Agro/ANSCI 931 Population Genetics

Gerardo Mamani

Lincoln, 11/1/2017

1. A bi-allelic locus affects oil content (%) in rye with the following genotypic values:

A1A1: 9.0% A1A2: 8.9% A2A2: 8.0%.

The frequency of A1 is 0.30 and the population is in Hardy-Weinberg Equilibrium.

a. Given these genotypic values, what are the values of a and d?

mid-point $= (A_1A_1 + A_2A_2)/2 = 8.5\%$

a $= A_1A_1 - mid - point = 9.0\% - 8.5\% = 0.5$

d $= A_1A_2 - mid - point = 8.9\% - 8.5\% = 0.4$

b. What is the population mean?

$M = a(p - q) + 2dpq$

```
p<-0.3 # Frequency of A1
q<-0.7 # Frequency of A2
a<-0.5 # value of a
d<-0.4 # value of d
```

```
mean<-a*(p-q)+(2*d*p*q)
mean
```

```
[1] -0.032
```

c. What is the mode of inheritance?

Partial dominance

2. Consider two more loci that affect oil content, B and C. Assume B1 and C1 have exactly the same average effect as A1. Likewise, B2 and C2 have exactly the same average effect as A2.

a. What is the breeding value of an individual that is $A_1A_2B_1B_2C_1C_2$?

```
p<-0.3 # Frequency of A1
q<-0.7 # Frequency of A2
a<-0.5 # value of a
d<-0.4 # value of d
```

Average value of allele $A_1, \alpha_1 = q[a + d(q - p)]$

```
ave_A1<-q*(a+d*(q-p))
ave_A1
```

```
[1] 0.462
```

Average value of allele A_2, $\alpha_2 = -p[a + d(q - p)]$:

```
ave_A2<--p*(a+d*(q-p))
ave_A2
```

```
[1] -0.198
```

Breeding Value of genotype $A_1 A_2 = \alpha_1 + \alpha_2$:

```
BV_A1A2<-ave_A1+ave_A2
BV_A1A2
```

```
[1] 0.264
```

Based in the assuntion, then the breeding value of $A_1 A_2 B_1 B_2 C_1 C_2$ is:

```
BV_A1A2B1B2C1C2=3*ave_A1 + 3*ave_A2
BV_A1A2B1B2C1C2
```

```
[1] 0.792
```

b. What is the breeding value of an individual that is $A_1 A_1 B_1 B_1 C_1 C_1$?

Based in the assuntion $A_1 = B_1 = C_1$, then the breeding value of $A_1 A_1 B_1 B_1 C_1 C_1$ is:

```
BV_A1A1B1B1C1C1=6*ave_A1
BV_A1A1B1B1C1C1
```

```
[1] 2.772
```

c. What does the calculated value in part b mean in terms of offspring performance of the individual in question?

For part b is

$$\frac{2.772}{2} = 1.386$$

. The offspring should have 1.386 percentage points more oil than the average of the population.

d. What is the expected breeding value of an offspring derived from a mating between the two individuals in parts a and b?

```
BV_offspring<-0.5*(BV_A1A2B1B2C1C2 + BV_A1A1B1B1C1C1)
BV_offspring
```

```
[1] 1.782
```

Offspring between **a** and **b** will be

$$\frac{0.792}{2} + \frac{2.772}{2} = 1.782$$

3. Why is the concept of the average effect of an allele important? Why not just deal with the average effect of a genotype, rather than the average effect of an allele?

Because parent pass on their alleles and not their genotypes to the next generation, genotypes being created afresh in each generation.

4. Consider a large population of beef cattle that has been genotyped for Myostatin and phenotyped for two traits, fat and loin muscle area, with the following results.

Table

Genotype	Frequency	Fat (cm)	Loin Muscle Area (cm^2)
AA	0.5	0.51	116.1
Aa	0.4	0.76	103.0
aa	0.1	1.5	83.9

a. Compute the expected value of Fat.

Genotype	Frequency	Fat (cm)	Loin Muscle Area (cm^2)	Freq × Value Fat	Freq × Value Loin
AA	0.5	0.51	116.1	0.255	58.05
Aa	0.4	0.76	103.0	0.304	41.2
aa	0.1	1.5	83.9	0.15	8.39
			Sum	0.709	107.64

$E(Fat) = \sum Frequency \times Genotypic value$

$E(Fat) = 0.5 \times 0.51 + 0.4 \times 0.76 + 0.1 \times 1.5 = 0.709$

b. Compute the variance of trait Fat

$Var(Fat) = E(Fat^2) - E(Fat)^2$

or it can be estimated directly as below

$Var(Fat) = 0.5(0.51 - 0.709)^2 + 0.4(0.76 - 0.709)^2 + 0.1(1.5 - 0.709)^2$

$Var(Fat) = 0.083409 cm^2$

c. Compute the expected value of Fat given that the individual carries at least one 'a' allele?

$Pr(Aa, a)/Pr(a) = \frac{0.4}{0.5} = 0.8$

$Pr(aa, a)/Pr(a) = \frac{0.1}{0.5} = 0.2$

$E(Fat/a) = 0.8 \times 0.76 + 0.2 \times 1.5 = 0.908 cm$

d. Compute the correlation between Fat and Loin Muscle Area?

```
fat<-c(0.51, 0.76, 1.5)
mus<-c(116.1, 103, 83.9)
cov<-(sum((fat-mean(fat))*(mus-mean(mus))))/(length(fat)-1)
var1<-(sum((fat-mean(fat))^2))/(length(fat)-1)
var2<-(sum((mus-mean(mus))^2))/(length(mus)-1)
r<-cov/sqrt(var1*var2)
r
```

```
[1] -0.9853857
```

$$r_{fat.mus} = \frac{Cov(Fat,Mus)}{\sqrt{Var(Fat)Var(Mus)}} = \frac{-8.2145}{\sqrt{0.2650333 \times 262.21}} = -0.9853857$$

e. Compute the coefficient of regression of Fat on Loin Muscle Area?

```r
cov<-(sum((fat-mean(fat))*(mus-mean(mus))))/(length(fat)-1)
var2<-(sum((mus-mean(mus))^2))/(length(mus)-1)
b<-cov/var2
b
```

```
[1] -0.03132794
```

$$b_{fat.mus} = \frac{Cov(Fat,Mus)}{\sqrt{Var(Mus)}} = \frac{-8.2145}{\sqrt{262.21}} = -0.03132794$$

f. Compute the predicted value for Loin Muscle Area for an individual from this population that has Fat=1.0?

```r
data<-data.frame(fat,mus)
mean(data$fat)
```

```
[1] 0.9233333
```

```r
mean(data$mus)
```

```
[1] 101
```

```r
lm(mus~fat)
```

```
Call:
lm(formula = mus ~ fat)

Coefficients:
(Intercept)          fat
     129.62       -30.99
```

$$E(Mus|Fat = 1) = a + b.fat$$

$$E(Mus|Fat = 1) = 129.6 - 30.99 \times 1 = 98.63 cm^2$$

g. Compute the variance of a new variable Z, which is equal to $Z = 3Fat - 0.5$LoinMuscleArea

$$Z = 3 * Fat - 0.5 * Mus$$

```r
Z<-3*fat - 0.5*mus
Z
```

```
[1] -56.52 -49.22 -37.45
```

Values for Z by genotype are :

$$AA = -56.52$$

$$Aa = -49.22$$

$$aa = -37.45$$

```
var_Z<-0.5*(-56.52- mean(Z))^2+0.4*(-49.22- mean(Z))^2+0.1*(-37.45- mean(Z))^2
var_Z<-0.5*(-56.52- (mean(Z)))^2+0.4*(-49.22- (mean(Z)))^2+0.1*(-37.45- (mean(Z)))^2
var_Z
```

```
[1] 50.08793
```

$Var(Z) = 0.5(0.51 - (-47.73))^2 + 0.4(0.76 - (-47.73)^2 + 0.1(1.5 - (-47.73)^2$

$Var(Z) = 50.08793$

h. Compute Cov(Z, Fat)

```
cov_Zfat<-(sum((Z-mean(Z))*(fat-mean(fat))))/(length(Z)-1)
cov_Zfat
```

```
[1] 4.90235
```

5. Assume that passing a course, denoted as P, is a random process that follows a Bernoulli distribution and that the $Pr(P) = 0.7$.

a. What is $E(P)$ (1pt.)

$E(P) = 0.7$

b. What is $var(P)$ (1pt.)

$0.7 * 0.3 = 0.21$