<-rnorm(n,0,1)
  y<-(C[i]*x+rnorm(n,0,.5))

  plot(x,y,main=C[i])
}
```



we started this aside by saying that the covariance measures the linear dependence between two variables. Let's see what this means: Say that $Y = a+bX+\epsilon$ where the last term is noise that has zero mean and is independent of X. So $\bar{y} = a + b\bar{x}$ and plugging this into (3), we get $Cov(X,Y) = E[(X-\bar{x})(a+bX+\epsilon-(a+b\bar{x}))] = E[(X-\bar{x})(bX-b\bar{x})] = bVar(X)$. So if we want to recover the slope, we divide the covariance by the variance in x, i.e. $b = Cov(X,Y)/Var(X)$.

Here are a few handy facts about covariance: (@fact1)

$$Cov(aX, bY) = abCov(X, Y)$$

(@fact2)

$$Cov(X, Y + Z) = Cov(X, Y) + Cov(X, Z)$$

(@fact3)

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

Another way of thinking about the covariance between two variables is that it is the part the two variables have in commong. That is, if $U = X + Y$ and $V = X + Z$, where X, Y, and Z are all independent, the covariance between U and V is the variance in X, i.e. $Cov(U, V) = Cov(X + Y, X + Z) = Cov(X, X) + Cov(X, Z) + Cov(Y, X) + Cov(Y, Z) = Cov(X, X) = Var(X)$ ___

Single locus (Cont'd)

Coming back to the parent-offspring relationship, let's figure out the covariance, so that we can figure out the slope of the line relating parent and offspring *phenotypes*. We are going to use the foundational assumtion that the phenotype is $z = g + \epsilon$ and that the environmental contribution, ϵ, is independent of the genetic contribtuion, g. In addition, we will assume that the envrionmental components of parents ϵ_p and offspring ϵ_o are independent as well. These assumptions mean that $Cov(g, \epsilon) = 0$ and $Cov(\epsilon_p, \epsilon_o) = 0$. So, let's evaluate the phenotypic covariance between parents and offspring, $Cov(z_p, z_o)$. Plugging in the genetic and envrironmental contributions, we get $Cov(z_p, z_o) = Cov(g_p + \epsilon_p, g_o + \epsilon_o)$. Using handy fact (@fact2) this becomes $Cov(z_p, z_o) = Cov(g_p, g_o) + Cov(\epsilon_p, \epsilon_o) + Cov(g_p, \epsilon_o) + Cov(\epsilon_p, g_o)$. All of the terms on the right side except the first are zero by assumption, leaving, $Cov(z_p, z_o) = Cov(g_p, g_o)$.

To find $Cov(g_p, g_o)$, the easiest thing to do is to use handy fact (@fact3), calculate $E(g_p g_o)$ and subtract off $E(g_p)E(g_o)$ since we already found $E(g_p) = \gamma(p - q)$ and if there is no selection, the mean for offspring must be the same so that $E(g_p)E(g_o) = \gamma^2(p - q)^2$. To find $E(g_p g_o)$ we write down each possible combination of g_p and g_o, multiply by the frequency with which they occur, and sum up to get the mean. There are alot of terms and I find it easiest to avoid screwing up if I do this systematically. So we'll start from the aa parent, do all of it's offspring, then the Aa parent, etc. Most of the terms will end up being 0 because we've assumed there's no dominance, but I will leave them in the initial calculation so you can see how we would deal with $\delta \neq 0$.

$$E(g_p g_o) = (-\gamma)(-\gamma)q^3 + (-\gamma)(0)q^2 p + (0)(-\gamma)(2pq)(q/2) + (0)(0)(2pq)(1/2) + (0)(\gamma)(2pq)(p/2) + (\gamma)(0)p^2 q + (\gamma)(\gamma)p^3$$

And because of all those zeros this boils down to $E(g_p g_o) = \gamma^2(q^3 + p^3)$. To get the covariance we subtract off $E(g_p)E(g_o)$ to get $Cov(g_p, g_o) = \gamma^2(q^3 + p^3) - \gamma^2(p - q)^2 = pq\gamma^2$. This last simplification is not too hard if you remember that $p + q = 1$. So that's the covariance between one parent and its offspring.

If we want the slope of the line describing the relationship between the parent phenotype and offspring phenotype, we need $Cov(z_p, z_o)/Var(z_p)$. From the stuff we did at the beginning, we know that $Var(z_p) = Var(g_p + \epsilon_p) = Var(g_p) + Var(\epsilon_p) = 2pq\gamma^2 + V_e$, so the slope is

(4)

$$b_{o,p} = \frac{pq\gamma^2}{2pq\gamma^2 + V_e}$$

This is for a single parent and it's offspring. More often we use the relationship between the average parent (or 'mid-parent') and the offspring. Now, we could go back and calculate all the possible pairs of parents and their averages and the possible offspring. But we can also do this in a much shorter way, by making use of handy fact (@fact2). The average parent phenotype is $\bar{z}_p = (z_s + z_d)/2$ where z_s and z_d are the paternal (sire) and maternal (dam) phenotypes. Again, the slope is going to be the covariance divided by

the variance. So let's start by figuring the covariance. Using (@fact2), $Cov(\bar{z}_p, z_o) = Cov((z_s + z_d)/2, z_o) = Cov(z_s/2, z_o) + Cov(z_d/2, z_o) = \frac{1}{2}[Cov(z_s, z_o) + Cov(z_d, z_o)] = Cov(z_p, z_o)$. The second to last step made use of handy fact (@fact1) and the final step comes from the assumption that the genetic covariance between fathers and offspring is the same as the genetic covariance between mothers and offspring. This assumption will be violated if there are sex-linked traits, sex-specific selection, etc. But under these assumptions we've found that the covariance between the average parent phenotype and their offspring is the same as for one parent.

Now let's look at the variance. Since $\bar{z}_p = (z_s + z_d)/2$, the variance is $Var(\bar{z}_p) = Var((z_s + z_d)/2) = Var(z_s)/2 + Var(z_d)/2$ assuming that the parent phenotypes are independent. Parent phenotypes wont be independent if there is assortative mating or if the parents are related or if they were reared in a common environment. Lots of reasons. (Using (@fact2) can you figure out how to account for non-independent parent phenotypes?) But let's stick with the simple case for now and assume they are independent. This means that the slope (i.e. the covariance / variance) is given by

(5)

$$b_{o,\bar{p}} = \frac{pq\gamma^2}{\frac{1}{2}(2pq\gamma^2 + V_e)} = \frac{2pq\gamma^2}{2pq\gamma^2 + V_e}$$

which is the ratio of the genetic variance to the phenotypic variance, otherwise known as the *heritability*. This also means that the slope of the line for a single parent (4) is equal to 1/2 of the heritability.

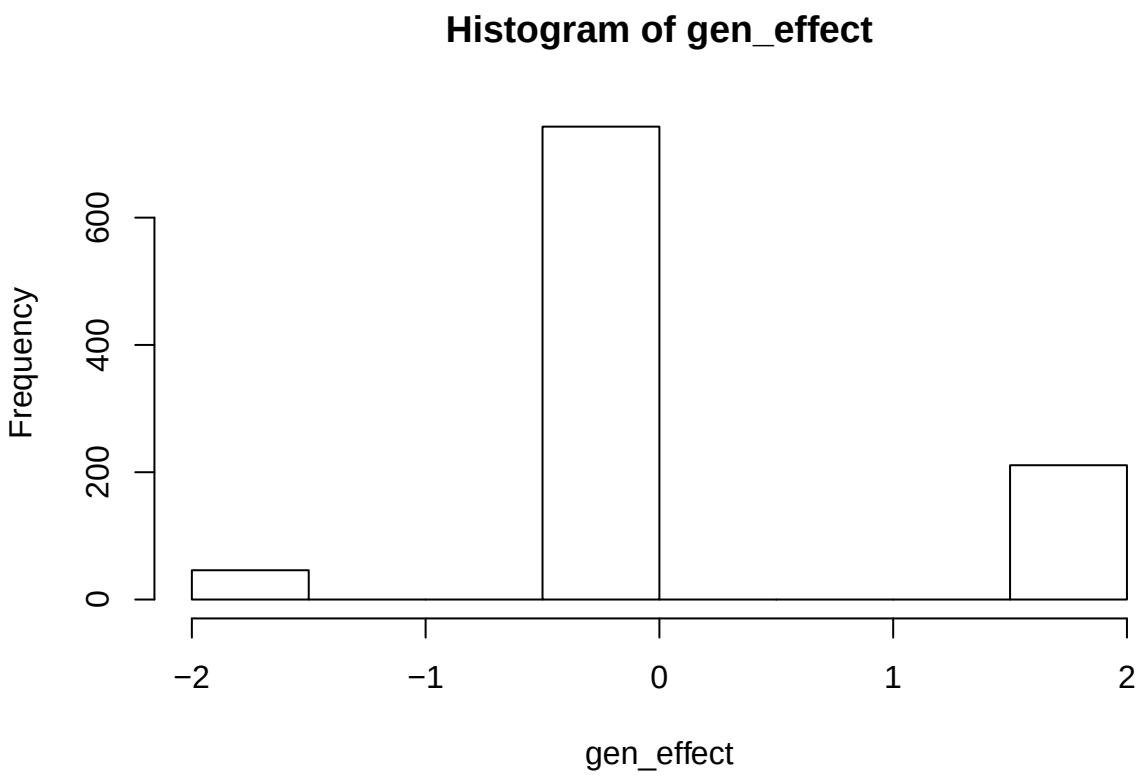
So far, we have assumed a) that genetic effects and environmental effects are independent and b) that 'genetic effects' are passed on to offspring via Mendelian genetics. These assumptions immediately imply that $Var(z) = Var(g) + Var(\epsilon)$ and that the covariance between phenotype and genotype is $Cov(z, g) = Var(g)$. Along the way, we found that the slope of a line between random variables X and Y is given by $b = Cov(X, Y)/Var(X)$. This, in turn, implies that the slope of the phenotype (x)-genotype (y) line will be $b = Cov(z, g)/Var(z) = Var(g)/Var(z)$ which is the 'heritability'. Using Mendelian genetics for a single locus, we deduced that the covariance between parents and offspring is $Cov(z_o, z_p) = \frac{1}{2}Var(g)$ and that the slope of the line relating the average parent phenotype to offspring phenotype is $b_{o,\bar{p}} = Var(g)/[Var(g) + V_e] = Var(g)/Var(z)$ which is also the heritability. Although we've gone through all this for just one locus, it will turn out that the results are analogous for many loci, under the assumption that they all contribute additively.

Many loci

Now, let's think about what happens when we have a bunch of loci that contribute additively. To try to make this concrete, let's simulate the distribution of genetic effects (i.e. the sum of the g_i's) for a large number of individuals. This first code chunk assumes 2 loci with random gene frequencies. You might want to run it a bunch of times to see how much it changes, then bump the number of loci to say 30 and see what happens. You might also want to change the range of gene frequencies to see how this affects the distribution from one time to the next.

```
N=1000 #number of individuals
nloci=2 #number of loci
gam=1 #all loci have same effect.
delta=1;

p_i<-runif(nloci,0,1)

g<-c(gam,delta,-gam)
pg<-function(p) c(p^2,2*p*(1-p),(1-p)^2)

z<-matrix(data=NA,nrow=nloci,ncol=N)
```

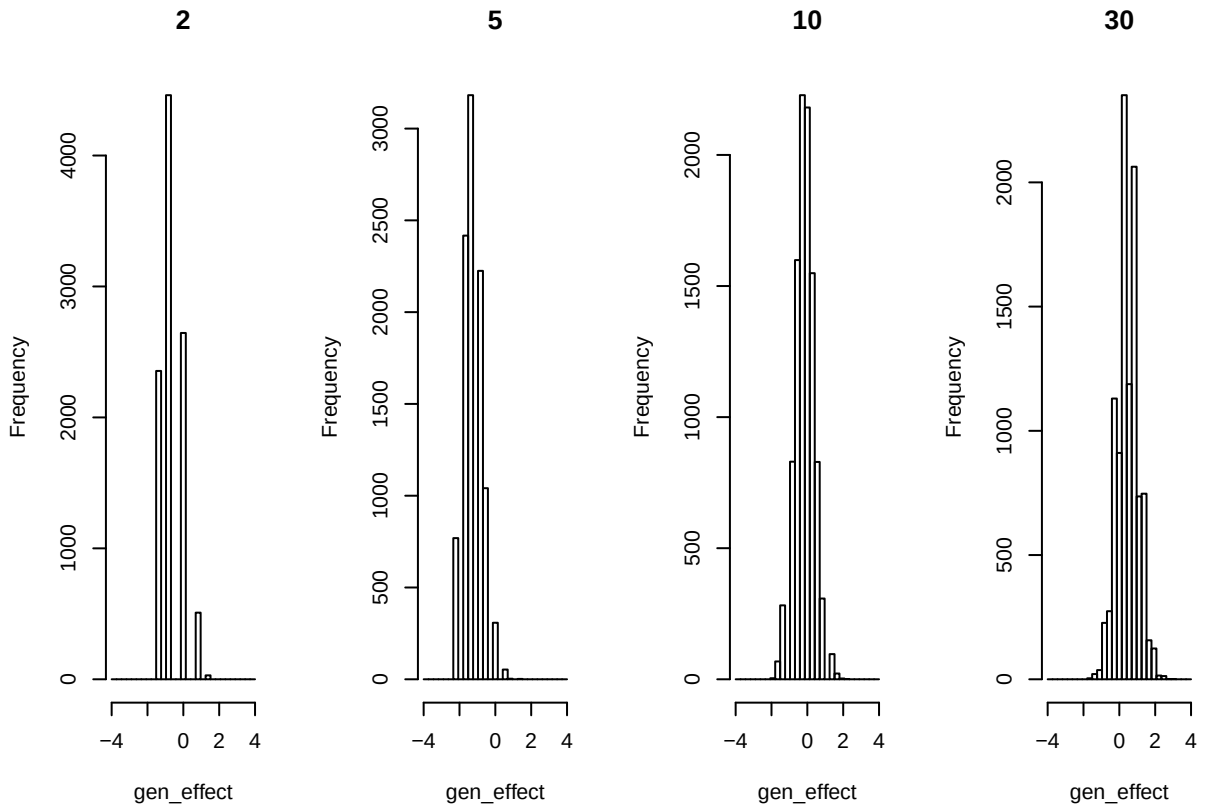
```r
for (i in 1:nloci){z[i,]=sample(g,N,prob=pg(p_i[i]),replace=TRUE)}
gen_effect=apply(z,2,sum)
hist(gen_effect)
```



In this two-locus case, variation in the gene frequencies can skew the distribution of genetic effects strongly in one direction or the other. We should expect this effect to decrease as we add up the effects of more and more loci - Here's a plot illustrating how the distribution of genetic effects changes as we increase the number of loci.

```r
N=10000 #number of individuals
delta=0;
gam=1 #all loci have same effect.



nloci<-c(2,5,10,30) #number of loci

par(mfrow=c(1,4))

for (l in 1:4){
  p_i<-runif(nloci[l],0,1)

  g<-c(gam,delta,-gam)/sqrt(nloci[l]) #rescaling gamma so that spread is same for all cases
  pg<-function(p) c(p^2,2*p*(1-p),(1-p)^2)

  z<-matrix(data=NA,nrow=nloci[l],ncol=N)
```

```
  for (i in 1:nloci[l]){z[i,]=sample(g,N,prob=pg(p_i[i]),replace=TRUE)}
  gen_effect=apply(z,2,sum)
  hist(gen_effect,breaks=seq(-4*gam,4*gam,length=30),xlim=c(-4*gam,4*gam),main=nloci[l])
}
```

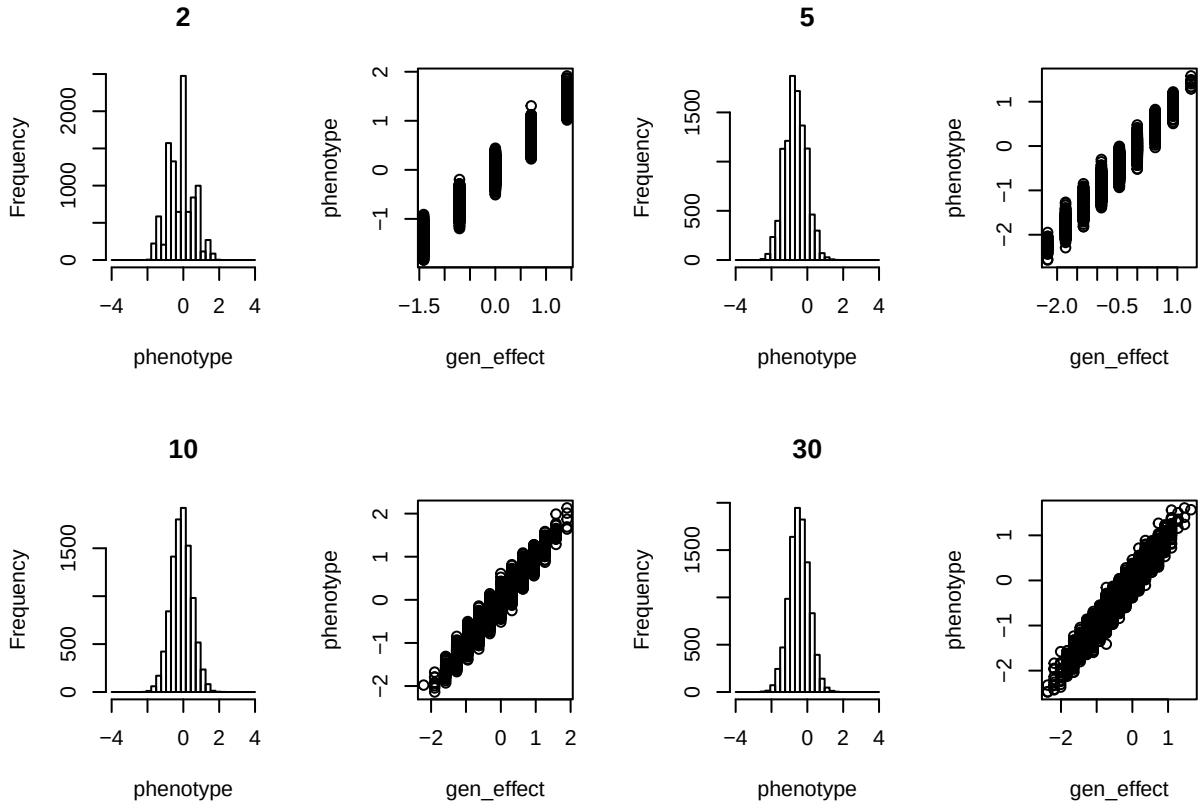


As we increase the number of loci, the distribution for the genetic effect get's closer and closer to Gaussian, provided of course that the gene frequencies are not all close to 0 or 1 (You might want to try this and see what happens). This is a consequence of the central limit theorem.

Next, let's recall that the individual's phenotype is not quite equal to its genotype. The default assumption here is that the phenotype is given by the sum of the genetic effect and some normally distributed noise, i.e. $\epsilon\ N(0, V_e)$. So, we'll re-do the simulation and plot histograms of phenotypes instead of genotypes as well as a plot of genotype v. phenotype.

```
N=10000 #number of individuals
delta=0;
gam=1 #all loci have same effect.

Ve=.02;

nloci<-c(2,5,10,30) #number of loci

par(mfrow=c(2,4))

for (l in 1:4){
  p_i<-runif(nloci[l],0,1)
```

```
  g<-c(gam,delta,-gam)/sqrt(nloci[l]) #rescaling gamma so that spread is same for all cases
  pg<-function(p) c(p^2,2*p*(1-p),(1-p)^2)

  z<-matrix(data=NA,nrow=nloci[l],ncol=N)

  for (i in 1:nloci[l]){z[i,]=sample(g,N,prob=pg(p_i[i]),replace=TRUE)}
  gen_effect=apply(z,2,sum)
  phenotype=gen_effect+rnorm(N,0,sqrt(Ve))
  hist(phenotype,breaks=seq(-4*gam,4*gam,length=30),xlim=c(-4*gam,4*gam),main=nloci[l])
  plot(gen_effect,phenotype)
}
```



There are several things to note from this. First, looking at the histograms, the differences that were so obvious between 2 loci and 30 loci in the earlier figure are now smoothed over by the addition of a relatively small amount of noise. To be specific, if $V_e>Second, the scatterplots show the relationship between genotype and phenotype in this model. For two and 5 loci there are obvious gaps between genotypes; the variation in phenotype within a genotype is due to 'envrionmental' noise.

Let's take a step back and try to think about these observations a bit more carefully. To start, what are the mean and variance in the genotype and phenotype? And did the covariance between parents and offspring change? Let's say that the genetic contribution to an individual's phenotype is $g = \sum_{i=1}^{L} g_i$ where L is the number of loci contributing to the trait. Let m_i be the mean genetic effect for locus i and V_{gi} be the genetic variance at locus i. Since we are assuming additivity and independence, the mean genotype is

(6)

$$E(g) = E\left[\sum_{i=1}^{L} g_i\right] = \sum_{i=1}^{L} E(g_i)$$

If we substitute our result for the mean from earlier, $E(g_i) = \gamma(p_i - q_i)$ we get $E(g) = \sum_i \gamma(p_i - q_i) = 2L\gamma(\bar{p} - \frac{1}{2})$. The intuition here is that the mean phenotype in the population depends not on the particular frequencies of individual loci, but on the average frequency of the A allele. This is a direct consequence of our assumption that the genetic contributions at all loci are the same.

To find the genetic variance in the population, we use the same sort of trick -assuming independence makes this all nice and tidy.

(7)

$$Var(g) = Var\left[\sum_{i=1}^{L} g_i\right] = \sum_{i=1}^{L} Var(g_i)$$

Inserting our earlier results on the genetic variance from a single locus, we get $Var(g) = \sum_{i=1}^{L} 2\gamma^2 p_i q_i = 2\gamma^2 L(\bar{p} - \bar{p^2})$ where $\bar{p^2}$ is the average value of p^2 (or the second moment). We could simplify this a bit more, but without much gain in intuition.

Similarly, the covariance between a parent and its offspring is given by

(8)

$$cov(g_o, g_p) = Cov\left[\sum_{i=1}^{L} g_{oi}, \sum_{j=1}^{L} g_{pj}\right] = \sum_{i=1}^{L}\sum_{j=1}^{L} Cov(g_{oi}, g_{pj}) = \frac{1}{2}\sum_{i=1}^{L} Var(g_{pi})$$

To get to the third expression, we apply handy fact (@fact2) and to get the last expression we make use of the assumption that the loci are independent (so all covariances for $i \neq j$ are 0). As you recall, the genetic covariance is the same whether we use one parent or the average of both.

Finally the slope of the line for the relationship between offspring phenotypes and the average parent phenotype is given by the genetic covariance divided by the phenotypic variance which is

(9)

$$b_{o,\bar{p}} = \frac{cov(z_o, \bar{z}_p)}{Var(\bar{z}_p)} = \frac{\frac{1}{2}\sum_{i=1}^{L} Var(g_{pi})}{\frac{1}{2}Var(z)} = \frac{Var(g)}{Var(z)}$$

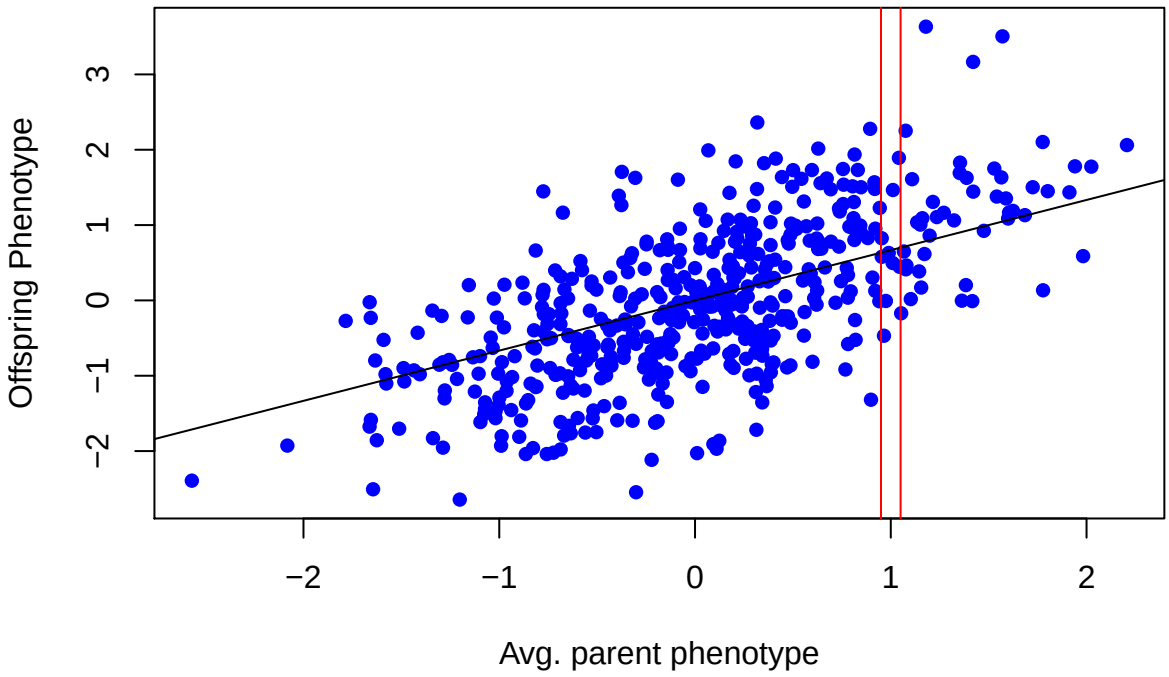
which is - again - the 'heritability.' It is worth point out that the heritability is the ratio of genetic variance to total phenotypic variance. As such, it will change as gene frequencies in the population change.

Before we move on, let's be clear that the reason the many-locus verion is nearly identical to the single locus version is because of the mountain of assumptions we've made getting here. Most importantly, we've assumed that all loci contribute additively and independently. If epistatic interactions (i.e. non-additivity) or linkage (i.e. non-independence) are present the results are considerably more complicated. It is also worth mentioning that what we've been calling genetic variance is typically referred to as the 'additive genetic variance', denoted by V_A because it is the part of the genetic variance that combines additively in determining phenotypes. This is distinct from the total genetic variance V_G which includes all those other, non-additive, effects like dominance and epistasis. So from here on in, we will use V_A to be consistent with the literature.

Selection

Now that we've laid the genetic foundation, let's start thinking about how we might use it to predict evolutionary responses in quantitative traits. To get going on this, let's take another look at the relationship

between parent and offspring phenotypes



The black line is the predicted relationship between parents and offspring using (9). Let's think about this a little more. This line gives us the average offspring phenotype for a given mid-parent phenotype. That is, we have $E(z_o|\bar{z}_p) = a + b_{o,\bar{p}}\bar{z}_p$. If $\bar{z}_p$ was at the population mean phenotype, $E(z) = m$, the mean for offspring would be $E(z_o|m) = a + b_{o,\bar{p}}m$ which, in the absence of selection ought to give us $E(z_o|m) = m$. So if we subtract these two equations, we get $E(z_o|\bar{z}_p) - m = b_{o,\bar{p}}(\bar{z}_p - m)$ which tells us that the difference between the offspring mean phenotype and the population mean is given by the slope times the difference between their parent's mean and that of the population. Recall that the slope is the heritability, i.e. $h^2 = b_{o,\bar{p}}$. So, in the figure above, $m = 0$ and if we were to breed parents that were in between the red lines, $\bar{z}_p \approx 1$, we would predict the average phenotype in their offspring to be $E(z_o|\bar{z}_p = 1) = h^2(1 - 0) = h^2$.

Tangent - It is kind of fun to think about how we might apply this to... us. My mother-in-law, who is very sweet but not a scientist, recently claimed that male children were always taller than their parents. This is clearly impossible, but it prompted me to try and evaluate whether my boys would be taller than I am. Here's a rough version of the analysis: The mean height in the US is roughly 5'7" (averaging across men and women) and the heritability for height is roughly 0.8 (The CDC collects height data and its not too hard to find a publication on human heritability). I am 5'11" and my wife is 5'1", so our 'mid-parent value' is 5'6". So, the best guess for my kids height is 5'7"+0.8(5'6"-5'7")=5'6.2" which is not only less than my height, but less than the population mean! (I actually re-did this calculation assuming sex-specific values for mean height and heritability which comes out a little different but is too long for a tangent - but the upshot is that my kids should still end up shorter!) ___

Now, since this equation for the mean phenotype of offspring is linear, we can extend it easily to the case where we allow more than one parent phenotype to reproduce. We do so by recalling the handy fact that $E(X) = E_Y[E(X|Y)]$ (aka the 'Law of Iterated Expectations') which in this case, means that we want $E(z_o) = E_{\bar{z}_p}[E(z_o|\bar{z}_p)] = m + h^2 E_{\bar{z}_p}[\bar{z}_p - m] = m + h^2[E(\bar{z}_p) - m]$. This gets us to the *breeder's equation* which is usually written as

(10)
$$\Delta R = h^2 \Delta S$$

where h^2 is the heritability, $\Delta S = E(\bar{z}_p) - m$ is known as the *selection differential*, and $\Delta R = E(z_o) - m$ is the *response to selection*. Another way of interpretting the heritability is that it is the fraction of the change in the mean phenotype in parents that is transmitted to the next generation.

Fitness

Let's pause for a bit, momentarily foresaking all our hard-earned knowledge about genetics, and think about selection more generally. We have previously defined 'fitness' in a variety of ways, using various proxies. To begin, let's think about fitness in terms of the survival of individuals in the parent generation prior to reproduction. If we have N total individuals and they have phenotypes $z_1, ...z_N$ the mean phenotype before selection is given by $\bar{z} = \frac{1}{N}\sum_i z_i$. Let's say that fitness (survival) is some function of the phenotype, i.e. $W(z)$ that takes values 0 or 1. The number of surviving individuals is then $N_s = \sum_i W(z_i)$ and the mean phenotype among survivors is $\bar{z}_s = \frac{1}{N_s}\sum_i z_i W(z_i)$. We can play a little trick here divide bot the numerator and denominator by N to get

$$\bar{z}_s = \frac{\frac{\sum_i z_i W(z_i)}{N}}{\frac{\sum_i W(z_i)}{N}} = \frac{\overline{zW}}{\overline{W}}$$

Now the denominator is the mean fitness (survival) and the numerator is the average of the product of the trait and fitness. We can re-arrange handy fact(@fact3) to see that $E(XY) = Cov(X,Y) + E(X)E(Y)$. Applying this identity to the mean of survivors gets us to

$$\bar{z}_s = \frac{\overline{zW}}{\overline{W}} = \frac{Cov(z,W) + \bar{z}\overline{W}}{\overline{W}} = \bar{z} + \frac{Cov(z,W)}{\overline{W}}$$

So the change in the mean due to a round of selection, i.e. the selection differential, is given by the covariance between the trait and fitness divided by the mean fitness, $\Delta S = \Delta\bar{z} = Cov(z,W)/\overline{W}$, which is equivalent to the covariance between the trait and *relative* fitness, $\Delta S = Cov(z,\omega)$, where relative fitness is defined as $\omega(z) = W(z)/\overline{W}$.

Let's do this again, but this time define fitness as the number of offspring produced. That is, parents with phenotype z_i, make $W(z_i)$ offspring. The total number of offspring produced is $N_b = \sum_i W(z_i)$. In light of our earlier work, we probably want to allow for the possibility that offspring are not identical to parents. Instead, we can think about the offspring phenotype as $z_i + \Delta z_i$. The mean phenotype among offspring is $\bar{z}_o = \frac{1}{N_b}\sum_i (z_i + \Delta z_i)W(z_i)$. The mean following a round of selection is then given by

(11)
$$\bar{z}_o = \frac{\overline{(z+\Delta z)W}}{\overline{W}} = \bar{z} + \frac{Cov(z,W)}{\overline{W}} + \frac{Cov(\Delta z, W)}{\overline{W}}$$

So what is Δz? It's the difference between the parent and offspring phenotypes. We can use our genetic model to approximate this as follows: $\bar{z}_o - m = h^2(\bar{z}_p - m)$ so $\Delta z = z_o - z_p \approx m + h^2(\bar{z}_p - m) - \bar{z}_p = (h^2 - 1)(z_p - m)$. Plugging this in, we get $Cov(\Delta z, W) \approx Cov((h^2-1)(z_p - m), W) = (h^2 - 1)Cov(z,W)$. We can use this in (11) to evaluate the mean phenotype for offspring following a round of selection

$$\bar{z}_o = \bar{z} + \frac{Cov(z,W)}{\overline{W}} + (h^2 - 1)\frac{Cov(z,W)}{\overline{W}} = \bar{z} + h^2\frac{Cov(z,W)}{\overline{W}}$$

Thinking about this in the context of the breeder's equation (10) we can make the analogy between ΔS and $Cov(z,W)/\overline{W}$. That is, the selection differential is again given by the covariance between fitness and phenotype.

It is worth noting that (11) is also known as **the Price Equation**. The Price equation is just an identity converting means of products into covariances. As such, it has some surprisingly ardent proponents and nay-sayers. For our purposes, it is a convenient trick for thinking about selection.

Now we have in hand an approach for measuring selection in natural populations. And since $\bar{z}' - \bar{z} = \Delta S$ is the selection differential, we can plug this in to the breeders equation (10) to find out the change in the mean in the next generation. Doing so we get the following sequence of equivalent expressions

(12)

$$\Delta R = \bar{z}_o - \bar{z} = h^2 \Delta S = h^2(\bar{z}' - \bar{z}) = h^2 \frac{Cov(z,W)}{\bar{W}} = \frac{V_A(z)}{V_P(z)} \frac{Cov(z,W)}{\bar{W}} = \frac{V_A(z)}{\bar{W}} \frac{Cov(z,W)}{V_P(z)} = \frac{V_A(z)}{\bar{W}} b_{W|z}$$

The last term , $b_{W|z} = Cov(z,W)/V_P(z)$ is known as the *selection gradient* (we will see why in a few paragraphs). It is the slope of the regression of fitness on phenotype. So the change in the mean of a trait after one generation is proportional to the slope of the linear regression of fitness on phenotypes. Since we have not assumed that fitness is linear in the phenotype, this might seem kind of odd. But remember that all we have said so far is that the change in the *mean* phenotype is driven by the average slope over the current distribution of phenotypes. Nonlinearities in the fitness function will induce changes in other moments of the phenotype distribution.

To see this, note that so far we havent said much about what constitutes a trait when calculating fitness. This is a strength of the Price equation; all of what we did with fitness and covariances was pretty generic, so it ought to apply regardless of how we define a trait. For example, what if instead of the mean we wanted to know something about changes in some function of z? To put the Price equation to work we just need to define a new trait $z' = f(z)$ and plug in. Let's say that we want to know about changes in variance due to selection on parents? We can find the selection differential for $z' = z^2$ by plugging this into (11) to find

$$\bar{z}_s^2 = \bar{z}^2 + \frac{Cov(z^2, W)}{\overline{W}}$$

And since $Var(z) = E(z^2) - E(z)^2$ we can find the change in variance from $\Delta Var(z) = \Delta E(z^2) - \Delta E(z)^2$ which is

$$\Delta Var(z) = \frac{Cov(z^2, W)}{\overline{W}} - \left[\frac{Cov(z,W)}{\overline{W}}\right]^2$$

You might be wondering why we didnt just plug in $(z - \bar{z})^2$ (which would be very clever of you) to find the change in the variance directly. The reason is that doing so doesnt allow for the change in $\bar{z}$ due to selection so the result tends to be biased high. Of course, if we already have an estimate of $\bar{z}'$, we *could* plug $(z - \bar{z}')^2$ into (11) to get the right answer.

An interesting thing happens if the "trait"" is *fitness*. That is, let's say we want to know how mean fitness changes across a generation, i.e. $\bar{W}_o = \bar{W}_p + h^2 \frac{Cov(W,W)}{\overline{W}}$. To simplify this, recall two things. First, note that $Cov(W,W)$ is the phenotypic variance in fitness, $V_p(W)$ from the definition of covariance. Second, the heritability is given by the additive genetic variation in the trait divided by the phenotypic variation. So, in this case, we could write this as $h^2 = V_A(W)/V_P(W)$ where the subscripts A and P differentiate the genetic and phenotypic variances. Putting these together, we get

$$\Delta \bar{W} = h^2 \frac{Cov(W,W)}{\bar{W}} = \frac{V_A(W)}{V_P(W)} \frac{V_P(W)}{\bar{W}} = \frac{V_A(W)}{\bar{W}}$$

Let's re-write this in terms of relative fitness which is $\omega = W/\bar{W}$. Recall that $Var(aX) = a^2 Var(X)$ so that $Var(\omega) = Var(W)/\bar{W}^2$. Substituting this into the equation above, we get

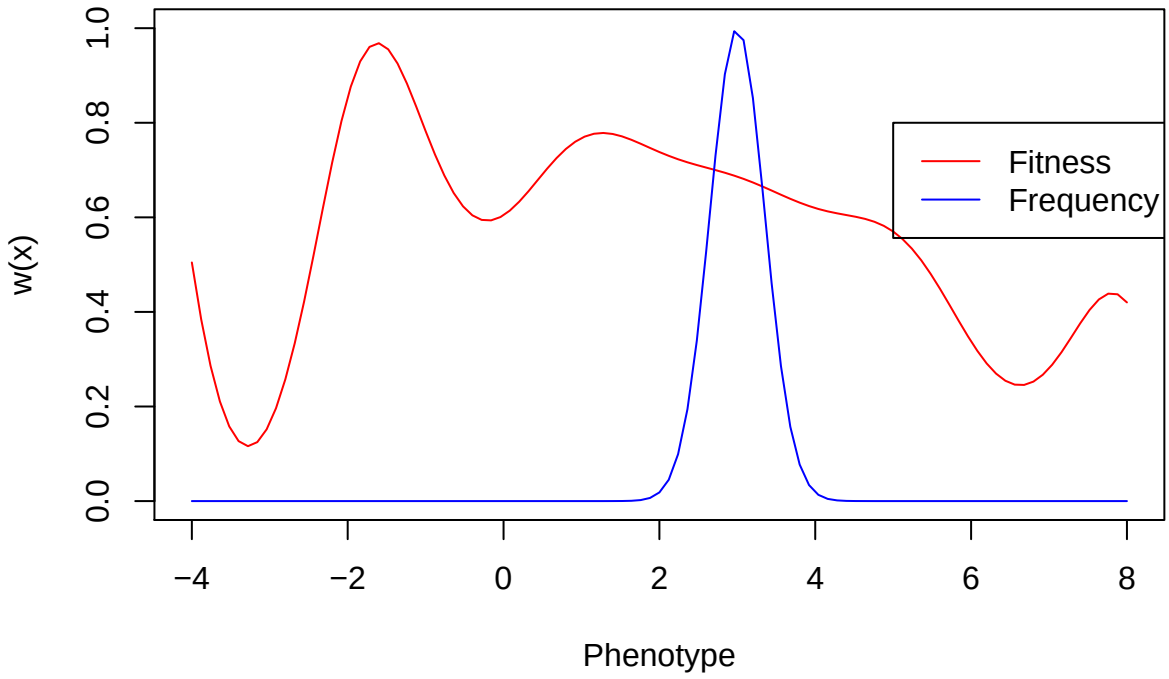
(13)

$$\Delta \bar{W}/\bar{W} = V_A(\omega)$$

This is **Fisher's fundamental theorem of natural selection** which says that the proportional change in mean fitness across a generation is equal to the additive genetic variance in relative fitness. Since variances are always greater than or equal to zero, (13) says that the change in mean fitness is always positive.

Applications

To use this stuff to make some more specific predictions, we need to define a fitness function. The most commonly used fitness functions in evolutionary quantitative genetics are linear, quadratic, and Gaussian. Now, real fitness functions are unlikely to be precisely linear, quadratic or Gaussian, but we can think of these as approximations to the real fitness function near the mean phenotype or the optimal phenotype. To see why this might be relevant, imagine that we have an arbitrary fitness function whose curvature is negligible over the observed range of phenotypes:



In this figure, the fitness function is pretty wiggly, but within the range of observed phenotypes it is actually nearly linear. Recall that covariance measures the linear relationship between two variables so we ought to get a pretty good approximation of $Cov(z, W)$ in this case if we used a low-order approximation to W in the vicinity of the mean. To see this let's approximate to second order around the mean phenotype, $\bar{z}$. To write our approximation we'll use the first two derivatives of W evaluated at $\bar{z}$. That is

(14)
$$W(z) \approx W(\bar{z}) + W'(\bar{z})(z - \bar{z}) + \frac{1}{2}W''(\bar{z})(z - \bar{z})^2$$

One thing that pops right out of this approximation is that mean fitness, $\bar{W}$, is approximately $\bar{W} \approx W(\bar{z}) + \frac{1}{2}W''(\bar{z})V_P(z)$ where $V_P(z)$ is the *phenotypic* variance. Moreover, since W is nearly linear in the vicinity of the population's current phenotype distribution, the second (and higher) derivatives must be pretty small. So we can ignore them which means

(15)
$$W(z) \approx W(\bar{z}) + W'(\bar{z})(z - \bar{z})$$

and this implies that $\bar{W} \approx W(\bar{z})$. To figure out the selection differential, we calculate

$$Cov(z, W) \approx Cov(z, W(\bar{z}) + W'(\bar{z})(z - \bar{z})) = W'(\bar{z})V_P(z)$$

And plugging this into (10) we find that the change in the mean when fitness is nearly linear over the current range of phenotypes is given by

(16)
$$\Delta\bar{z} = h^2\frac{Cov(z, W)}{\bar{W}} \approx h^2\frac{W'(\bar{z})V_P(z)}{\bar{W}} = \frac{V_A(z)}{V_P(z)}\frac{W'(\bar{z})V_P(z)}{\bar{W}} = V_A(z)\frac{W'(\bar{z})}{\bar{W}}$$

Comparing this last bit to (12) we can see where the name 'selection gradient' comes from since $b_{W|z} \approx W'(\bar{z})$.

To see why a Gaussian fitness function is often assumed, consider approximating $\ln[W]$ in the vicinty of some 'optimal' phenotype z^*. To simplify the notation let $r = ln[W]$ be log-fitness, and approximating this to second order we would get

(17)
$$r(z) = ln[W(z)] \approx r(z^*) + r'(z^*)(z - z^*) + r''(z^*)(z - z^*)^2$$

Since z^* is a local maximum in fitness, the first derivative term is 0 so that the approximation on the original scale is

(18)
$$W(z) \approx W(z^*)e^{\frac{1}{2}r''(z^*)(z-z^*)^2}$$

which looks alot like a normal or Gaussian distribution.

Selection on two traits

Although we've focused this far on the evolution of just one trait, it is far more likely that selection acts on many traits at the same time. Multivariate quantitative genetics attempts to take this into account. The math is a bit more involved than we have time for, but I'd like to highlight a few of the more counter-intuitive things that happen when we move beyond one trait.

To begin, let's think about what happens when there are two traits, call them z_1 and z_2. Again, we'll say that the phenotypes can be decomposed into genetic and environmental components, specifically, $z_1 = g_1 + \epsilon_1$ and $z_2 = g_2 + \epsilon_2$. In keeping with our earlier model, we'll assume that the genetic and environmental components are independent. If we went a step further and assumed that all the g's and ϵ's are independent, we dont need multivariate quantitative genetics at all - we can just use the single trait stuff we already developed, one trait at a time. But since we are talking about two traits in the same individual, it probably doesnt make much sense to assume that things are totally independent (and it's way more interesting if they arent). So let's evaluate the covariance between traits 1 and 2-

$$Cov(z_1, z_2) = Cov(g_1, g_2) + Cov(\epsilon_1, \epsilon_2)$$

The usual notation is to let $C_A = Cov(g_1, g_2)$ be the 'additive genetic covariance' and $C_E = Cov(\epsilon_1, \epsilon_2)$ be the environmental covariance. The two phenotypes can covary if either the underlying genotypes covary OR if the environmental effects covary. In the first case, this can happen because the traits share common genes, arise from a single locus with pleiotropic effects, or are simply linked on a chromosome. The environmental effects can covary because individuals experience different environments that affect both traits. Common examples include food availability and temperature which tend to affect lots of traits both directly and indirectly regardless of whether they share a genetic basis. A first hint of the wierdness that can happen in higher dimensions that can't happen in just one is that it is now possible to have two traits that appear

independent (or nearly so) because the genetic and environmental covariances cancel. That is, if we have a situation where $Cov(g_1, g_2) > 0$ and $Cov(\epsilon_1, \epsilon_2) < 0$ the phenotypic covariance will acutally be less than the genetic covariance...

Next, let's think about selection. To keep things simple, we'll look at the linear approximation to the fitness function (15) and extend this to two dimensions. To do so, we'll need the partial derivatives of fitness with respect to z_1 and z_2.

Mathematical Aside - Partial Derivatives

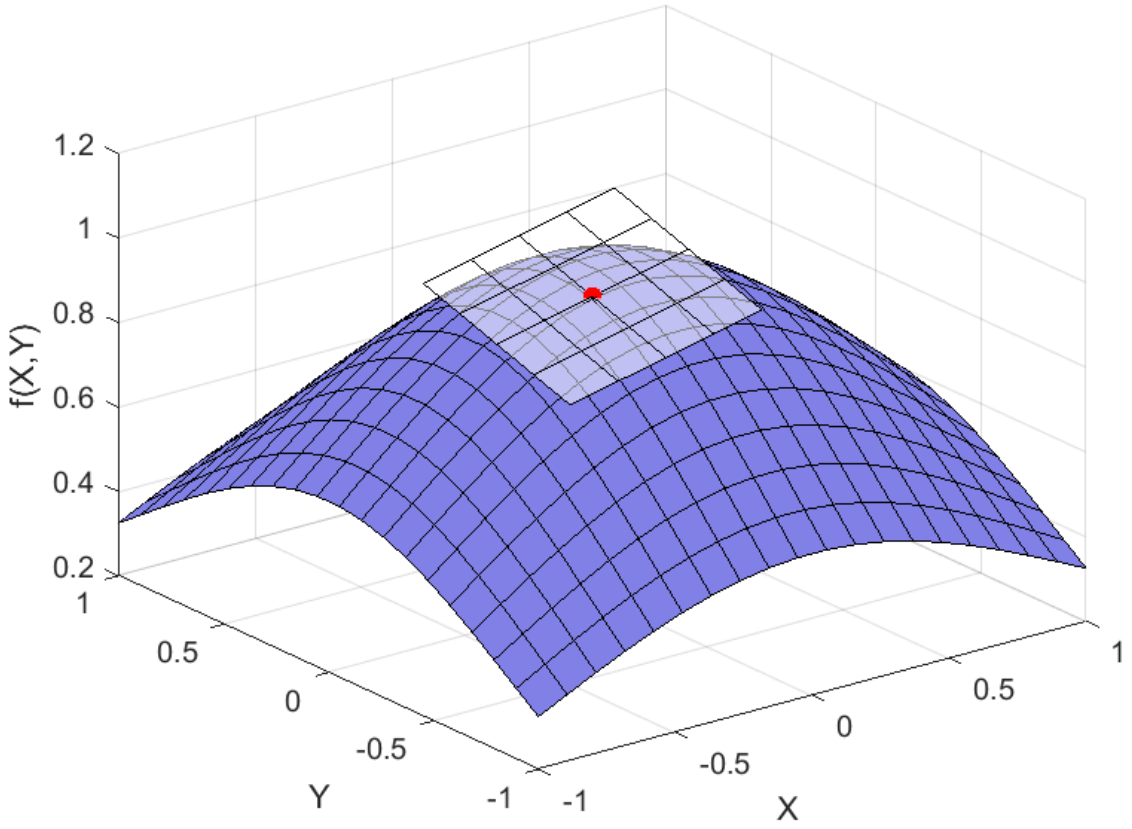
Recall that in 1-dimension, the derivative gave us the slope of the function evaluated at a point. Similarly, the partial derivative of a function of more than one variable is the slope in a given direction. Usually, we take partial derivatives in directions corresponding to coordinate axes. Let's think about that - if we have a function, $f(x, y)$ and we want to know the slope in the x-direction, we would think about the slope the slope of the line connecting two points that have different values of x and the same value of y, i.e. slope= $[f(x_2, y) - f(x_1, y)]/[x_2 - x_1]$. Setting $x_1 = x$ and $x_2 = x + \Delta x$, the slope is $[f(x + \Delta x, y) - f(x, y)]/\Delta x$ and taking the limit as $\Delta x \to 0$, we get the partial derivative of f with respect to x, which is written as

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$$

Similarly the partial derivative with respect to y is

$$\frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}$$

Operationally, we take the partial derivative with respect to one variable by treating the over variables as constants. This is illustrated in the figure below



___ ##Returning to selection##

We'll look at the linear approximation to the fitness function (15) and extend this to two dimensionsusing the partial derivatives of fitness with respect to z_1 and z_2 which is given by

(19)

$$W(z_1, z_2) \approx W(\bar{z}_1, \bar{z}_2) + \frac{\partial W}{\partial z_1}(z_1 - \bar{z}_1) + \frac{\partial W}{\partial z_2}(z_2 - \bar{z}_2)$$

where the partial derivatives are evaluated at the mean values for each trait. We then use this to determine the covariance between z_1 and fitness which gives us the selection differential on z_1:

$$Cov(z_1, W) \approx Cov(z_1, W(\bar{z}_1, \bar{z}_2) + \frac{\partial W}{\partial z_1}(z_1 - \bar{z}_1) + \frac{\partial W}{\partial z_2}(z_2 - \bar{z}_2)) = \frac{\partial W}{\partial z_1} Var(z_1) + \frac{\partial W}{\partial z_2} Cov(z_1, z_2)$$

From here we can see that the selection differential on trait 1 depends not only on the fitness gradient with respect to trait 1, but also the indirect effect of selection on trait 2. (Obviously, there is an analogous equation for selection on trait 2 which looks just like the above with all the 1's and 2's swapped). So even if fitness doesnt depend on z_1, we can have selection on trait 1 driven by $Cov(z_1, z_2)$. Even more counter-intuitively, we can have selection on trait 2 drag trait 1 in a direction in which fitness actually decreases. The condition for this is that the sign of $\frac{\partial W}{\partial z_1} Var(z_1) + \frac{\partial W}{\partial z_2} Cov(z_1, z_2)$ is opposite the sign of $\frac{\partial W}{\partial z_1}$. Assuming for example that this is positive, then the direction of selection is reversed when $\frac{\partial W}{\partial z_1} Var(z_1) < -\frac{\partial W}{\partial z_2} Cov(z_1, z_2)$ which requires that $frac \partial W \partial z_2$ and $Cov(z_1, z_2)$ have opposite signs (at a miminum).

The results for evolutionary changes following a round of selection are analogous:

(20)

$$\bar{z}_{1o} - \bar{z}_1 \approx \frac{\partial W}{\partial z_1} V_A(z_1) + \frac{\partial W}{\partial z_2} C_A(z_1, z_2)$$

Again, evolution of trait 1 is now influenced by selection on trait 2 but the effect is mediated through the genetic covariance rather than the phenotypic covariance.

Many Traits

I feel sort of compelled to go on with the general case, which requires the use of linear algebra. The generalization from the 2-d case is fairly straightforward. To begin, we need to define some notation. Let $\mathbf{z}$ be a vector of n different traits, i.e.

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$$

and let the population mean prior to selection be $\bar{\mathbf{z}}$. Define the genetic and environmental components of $mathbfz$ analogously, that is

$$\mathbf{z} = \mathbf{g} + \epsilon \rightarrow \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

In light of this, we will arrange the variances and covariances into matrices. The phenotypic covariance matrix, $\mathbf{P}$, is given by

$$\mathbf{P} = \begin{bmatrix} V(z_1) & Cov(z_1, z_2) & \cdots & Cov(z_1, z_n) \\ Cov(z_2, z_1) & V(z_2) & \cdots & Cov(z_2, z_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(z_n, z_1) & Cov(z_n, z_2) & \cdots & V(z_n) \end{bmatrix}$$

Similarly, the genetic covariance matrix, $\mathbf{G}$ is given by

$$\mathbf{G} = \begin{bmatrix} V_A(z_1) & Cov_A(z_1, z_2) & \cdots & Cov_A(z_1, z_n) \\ Cov_A(z_2, z_1) & V_A(z_2) & \cdots & Cov_A(z_2, z_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov_A(z_n, z_1) & Cov_A(z_n, z_2) & \cdots & V_A(z_n) \end{bmatrix}$$

The corresponding 'environmental' covariance matrix, $\mathbf{E}$, is defined similarly.

As in the single-trait version, the phenotypic variance-covariance matrix is decomposed as $\mathbf{P} = \mathbf{G} + \mathbf{E}$. This is easy to see if you recognize that the $(i,j)^{th}$ element of $\mathbf{P}$ **is given by the sum of the** $(i,j)^{th}$ **elements of G and E. That is, when** $i \neq j$**, we have** $Cov(z_i, z_j) = Cov(g_i, g_j) + Cov(\epsilon_i, \epsilon_j)$ **just it like did in the two trait case, and when** $i = j$ $Var(z_i) = V_A + V_E$**.**

Making the analogy with (10), we can think about a multivariate version of the breeder's equation. To do so, we need a multivariate version of heritability and a vector of selection differentials, $\Delta\mathbf{S}$**. Recalling that in 1-d heritability is** $h^2 = V_A/V_P$**, the multivariate analog is** $\mathbf{H} = \mathbf{GP}^{-1}$ **where** $\mathbf{P}^{-1}$ **is the 'inverse' of the phenotypic covariance matrix. The inverse satisfies**

the identity $\mathbf{P^{-1}P = I}$ where I is the 'identity' matrix. I is a matrix with ones on the diagonal and zeros everywhere else which plays the role of 1 in the 1-d case (i.e. $1x = x \rightarrow \mathbf{IX = X}$ and $(1/x)x = 1 \rightarrow \mathbf{X^{-1}X = I}$).

The vector of selection differentials is given by $\mathbf{\Delta S = \bar{z}' - \bar{z}}$

$$\mathbf{\Delta S} = \begin{bmatrix} \Delta S_1 \\ \Delta S_2 \\ \vdots \\ \Delta S_n \end{bmatrix}$$

Putting these together we have the multivariate breeders equation

(21)

$$\mathbf{\bar{z}_o - \bar{z} = \Delta R = H\Delta S = GP^{-1}\Delta S}$$

It is worth pointing out that almost no one uses H for multivariate heritability; the preference seems to be to keep $\mathbf{GP^{-1}}$ distinct.

To measure selection, we still use the covariance between the traits and fitness, though there is now a vector of covariances. To keep things intuitive, we will stick with the linear approximation to fitness analogous to (19) but with n traits. We can write this approximation formally as

$$W(\mathbf{z}) \approx W(\bar{\mathbf{z}}) + \nabla \mathbf{W} \cdot (\mathbf{z} - \bar{\mathbf{z}})$$

where the notation $\nabla W \cdot (\mathbf{z} - \bar{\mathbf{z}})$ is the multivariate calculus / linear algebra version of the sum in (19), that is

$$\nabla \mathbf{W} \cdot (\mathbf{z} - \bar{\mathbf{z}}) = \sum_{i=1}^{n} \frac{\partial W}{\partial z_i}(z_i - \bar{z}_i)$$

Evaluating the covariance between the trait vector z and W using this approximation, we get

(22)

$$\mathbf{\Delta S = P \nabla W}$$

which is a super-compact way of saying

$$\Delta S_i = \sum_{i=1}^{n} \frac{\partial W}{\partial z_i} Cov(z_i, z_j)$$

Finally, we can plug this into the multivariate breeders equation (21) to get

$$\mathbf{\bar{z}_o - \bar{z} = GP^{-1}\Delta S = GP^{-1}P\nabla W = G\nabla W}$$

This is actually how we derived (20) - without the linear algebra, it takes alot of tedious algebra to cancel all the phenotypic covariances! Also, we can see that the chain of reasoning is precisely what it was in the 1-d case. In fact, we can see from the second and fourth terms that $\mathbf{P^{-1}\Delta S = \nabla W}$. This is actually the origin of the term 'selection gradient' for the multiple regression of fitness on traits, since $\mathbf{b_{W|Z} = P^{-1}\Delta S}$

There is, of course, alot more we could do, with this stuff, but this is all I can write for now.