

Bowen Wang

STAT 551: Statistical Genetics 2: Quantitative Genetics

Homework 3

Due 11/11/2014

1.**(a)**

$$\begin{aligned}\mu_G &= \sum_{ijkl} P(ijkl)G_{ijkl} \\ &= 3.2 \times 0.02 + 5.2 \times 0.04 + \dots + 13.4 \times 0.17 \\ &= 7.901.\end{aligned}$$

(b)Let P_A and P_a be the allele frequencies of A and a .

$$\begin{aligned}P_A &= (0.02 + 0.04 + 0.03) + \frac{1}{2}(0.07 + 0.2 + 0.15) = 0.3 \\ P_a &= (0.08 + 0.24 + 0.17) + \frac{1}{2}(0.07 + 0.2 + 0.15) = 0.7.\end{aligned}$$

Then

$$\begin{aligned}G_{A\dots} &= \sum_{jkl} P(jkl|A)G_{Aijkl} \\ &= \sum_{jkl} P(Ajkl)G_{Aijkl}/P_A \\ &= [(3.2 \times 0.02 + 5.2 \times 0.04 + 6.8 \times 0.03) + \frac{1}{2}(4.8 \times 0.07 + 5.8 \times 0.2 + 7.7 \times 0.15)]/0.3 \\ &= 6.005, \\ G_{\dots a} &= [(4.2 \times 0.08 + 9 \times 0.24 + 13.4 \times 0.17) + \frac{1}{2}(4.8 \times 0.07 + 5.8 \times 0.2 + 7.7 \times 0.15)]/0.7 \\ &\approx 8.714, \\ \alpha_A &= G_{A\dots} - \mu_G \approx -1.896 \\ \alpha_a &= G_{\dots a} - \mu_G \approx 0.813.\end{aligned}$$

Similarly,

$$\begin{aligned}P_B &= (0.02 + 0.07 + 0.08) + \frac{1}{2}(0.04 + 0.2 + 0.24) = 0.41 \\ P_b &= (0.03 + 0.15 + 0.17) + \frac{1}{2}(0.04 + 0.2 + 0.24) = 0.59 \\ G_{\dots B} &= [(3.2 \times 0.02 + 4.8 \times 0.07 + 4.2 \times 0.08) + \frac{1}{2}(5.2 \times 0.04 + 5.8 \times 0.2 + 9 \times 0.24)]/0.41 \\ &\approx 6.098 \\ G_{\dots b} &= [(6.8 \times 0.03 + 7.7 \times 0.15 + 13.4 \times 0.17) + \frac{1}{2}(5.2 \times 0.04 + 5.8 \times 0.2 + 9 \times 0.24)]/0.59 \\ &\approx 9.154 \\ \alpha_B &= G_{\dots B} - \mu_G \approx -1.803 \\ \alpha_b &= G_{\dots b} - \mu_G \approx 1.253.\end{aligned}$$

The additive variance component is given by

$$\begin{aligned}\sigma_A^2 &= \sigma^2(\alpha_i) + \sigma^2(\alpha_j) + \sigma^2(\alpha_k) + \sigma^2(\alpha_l) \\ &= 2(\alpha_A^2 P_A + \alpha_a^2 P_a + \alpha_B^2 P_B + \alpha_b^2 P_b) \\ &\approx 7.602.\end{aligned}$$

(c)

Let P_{AA} be the probability of having homozygote AA at the locus 1 etc.

$$\begin{aligned}P_{AA} &= 0.02 + 0.04 + 0.03 = 0.09 \\ P_{Aa} &= 0.07 + 0.2 + 0.15 = 0.42 \\ P_{aa} &= 0.08 + 0.24 + 0.17 = 0.49 \\ G_{AA..} &= (3.2 \times 0.02 + 5.2 \times 0.04 + 6.8 \times 0.03)/0.09 = 5.289 \\ G_{Aa..} &= (4.8 \times 0.07 + 5.8 \times 0.2 + 7.7 \times 0.15)/0.42 = 6.312 \\ G_{aa..} &= (4.2 \times 0.08 + 9 \times 0.24 + 13.4 \times 0.17)/0.49 = 9.743 \\ \delta_{AA} &= G_{AA..} - \mu_G - 2\alpha_A = 1.180 \\ \delta_{Aa} &= G_{Aa..} - \mu_G - \alpha_A - \alpha_a = -0.506 \\ \delta_{aa} &= G_{aa..} - \mu_G - 2\alpha_a = 0.217.\end{aligned}$$

Similarly,

$$\begin{aligned}P_{BB} &= 0.02 + 0.07 + 0.08 = 0.17 \\ P_{Bb} &= 0.04 + 0.2 + 0.24 = 0.48 \\ P_{bb} &= 0.03 + 0.15 + 0.17 = 0.35 \\ G_{..BB} &= (3.2 \times 0.02 + 4.8 \times 0.07 + 4.2 \times 0.08)/0.17 \approx 4.329 \\ G_{..Bb} &= (5.2 \times 0.04 + 5.8 \times 0.2 + 9 \times 0.24)/0.48 \approx 7.35 \\ G_{..bb} &= (6.8 \times 0.03 + 7.7 \times 0.15 + 13.4 \times 0.17)/0.35 \approx 10.391 \\ \delta_{BB} &= G_{..BB} - \mu_G - 2\alpha_B \approx 0.035 \\ \delta_{Bb} &= G_{..Bb} - \mu_G - \alpha_B - \alpha_b \approx -7.983 \times 10^{-4} \\ \delta_{bb} &= G_{..bb} - \mu_G - 2\alpha_b \approx -0.016.\end{aligned}$$

The dominance variance component is given by

$$\begin{aligned}\sigma_D^2 &= \sigma^2(\delta_{ij}) + \sigma^2(\delta_{kl}) \\ &= \delta_{AA}^2 P_{AA} + \delta_{Aa}^2 P_{Aa} + \delta_{aa}^2 P_{aa} + \delta_{BB}^2 P_{BB} + \delta_{Bb}^2 P_{Bb} + \delta_{bb}^2 P_{bb} \\ &\approx 0.256.\end{aligned}$$

(d)

Let P_{AB} be the probability that a randomly chosen haplotype has allele A at locus 1 and allele B at locus 2 etc. Assuming unlinked loci.

$$\begin{aligned}P_{AB} &= 0.02 + 0.5 \times (0.04 + 0.07) + 0.25 \times 0.2 = 0.125 \\ P_{Ab} &= 0.03 + 0.5 \times (0.04 + 0.15) + 0.25 \times 0.2 = 0.175 \\ P_{aB} &= 0.08 + 0.5 \times (0.07 + 0.24) + 0.25 \times 0.2 = 0.285 \\ P_{ab} &= 0.17 + 0.5 \times (0.24 + 0.15) + 0.25 \times 0.2 = 0.415 \\ G_{A.B.} &= [3.2 \times 0.02 + 0.5 \times (5.2 \times 0.04 + 4.8 \times 0.07) + 0.25 \times 5.8 \times 0.2]/0.125 \approx 5.008 \\ G_{A..b} &= [6.8 \times 0.03 + 0.5 \times (5.2 \times 0.04 + 7.7 \times 0.15) + 0.25 \times 5.8 \times 0.2]/0.175 \approx 6.717 \\ G_{.aB.} &= [4.2 \times 0.08 + 0.5 \times (4.8 \times 0.07 + 9 \times 0.24) + 0.25 \times 5.8 \times 0.2]/0.285 \approx 6.575 \\ G_{.a.b} &= [13.4 \times 0.17 + 0.5 \times (9 \times 0.24 + 7.7 \times 0.15) + 0.25 \times 5.8 \times 0.2]/0.415 \approx 10.182.\end{aligned}$$

Then

$$\begin{aligned}\alpha\alpha_{AB} &= G_{A.B.} - \mu_G - \alpha_A - \alpha_B \approx 0.806 \\ \alpha\alpha_{Ab} &= G_{A..b} - \mu_G - \alpha_A - \alpha_b \approx -0.541 \\ \alpha\alpha_{aB} &= G_{.aB.} - \mu_G - \alpha_a - \alpha_B \approx -0.335 \\ \alpha\alpha_{ab} &= G_{.a.b} - \mu_G - \alpha_a - \alpha_b \approx 0.215.\end{aligned}$$

The additive \times additive variance component is given by

$$\begin{aligned}\sigma_{AA}^2 &= \sigma^2[(\alpha\alpha)_{ik}] + \sigma^2[(\alpha\alpha)_{il}] + \sigma^2[(\alpha\alpha)_{jk}] + \sigma^2[(\alpha\alpha)_{jl}] \\ &= 4(\alpha\alpha_{AB}^2 P_{AB} + \alpha\alpha_{Ab}^2 P_{Ab} + \alpha\alpha_{aB}^2 P_{aB} + \alpha\alpha_{ab}^2 P_{ab}) \\ &\approx 0.735.\end{aligned}$$

Note: R code for calculation of this problem is at the back.

2.

(a)

Outbred second cousins share 1 allele IBD with probability $\frac{1}{16}$ and 0 allele IBD with probability $\frac{15}{16}$. State 1 is possible only if the parents share 1 allele IBD.

$$\Delta 1 = \frac{1}{16} \cdot \left(\frac{1}{2}\right)^4 = \frac{1}{256} \approx 3.906 \times 10^{-3}.$$

State 2 requires that the parents share 2 different alleles IBD, which is not possible for second cousins.

$$\Delta 2 = 0.$$

For state 3 through 6 to be possible, parents must share 1 allele IBD.

$$\begin{aligned}\Delta 3 &= \frac{1}{16} \cdot 2 \cdot \left(\frac{1}{2}\right)^4 = \frac{2}{256} \approx 7.813 \times 10^{-3} \\ \Delta 4 &= \frac{1}{16} \cdot \left(\frac{1}{2}\right)^4 = \frac{1}{256} \approx 3.906 \times 10^{-3} \\ \Delta 5 &= \frac{1}{16} \cdot 2 \cdot \left(\frac{1}{2}\right)^4 = \frac{2}{256} \approx 7.813 \times 10^{-3} \\ \Delta 6 &= \frac{1}{16} \cdot \left(\frac{1}{2}\right)^4 = \frac{1}{256} \approx 3.906 \times 10^{-3}.\end{aligned}$$

For state 7 and 8 to be possible, parents can share either 1 or 0 allele IBD.

$$\begin{aligned}\Delta 7 &= \frac{1}{16} \cdot 3 \cdot \left(\frac{1}{2}\right)^4 + \frac{15}{16} \cdot \frac{1}{4} = \frac{63}{256} \approx 0.246 \\ \Delta 8 &= \frac{1}{16} \cdot 6 \cdot \left(\frac{1}{2}\right)^4 + \frac{15}{16} \cdot \frac{1}{2} = \frac{126}{256} \approx 0.492.\end{aligned}$$

State 9 is possible only if the parents share no allele IBD.

$$\Delta 9 = \frac{15}{16} \cdot \frac{1}{4} = \frac{60}{256} \approx 0.234.$$

(b)

Kinship coefficient for Sib1 and Sib2 is given by

$$\begin{aligned}\theta_{12} &= \Delta 1 + \frac{1}{2}(\Delta 3 + \Delta 5 + \Delta 7) + \frac{1}{4}\Delta 8 \\ &= \frac{1}{256} + \frac{1}{2}\left(\frac{2}{256} + \frac{2}{256} + \frac{63}{256}\right) + \frac{1}{4} \cdot \frac{126}{256} \\ &= \frac{66}{256} \\ &\approx 0.258.\end{aligned}$$

(c)

Inbreeding coefficient for Sib1 is just the kinship coefficient for the parents. Since the parents are second cousins, $F_1 = \frac{1}{64}$.

(d)

We can condition on whether the parents share 1 allele IBD. When they do, the child can be inbred or outbred possibly sharing 2 different alleles IBD with the father. Then it follows

$$\theta = \frac{1}{16} \cdot \left(\frac{1}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4}\right) + \frac{15}{16} \cdot \frac{1}{4} = \frac{3+30}{128} \approx 0.258.$$

R code for problem 1

```
# part (a)
prob=matrix(c(0.02,0.04,0.03,0.07,0.2,0.15,0.08,0.24,0.17),nrow=3,
            ncol=3,byrow=T)
geno=matrix(c(3.2,5.2,6.8,4.8,5.8,7.7,4.2,9,13.4),nrow=3,ncol=3,
            byrow=T)
muG=sum(geno*prob)
muG

# part (b)
PA=sum(prob[1,])+0.5*sum(prob[2,])
Pa=sum(prob[3,])+0.5*sum(prob[2,])
PA
Pa

GA...=(sum((geno*prob)[1,])+sum((geno*prob)[2,])*0.5)/PA
GA...
G.a...=(sum((geno*prob)[3,])+sum((geno*prob)[2,])*0.5)/Pa
G.a...
alphaA=GA...-muG
alphaA
alphaa=G.a...-muG
alphaa

PB=sum(prob[,1])+0.5*sum(prob[,2])
PB
Pb=sum(prob[,3])+0.5*sum(prob[,2])
Pb

G..B.=(sum((geno*prob)[,1])+sum((geno*prob)[,2])*0.5)/PB
G..B.
G...b=(sum((geno*prob)[,3])+sum((geno*prob)[,2])*0.5)/Pb
G...b
alphaB=G..B.-muG
alphaB
alphab=G...b-muG
alphab

sigmaA=2*(alphaA^2*PA+alphaa^2*Pa+alphaB^2*PB+alphab^2*Pb)
sigmaA

# part (c)
PAA=sum(prob[1,])
PAA
PAA=sum(prob[2,])
PAA
Paa=sum(prob[3,])
Paa

GAA...=sum(geno[1,]*prob[1,])/PAA
GAA...
GAA...=sum(geno[2,]*prob[2,])/PAA
GAA...
```

```

Gaa.=sum(geno[3,]*prob[3,])/Paa
Gaa.
deltaAA=GAA.-muG-2*alphaA
deltaAA
deltaAa=GAA.-muG-alphaA-alphaa
deltaAa
deltaaaa=Gaa.-muG-2*alphaa
deltaaaa

PBB=sum(prob[,1])
PBB
PBb=sum(prob[,2])
PBb
Pbb=sum(prob[,3])
Pbb

G..BB=sum(geno[,1]*prob[,1])/PBB
G..BB
G..Bb=sum(geno[,2]*prob[,2])/PBb
G..Bb
G..bb=sum(geno[,3]*prob[,3])/Pbb
G..bb
deltaBB=G..BB-muG-2*alphaB
deltaBB
deltaBb=G..Bb-muG-alphaB-alphab
deltaBb
deltabb=G..bb-muG-2*alphab
deltabb

sigmaD=deltaAA^2*PAA+deltaAa^2*PAa+deltaaaa^2*Paa+deltaBB^2*PBB+
deltaBb^2*PBb+deltabb^2*Pbb
sigmaD

# part (d)
PAB=sum(prob[-3,-3]*matrix(c(1,0.5,0.5,0.25),nrow=2,ncol=2))
PAB
PAb=sum(prob[-3,-1]*matrix(c(0.5,0.25,1,0.5),nrow=2,ncol=2))
PAb
PaB=sum(prob[-1,-3]*matrix(c(0.5,1,0.25,0.5),nrow=2,ncol=2))
PaB
Pab=sum(prob[-1,-1]*matrix(c(0.25,0.5,0.5,1),nrow=2,ncol=2))
Pab

GA.B.=sum(geno[-3,-3]*prob[-3,-3]*matrix(c(1,0.5,0.5,0.25),
nrow=2,ncol=2))/PAB
GA..b=sum(geno[-3,-1]*prob[-3,-1]*matrix(c(0.5,0.25,1,0.5),
nrow=2,ncol=2))/PAb
G.aB.=sum(geno[-1,-3]*prob[-1,-3]*matrix(c(0.5,1,0.25,0.5),
nrow=2,ncol=2))/PaB
G.a.b=sum(geno[-1,-1]*prob[-1,-1]*matrix(c(0.25,0.5,0.5,1),
nrow=2,ncol=2))/Pab

GA.B.
GA..b
G.aB.

```

G.a.b

$\alpha_{AB} = GA.B - \mu_G - \alpha_A - \alpha_B$

α_{AB}

$\alpha_{Ab} = GA..b - \mu_G - \alpha_A - \alpha_b$

α_{Ab}

$\alpha_{aB} = G.aB. - \mu_G - \alpha_a - \alpha_B$

α_{aB}

$\alpha_{ab} = G.a.b - \mu_G - \alpha_a - \alpha_b$

α_{ab}

$\sigma_{AA} = 4 * (\alpha_{AB}^2 * P_{AB} + \alpha_{Ab}^2 * P_{Ab} + \alpha_{aB}^2 * P_{aB} + \alpha_{ab}^2 * P_{ab})$

σ_{AA}