

SOLUTIONS TO EXERCISES FOR CHAPTER 9

**Exercise 9.1.**

a. Algebraic method for the grandparent-grandoffspring covariance (see parent-offspring covariance, pp. 9.6 to 9.13):

One needs to set up the following table to express the covariance between a grandparent and a grandoffspring.

Grandparent		Offspring			Conditional probability of grandoffspring given offspring			Joint frequencies			
Geno	Freq	Geno	Cond freq	Freq	$A_1A_1$	$A_1A_2$	$A_2A_2$	$A_1A_1$	$A_1A_2$	$A_2A_2$	Total
$A_1A_1$	$p_1^2$	$A_1A_1$	$p_1$	$p_1^3$	$p_1$	$p_2$		$p_1^4$	$p_1^3 p_2$		
		$A_1A_2$	$p_2$	$\frac{p_1^2 p_2}{p_1^2}$	$\frac{1}{2} p_1$	$\frac{1}{2}$	$\frac{1}{2} p_2$	$\frac{1}{2} p_1^3 p_2$	$\frac{1}{2} p_1^2 p_2$	$\frac{1}{2} p_1^2 p_2^2$	
				$p_1^2$				$p_1^3 (p_1 + \frac{1}{2} p_2)$	$p_1^2 p_2 (p_1 + \frac{1}{2})$	$\frac{1}{2} p_1^2 p_2^2$	$p_1^2$
$A_1A_2$	$2 p_1 p_2$	$A_1A_1$	$\frac{1}{2} p_1$	$\frac{p_1^2 p_2}{2 p_1 p_2}$	$p_1$	$p_2$		$p_1^3 p_2$	$p_1^2 p_2^2$		
		$A_1A_2$	$\frac{1}{2}$	$\frac{p_1 p_2}{2 p_1 p_2}$	$\frac{1}{2} p_1$	$\frac{1}{2}$	$\frac{1}{2} p_2$	$\frac{1}{2} p_1^2 p_2$	$\frac{1}{2} p_1 p_2$	$\frac{1}{2} p_1 p_2^2$	
		$A_2A_2$	$\frac{1}{2} p_2$	$\frac{p_1 p_2^2}{2 p_1 p_2}$		$p_1$	$p_2$		$\frac{p_1^2 p_2^2}{2 p_1 p_2}$	$\frac{p_1 p_2^3}{2 p_1 p_2}$	
				$2 p_1 p_2$				$p_1^2 p_2 (p_1 + \frac{1}{2})$	$p_1 p_2^2 (\frac{1}{2} + p_2)$	$2 p_1 p_2$	
								$p_1 p_2 (p_1 p_2 + \frac{1}{2} + p_1 p_2)$			
$A_2A_2$	$p_2^2$	$A_1A_2$	$p_1$	$\frac{p_1 p_2^2}{p_2^2}$	$\frac{1}{2} p_1$	$\frac{1}{2}$	$\frac{1}{2} p_2$	$\frac{1}{2} p_1^2 p_2^2$	$\frac{1}{2} p_1 p_2^2$	$\frac{1}{2} p_1 p_2^3$	
		$A_2A_2$	$p_2$	$\frac{p_2^3}{p_2^2}$		$p_1$	$p_2$		$\frac{p_1 p_2^3}{p_2^2}$	$\frac{p_2^4}{p_2^2}$	
				$p_2^2$					$\frac{1}{2} p_1^2 p_2^2$	$p_1 p_2^2 (\frac{1}{2} + p_2)$	$p_2^3 (\frac{1}{2} p_1 + p_2)$

The desired grandparent-grandoffspring covariance is

$$C(X, Y) = (2\alpha_1 + \delta_{11}) \left[ p_1^3 (p_1 + \frac{1}{2} p_2) (2\alpha_1 + \delta_{11}) + p_1^2 p_2 (p_1 + \frac{1}{2}) (\alpha_1 + \alpha_2 + \delta_{12}) + \frac{1}{2} p_1^2 p_2^2 (2\alpha_2 + \delta_{22}) \right] \\ + (\alpha_1 + \alpha_2 + \delta_{12}) \left[ p_1^2 p_2 (p_1 + \frac{1}{2}) (2\alpha_1 + \delta_{11}) + p_1 p_2 (2 p_1 p_2 + \frac{1}{2}) (\alpha_1 + \alpha_2 + \delta_{12}) + p_1 p_2^2 (\frac{1}{2} + p_2) (2\alpha_2 + \delta_{22}) \right]$$

$$+(2\alpha_2 + \delta_{22})\left[\frac{1}{2}p_1^2 p_2^2 (2\alpha_1 + \delta_{11}) + p_1 p_2^2 \left(\frac{1}{2} + p_2\right)(\alpha_1 + \alpha_2 + \delta_{12}) + p_2^3 \left(\frac{1}{2} p_1 + p_2\right)(2\alpha_2 + \delta_{22})\right]$$

The problem before us is to simplify the above, which can be a very tedious frustrating problem. The challenge before us is to group quantities which can be simplified because  $p_1 + p_2 = 1$ , or because  $p_1\alpha_1 + p_2\alpha_2 = 0$ ,  $p_1\delta_{11} + p_2\delta_{12} = 0$ , or  $p_1\delta_{12} + p_2\delta_{22} = 0$ . Although we will present the simplification of all major terms in the whole expression, there are often symmetries in these kinds of problems which can reduce the labor involved. For example, the expression enclosed in brackets in the first major term is symmetrical to that in the third major term, i.e., by simply changing all subscript 1s to 2s and 2s to 1s on all quantities in the first major term, one obtains that in the third term. Hence, one really needs to simplify only the quantities enclosed in brackets in the first major term and then change the subscripts from 1 to 2 and 2 to 1 in the simplified expression from the first major term to obtain the expression for the third major term.

$$\begin{aligned} C(X, Y) &= (2\alpha_1 + \delta_{11})\left\{p_1^2 \left[p_1 \left(p_1 + \frac{1}{2} p_2\right)(2\alpha_1 + \delta_{11}) + p_2 \left(p_1 + \frac{1}{2}\right)(\alpha_1 + \alpha_2 + \delta_{12}) + \frac{1}{2} p_2^2 (2\alpha_2 + \delta_{22})\right]\right\} \\ &\quad + (\alpha_1 + \alpha_2 + \delta_{12})\left\{p_1 p_2 \left[p_1 \left(p_1 + \frac{1}{2}\right)(2\alpha_1 + \delta_{11}) + \left(2p_1 p_2 + \frac{1}{2}\right)(\alpha_1 + \alpha_2 + \delta_{12}) + p_2 \left(\frac{1}{2} + p_2\right)(2\alpha_2 + \delta_{22})\right]\right\} \\ &\quad + (2\alpha_2 + \delta_{22})\left\{p_2^2 \left[\frac{1}{2} p_1^2 (2\alpha_1 + \delta_{11}) + p_1 \left(\frac{1}{2} + p_2\right)(\alpha_1 + \alpha_2 + \delta_{12}) + p_2 \left(\frac{1}{2} p_1 + p_2\right)(2\alpha_2 + \delta_{22})\right]\right\} \\ &= (2\alpha_1 + \delta_{11})\left\{p_1^2 \left[\left(p_1^2 + \frac{1}{2} p_1 p_2\right)(2\alpha_1 + \delta_{11}) + \left(p_1 p_2 + \frac{1}{2} p_2\right)(\alpha_1 + \alpha_2 + \delta_{12}) + \frac{1}{2} p_2^2 (2\alpha_2 + \delta_{22})\right]\right\} \\ &\quad + (\alpha_1 + \alpha_2 + \delta_{12})\left\{p_1 p_2 \left[\left(p_1^2 + \frac{1}{2} p_1\right)(2\alpha_1 + \delta_{11}) + \left(2p_1 p_2 + \frac{1}{2}\right)(\alpha_1 + \alpha_2 + \delta_{12}) + \left(\frac{1}{2} p_2 + p_2^2\right)(2\alpha_2 + \delta_{22})\right]\right\} \\ &\quad + (2\alpha_2 + \delta_{22})\left\{p_2^2 \left[\frac{1}{2} p_1^2 (2\alpha_1 + \delta_{11}) + \left(\frac{1}{2} p_1 + p_1 p_2\right)(\alpha_1 + \alpha_2 + \delta_{12}) + \left(\frac{1}{2} p_1 p_2 + p_2^2\right)(2\alpha_2 + \delta_{22})\right]\right\} \end{aligned}$$

Then grouping like terms, we have

$$\begin{aligned} C(X, Y) &= (2\alpha_1 + \delta_{11})\left\{p_1^2 \left[\left(\frac{1}{2} p_1^2 + \frac{1}{2} p_1^2 + \frac{1}{2} p_1 p_2\right)(2\alpha_1) + \left(p_1 p_2 + \frac{1}{2} p_2\right)\alpha_1 + \left(p_1 p_2 + \frac{1}{2} p_2\right)\alpha_2 + \frac{1}{2} p_2^2 (2\alpha_2)\right.\right. \\ &\quad \left.\left.+ \left(\frac{1}{2} p_1^2 + \frac{1}{2} p_1^2 + \frac{1}{2} p_1 p_2\right)\delta_{11} + \left(\frac{1}{2} p_1 p_2 + \frac{1}{2} p_1 p_2 + \frac{1}{2} p_2\right)\delta_{12} + \frac{1}{2} p_2^2 \delta_{22}\right]\right\} \\ &\quad + (\alpha_1 + \alpha_2 + \delta_{12})\left\{p_1 p_2 \left[\left(p_1^2 + \frac{1}{2} p_1\right)2\alpha_1 + \left(2p_1 p_2 + \frac{1}{2}\right)\alpha_1 + \left(2p_1 p_2 + \frac{1}{2}\right)\alpha_2 + \left(\frac{1}{2} p_2 + p_2^2\right)2\alpha_2\right.\right. \\ &\quad \left.\left.+ \left(p_1^2 + \frac{1}{2} p_1\right)\delta_{11} + \left(p_1 p_2 + p_1 p_2 + \frac{1}{2}\right)\delta_{12} + \left(\frac{1}{2} p_2 + p_2^2\right)\delta_{22}\right]\right\} \\ &\quad + (2\alpha_2 + \delta_{22})\left\{p_2^2 \left[\frac{1}{2} p_1^2 (2\alpha_1) + \left(\frac{1}{2} p_1 + p_1 p_2\right)\alpha_1 + \left(\frac{1}{2} p_1 + p_1 p_2\right)\alpha_2 + \left(\frac{1}{2} p_1 p_2 + \frac{1}{2} p_2^2 + \frac{1}{2} p_2^2\right)(2\alpha_2)\right.\right. \end{aligned}$$



$$\begin{aligned}
C(X, Y) &= (2\alpha_1 + \delta_{11}) \left\{ p_1^2 \left[ \frac{1}{2} p_1 \alpha_1 + \frac{1}{2} p_1 \alpha_1 + \frac{1}{2} p_2 \alpha_1 + \frac{1}{2} p_2 \alpha_2 + \frac{1}{2} (p_1 \delta_{11} + p_2 \delta_{12}) \right] \right\} \\
&\quad + (\alpha_1 + \alpha_2 + \delta_{12}) \left\{ p_1 p_2 \left[ \frac{1}{2} (\alpha_1 + \alpha_2) \right] \right\} \\
&\quad + (2\alpha_2 + \delta_{22}) \left\{ p_2^2 \left[ \frac{1}{2} p_1 \alpha_1 + \frac{1}{2} p_1 \alpha_2 + \frac{1}{2} p_2 \alpha_2 + \frac{1}{2} p_2 \alpha_2 + \frac{1}{2} (p_1 \delta_{12} + p_2 \delta_{22}) \right] \right\} \\
&= (2\alpha_1 + \delta_{11}) \left\{ p_1^2 \left[ \frac{1}{2} (p_1 + p_2) \right] \alpha_1 \right\} \\
&\quad + (\alpha_1 + \alpha_2 + \delta_{12}) \left\{ p_1 p_2 \left[ \frac{1}{2} (\alpha_1 + \alpha_2) \right] \right\} \\
&\quad + (2\alpha_2 + \delta_{22}) \left\{ p_2^2 \left[ \frac{1}{2} (p_1 + p_2) \alpha_2 \right] \right\} \\
&= (2\alpha_1 + \delta_{11}) \left( \frac{1}{2} p_1^2 \alpha_1 \right) + (\alpha_1 + \alpha_2 + \delta_{12}) \left[ \frac{1}{2} p_1 p_2 (\alpha_1 + \alpha_2) \right] + (2\alpha_2 + \delta_{22}) \left( \frac{1}{2} p_2^2 \alpha_2 \right) \\
&= \frac{1}{2} \left[ (2\alpha_1 + \delta_{11}) \left( p_1^2 \alpha_1 \right) + (\alpha_1 + \alpha_2 + \delta_{12}) \left[ p_1 p_2 (\alpha_1 + \alpha_2) \right] + (2\alpha_2 + \delta_{22}) \left( p_2^2 \alpha_2 \right) \right] \\
&= \frac{1}{2} C(P, O) \text{ as derived in Box 9.1 (middle of p. 9.10)} \\
&= \frac{1}{2} \left( \frac{1}{2} \sigma_A^2 \right) \\
&= \frac{1}{4} \sigma_A^2
\end{aligned}$$

b. The expectation method (see pp. 9.22 to 9.23):

First, we must derive the deviation of the average genotypic value of the grandoffspring in terms of the average effects of the grandparent, i.e.,  $\alpha_i$  and  $\alpha_j$

$$\text{gametic array of grandparent} = \frac{1}{2} A_i + \frac{1}{2} A_j = \frac{1}{2} (A_i + A_j)$$

$$\text{gametic array of population} = \sum_{k=1}^2 p_k A_k$$

$$\text{genotypic array of offspring} = \frac{1}{2} (A_i + A_j) \left( \sum_{k=1}^2 p_k A_k \right)$$

$$\begin{aligned}
\text{gametic array of offspring} &= \frac{1}{2} \left[ \frac{1}{2} (A_i + A_j) + \sum_{k=1}^2 p_k A_k \right] \\
&= \frac{1}{4} A_i + \frac{1}{4} A_j + \frac{1}{2} \sum_{k=1}^2 p_k A_k
\end{aligned}$$

$$\begin{aligned} \text{genotypic array of grandoffspring} &= \left( \frac{1}{4} A_i + \frac{1}{4} A_j + \frac{1}{2} \sum_{k=1}^2 p_k A_k \right) \left( \sum_{k=1}^2 p_k A_k \right) \\ &= \frac{1}{4} \sum_{k=1}^2 p_k A_i A_k + \frac{1}{4} \sum_{k=1}^2 p_k A_j A_k + \frac{1}{2} \left( \sum_{k=1}^2 p_k A_k \right) \left( \sum_{k=1}^2 p_k A_k \right) \end{aligned}$$

$$\text{average genotypic value of grandoffspring} = \frac{1}{4} \sum_{k=1}^2 p_k G_{ik} + \frac{1}{4} \sum_{k=1}^2 p_k G_{jk} + \frac{1}{2} \sum_{k=1}^2 \sum_{k'=1}^2 p_k p_{k'} G_{kk'}$$

$$\begin{aligned} \left( \begin{array}{c} \text{deviation of average genotypic value} \\ \text{of grandoffspring} \end{array} \right) &= \frac{1}{4} \left( \sum_{k=1}^2 p_k G_{ik} - G_{..} \right) + \frac{1}{4} \left( \sum_{k=1}^2 p_k G_{jk} - G_{..} \right) + \frac{1}{2} \left( \sum_{k=1}^2 \sum_{k'=1}^2 p_k p_{k'} G_{kk'} - G_{..} \right) \\ &= \frac{1}{4} \alpha_i + \frac{1}{4} \alpha_j \\ &= \frac{1}{4} (\alpha_i + \alpha_j) \end{aligned}$$

One may now express the covariance between grandparent and grandoffspring by the expectation method:

$$\begin{aligned} C(X, Y) &= E(X - \mu)(Y - \mu) \\ &= E(\alpha_i + \alpha_j + \delta_{ij}) \left[ \frac{1}{4} (\alpha_i + \alpha_j) \right] \\ &= \frac{1}{4} E(\alpha_i^2 + 2\alpha_i \alpha_j + \alpha_j^2 + \alpha_i \delta_{ij} + \alpha_j \delta_{ij}) \\ &= \frac{1}{4} E(\alpha_i^2) + \frac{2}{4} E(\alpha_i \alpha_j) + \frac{1}{4} E(\alpha_j^2) + \frac{1}{4} E(\alpha_i \delta_{ij}) + \frac{1}{4} E(\alpha_j \delta_{ij}) \\ &= \frac{1}{4} \sigma_\alpha^2 + \frac{1}{4} \sigma_\alpha^2 \\ &= \frac{1}{4} (2\sigma_\alpha^2) \\ &= \frac{1}{4} \sigma_A^2 \end{aligned}$$

c. Grandparent-grandoffspring covariance by the general, identity-by-descent method by use of equation (9.63):

$$\begin{aligned} \theta_{XY} = \theta_{X(O.A)} &= \frac{1}{2} (\theta_{XO} + \theta_{XA}) \\ &= \frac{1}{2} \theta_{XO} + \frac{1}{2} \theta_{XA} \\ &= \frac{1}{2} \theta_{XO} + 0 \\ &= \frac{1}{2} \left( \frac{1}{4} \right) = \frac{1}{8} \quad \delta_{d_{XY}} = 0 \end{aligned}$$

$$\begin{aligned} C(X, Y) &= 2\theta_{XY} \sigma_A^2 + \delta_{d_{XY}} \sigma_D^2 \\ &= 2 \left( \frac{1}{8} \right) \sigma_A^2 \\ &= \frac{1}{4} \sigma_A^2 \end{aligned}$$

**Exercise 9.2.**

The reason is that for any particular parental genotype the weighted sum of the deviations of the offspring genotypic values from the true offspring mean for that parental genotype sums to zero. Hence, their product with the parental genotypic value contributes nothing to the overall covariance. That is, the same value for the parent-offspring covariance is obtained using different genotypic values for the different offspring genotypes associated with that particular parental genotype, as one obtains using the single mean offspring genotypic value itself for the different genotypes associated with that particular parental genotype. This is discussed on pp. 9.11 and 9.12.

**Exercise 9.3.**

a. In each of the three cases, two alleles, three alleles, and four alleles, we are concerned with the probability of one of the possible states of identity by descent involving two, three or four random alleles, respectively. In the case of two-allele probability function the two random alleles may be in the same individual in which case the two-allele probability function is known as the coefficient of inbreeding or the two random alleles may be in two individuals (the same or different individuals) in which case the two-allele probability function is known as the coefficient of coancestry. In the case of three-allele probability function the three random alleles may be from three different individuals, two in one individual and a random gene in a second individual. In the case of four-allele probability function the four random alleles may be from four different individuals, from three individuals with two alleles in one individual and a random gene from each of two other individuals, and from two individuals with two alleles in one individual and two alleles in a second individual.

b. There are 2 states of identity by descent for 2-allele probability function (see Section 4.7, pp. 4.58 to 4.62). There are 5 states of identity by descent for 3-allele probability function (see Section 4.8, pp. 4.62 to 4.67). There are 15 states of identity by descent for 4-allele probability function (see Section 4.9, pp. 4.67 to 4.77). ERRORS on p. 4.75 in equation (4.82): 1) on  $\Delta_2$  line, change " $= 2\delta_{\ddot{X}\ddot{Y}}$ " to " $= 2\delta_{\ddot{X}\ddot{Y}}$ ", and 2) on  $\Delta_9$  line, delete " $+ \delta_{13} + \delta_{14}$ ". See Errata for Statistical Genetics Notes.

c. For three-allele probability function there are five states so we need 4 parameters which is one less than the number of states. Since there are three alleles we may use  $\binom{3}{2} = 3$  two-allele parameters, leaving only one parameter undefined. We define this remaining parameter as the one for all three alleles identical by descent.

d. For four-allele probability functions there are 15 states so we need 14 parameters which is one less than the number of states. Using previously defined two-allele and three-allele probability functions we may express 10 of the 14 parameters as functions of two- and three-allele probability functions. With four alleles there are  $\binom{4}{2} = 6$  two-allele probability functions and  $\binom{4}{3} = 4$  three-allele probability functions, summing to 10.

e. Coefficient of dominance coancestry is a four-allele probability function and is the sum of two of the 15 states of identity by descent, namely,

$$\delta_{d_{XY}} = P_4(a_X^m \equiv a_Y^m; a_X^f \equiv a_Y^f) + P_4(a_X^m \equiv a_Y^f; a_X^f \equiv a_Y^m) = \Delta_5 = 2\delta_{\ddot{X}+\ddot{Y}} = \delta_{ac-bd} + \delta_{ad-bc}$$

f. We have considered various two-locus, identical-by-descent functions:

1) Two genes identical by descent at two loci:

$F_{11}$  = joint probability of identity by descent at two loci in the same individual (see equation (8.175D))

$\theta_{2,mm}$  = joint probability of identity by descent of random genes in two individuals at each of two loci (I

may not have introduced this symbol per se in my notes). In this case, we have specifically considered all four genes to be from the male parent, namely,

$$\theta_{2,mm} = P\left(a_X^m \equiv a_Y^m; b_X^m \equiv b_Y^m\right) \quad [\text{see equation (9.150)}]$$

2) Four genes at one locus and two genes at another locus:

$$P\left(a_X^m \equiv a_Y^m; a_X^f \equiv a_Y^f; b_X^m \equiv b_Y^m\right)$$

4) Four genes at one locus and four genes at another locus:

$$P\left(a_X^m \equiv a_Y^m; a_X^f \equiv a_Y^f; b_X^m \equiv b_Y^m; b_X^f \equiv b_Y^f\right)$$

#### Exercise 9.4.

a. The two events are mutually exclusive and hence their combined probability can be written as a sum of the two events. In the first case, the male genes in  $X$  and  $Y$  are IBD and likewise the female genes are IBD. In the second case the male gene in  $X$  is IBD to the female gene in  $Y$  and female gene in  $X$  is IBD to the male gene in  $Y$ .

b. random mating

c. independence between loci, i.e.,  $\rho_1 = 0.5$

d. linkage equilibrium

e. Hardy Weinberg equilibrium, i.e., equal gene frequencies in both males and females. Specifically the frequencies of the  $B$  alleles in the males are the same as the  $B$  alleles in the females.

#### Exercise 9.5.

In general, the condition is that all the genes that determine the effect in one relative are identical by descent to all the corresponding genes in the other relative.

For one pair of genes, the two genes, one in  $X$  and one in  $Y$ , must be identical by descent before their corresponding effects can contribute to the covariance between relatives. This is applicable to the additive variance, i.e., the gene causing the additive effect in individual  $X$  must be identical by descent to the gene causing the additive effect in individual  $Y$ , before a contribution to the additive variance in the covariance between relatives can occur. When the two genes are identical by descent, the effects are the same and upon squaring them in the covariance expression between  $X$  and  $Y$ , the additive variance results. The probability of the two genes being identical by descent in the covariance expression is equal to the fraction of the additive variance in the covariance of relatives.

For two pairs of genes, both of the pairs of genes must be identical by descent before their corresponding effects can contribute to the covariance between relatives, i.e., one gene in  $X$  must be identical by descent to a gene in  $Y$ , and another gene in  $X$  must be identical by descent to another gene in  $Y$ . When this occurs, the effect in individual  $X$  is the same as the effect in individual  $Y$ . The two pairs may involve the same locus or two different loci. This is applicable to the dominance and additive-by-additive variances. For the dominance variance, two genes at the same locus determine the dominance effect in an individual. One gene in  $X$  must be identical by descent to a gene in  $Y$ , and another gene in  $X$  at the same locus must be identical by descent to another gene in  $Y$ . When each of the two pairs of genes is identical by descent, the dominance effects are the same in individuals  $X$  and  $Y$  and upon squaring them in the covariance expression, the dominance variance results. The probability of the two pairs of genes being identical by descent in the covariance expression is equal to the fraction of the dominance variance in the covariance of relatives. For the additive-by-additive variance, two genes determine the additive-by-additive effect in an individual, one gene at one locus and another gene at a second locus. One gene in  $X$  at the first locus must be identical by descent to a gene in  $Y$  at that first locus, and another gene in  $X$  at the second locus must be identical by descent to another gene in  $Y$  at the second locus. When each of the two pairs of genes is identical by descent—one pair at the first locus and the second pair at the second locus, the additive-by-additive effects are the same and upon squaring them in the covariance expression, the additive-by-additive variance results. The probability of the two pairs of genes being identical by descent in the covariance expression is equal to the fraction of the additive-by-additive variance in the covariance of relatives.

When three pairs of genes are identical by descent—two pairs at one locus and another pair at another locus, it determines the fraction of the dominance-by-additive or additive-by-dominance variance.

When four pairs of genes are identical by descent—two pairs at one locus and two pairs at another locus, it determines the fraction of the dominance-by-dominance variance.

### Exercise 9.6.

The more important conditions and/or assumptions that the general formula for the covariance of relatives is based on are:

--parents are random, unrelated (so that the relatives are noninbred), and noninbred, partially, or fully inbred individuals from a reference population

--the reference population assumes:

--Mendelian, disomic inheritance

--Hardy-Weinberg equilibrium (random mating)

--linkage equilibrium for all two, three, four, or more loci

--arbitrary number of segregating loci

--two or more (multiple) alleles per locus

--independence between all pairs of loci

--arbitrary gene action (arbitrary level of dominance and arbitrary epistasis)

Other assumptions are listed on pp. 9.68 and 9.69.

### Exercise 9.7.

$$\theta_{mm} = \theta_{DG} = 0$$

$$\theta_{mf} = \theta_{DF} = \left(\frac{1}{2}\right)^4 (1 + F_A) \text{ by equation (4.57), or alternatively}$$

$$= \theta_{DF} = \theta_{(C-)F} = \frac{1}{2}\theta_{CF} = \frac{1}{2}\theta_{(A-)F} = \frac{1}{4}\theta_{AF} = \frac{1}{4}\theta_{A(A-)} = \frac{1}{8}\theta_{AA} = \frac{1}{16}(1 + F_A)$$

$$\theta_{fm} = \theta_{EG} = \left(\frac{1}{2}\right)^3 (1 + F_B) \text{ by equation (4.57), or alternatively}$$

$$= \theta_{EG} = \theta_{(-B)G} = \frac{1}{2}\theta_{BG} = \frac{1}{2}\theta_{B(B-)} = \frac{1}{4}\theta_{BB} = \frac{1}{8}(1 + F_B)$$

$$\theta_{ff} = \theta_{EF} = 0$$

$$\delta_{d_{XY}} = \theta_{mf}\theta_{fm} = \left(\frac{1}{4}\right)^4 (1 + F_A)\left(\frac{1}{2}\right)^3 (1 + F_B)$$

### Exercise 9.8.

a. We use equation (4.48)

$$\theta_{XY} = \frac{1}{4}(\theta_{FL} + \theta_{FM} + \theta_{GL} + \theta_{GM})$$

and calculate  $\theta_{FL}$ ,  $\theta_{FM}$ ,  $\theta_{GL}$ , and  $\theta_{GM}$  below by the chain of coancestry method [equation (4.57)].

$$\theta_{FL} = \theta_{mm} = \left(\frac{1}{2}\right)^4 (1 + F_B) = \frac{1 + F_B}{16}$$

$$\theta_{FM} = \theta_{mf} = 0$$

$$\theta_{GL} = \theta_{fm} = 0$$

$$\theta_{GM} = \theta_{ff} = \left(\frac{1}{2}\right)^4 (1 + F_C) + \left(\frac{1}{2}\right)^4 (1 + F_D) = \frac{1 + F_C + 1 + F_D}{16} = \frac{2 + F_C + F_D}{16}$$

$$\theta_{XY} = \frac{1}{4}(\theta_{FL} + \theta_{FM} + \theta_{GL} + \theta_{GM}) = \frac{1}{4}\left(\frac{1 + F_B}{16} + 0 + 0 + \frac{2 + F_C + F_D}{16}\right) = \frac{1}{64}(3 + F_B + F_C + F_D)$$

b. The coefficient of dominance coancestry calculated by the simplified method is [see equation (9.82)]

$$\delta_{dXY} = \Delta_5 = \theta_{mm}\theta_{ff} = \left(\frac{1+F_B}{16}\right)\left(\frac{2+F_C+F_D}{16}\right)$$

The basic idea behind the simplified method of calculating the coefficient of dominance coancestry is (see Section 9.3.2, pp. 9.42 and 9.43): If the coefficient of coancestry between X and Y on the male side of the pedigree is independent or completely separate (uninfluenced) from the coefficient of coancestry between X and Y on the female side of the pedigree, then the coefficient of dominance coancestry can be calculated by simply multiplying the coefficients for the two sides together.

c. For  $F_B = F_C = F_D = 0$ , the covariance between relatives is [see equation (9.144)]

$$2\theta_{XY} = 2\left[\frac{1}{64}(3+F_B+F_C+F_D)\right] = \frac{6}{64} = \frac{3}{32}$$

$$\delta_{dXY} = \left(\frac{1+F_B}{16}\right)\left(\frac{2+F_C+F_D}{16}\right) = \left(\frac{1}{16}\right)\left(\frac{2}{16}\right) = \frac{2}{256} = \frac{1}{128}$$

$$\text{Cov}(X, Y) = \frac{3}{32}\sigma_A^2 + \frac{1}{128}\sigma_D^2 + \left(\frac{3}{32}\right)^2\sigma_{AA}^2 + \left(\frac{3}{32}\right)\left(\frac{1}{128}\right)\sigma_{AD}^2 + \left(\frac{1}{128}\right)^2\sigma_{DD}^2 + \dots$$

For  $F_B = 0, F_C = \frac{1}{2}, F_D = 1$ , the covariance between relatives is

$$2\theta_{XY} = 2\left[\frac{1}{64}(3+F_B+F_C+F_D)\right] = 2\left[\frac{1}{64}\left(3+0+\frac{1}{2}+1\right)\right] = 2\frac{1}{64}\frac{9}{2} = \frac{9}{64}$$

$$\delta_{dXY} = \left(\frac{1+F_B}{16}\right)\left(\frac{2+F_C+F_D}{16}\right) = \left(\frac{1+0}{16}\right)\left(\frac{2+\frac{1}{2}+1}{16}\right) = \left(\frac{1}{16}\right)\left(\frac{7}{32}\right) = \frac{7}{512}$$

$$\text{Cov}(X, Y) = \frac{9}{64}\sigma_A^2 + \frac{7}{512}\sigma_D^2 + \left(\frac{9}{64}\right)^2\sigma_{AA}^2 + \left(\frac{9}{64}\right)\left(\frac{7}{512}\right)\sigma_{AD}^2 + \left(\frac{7}{512}\right)^2\sigma_{DD}^2 + \dots$$

### Exercise 9.9.

a. Two-locus linkage equilibrium is defined as  $p_{A_i B_j} = p_{A_i} p_{B_j}$ . Three-locus linkage disequilibrium is defined as  $p_{A_i B_j C_k} \neq p_{A_i} p_{B_j} p_{C_k}$ . Yes, one can have two-locus linkage equilibrium between all pairs but not have three-locus linkage equilibrium. See pp. 3.85 and 3.86.

b. Linkage disequilibrium is the measure of association of genes at different loci in a population whereas genetic linkage refers to the association of genes at different loci that occurs in the gametic output from an individual. Yes, one can have linkage disequilibrium between genes at two loci on separate chromosomes.

c. Genetic linkage relates to linkage disequilibrium only to the extent that it affects the rate of dissipation of linkage disequilibrium.

d. Linkage disequilibrium has no effect upon the population mean in the absence of epistasis. Linkage disequilibrium does affect the population mean when epistasis is present as shown in equation (8.171).

e. Linkage disequilibrium affects the gametic genetic variance and hence the genotypic variance whether epistasis is absent or present. When epistasis is absent, the gametic genetic variance is

$$\begin{aligned} \sigma_g^2 &= \sigma_{\alpha_a}^2 + \sigma_{\alpha_b}^2 + 2\text{Cov}(\alpha_a, \alpha_b) \\ &= \sigma_{\alpha_a}^2 + \sigma_{\alpha_b}^2 + \frac{2\Delta}{\sqrt{p_{A_1} p_{A_2} p_{B_1} p_{B_2}}} \sigma_{\alpha_a} \sigma_{\alpha_b} \end{aligned} \quad \text{equation (8.173)}$$

When epistasis is present, the gametic genetic variance is also affected, namely,

$$\sigma_g^2 = \sum_i \sum_k \left[ p_{A_i} p_{B_k} + (-1)^{i+k} \Delta \right] (G_{i.k.} - G_{\dots})^2 \quad \text{equation (8.172)} \quad \left[ \sum_i \sum_k \text{ was omitted erroneously from equation} \right]$$

f. Gametic genetic variance is the variance among the genotypic values associated with the array of gametes. Each gamete is considered in combination with all possible gametes in the pool, namely,

$$\sigma_g^2 = \sum_i \sum_k p_{A_i} p_{B_k} (G_{i.k.} - G_{\dots})^2$$

**Exercise 9.10.**

a.  $\text{Cov}(\text{H.S.}) = \frac{1}{4} \sigma_A^2 + \frac{1}{16} \sigma_{AA}^2 = \frac{1}{4}(0.36107904) + \frac{1}{16}(0.01161216) = 0.09099552$  [see equation (9.147)]

b.  $\text{Cov}(\text{H.S.}) = \frac{1}{4} \sigma_A^2 + \frac{1+\lambda_1^2}{16} \sigma_{AA}^2 = \frac{1}{4}(0.36107904) + \frac{1+(0.6)^2}{16}(0.01161216)$  [see equation (9.158)]  
 $= 0.09125679$

There is no change in the additive variance component with linkage. The change in the additive-by-additive variance component is an increase as follows:

$$\left( \frac{1+\lambda_1^2}{16} \right) / \left( \frac{1}{16} \right) = 1 + \lambda_1^2 = 1 + (0.6)^2 = 1.36 \quad \text{or 36\% increase}$$

The change in the half-sib covariance itself is

$$\frac{0.09125679}{0.09099552} = 1.00287128 \quad \text{or 0.287\% increase}$$

c.  $\text{Cov}(\text{F.S.}) = \frac{1}{2} \sigma_A^2 + \frac{1}{4} \sigma_D^2 + \frac{1}{4} \sigma_{AA}^2 + \frac{1}{8} \sigma_{AD}^2 + \frac{1}{16} \sigma_{DD}^2$  [see equation (9.148)]  
 $= \frac{1}{2}(0.36107904) + \frac{1}{4}(0.31526208) + \frac{1}{4}(0.01161216) + \frac{1}{8}(0.02225664) + \frac{1}{16}(0.01016064)$   
 $= 0.2656752$

d.

$$\text{Cov}(\text{F.S.}) = \frac{1}{2} \sigma_A^2 + \frac{1}{4} \sigma_D^2 + \frac{1}{4} \left( 1 + \frac{1}{2} \lambda^2 \right) \sigma_{AA}^2 + \frac{1}{8} (1 + \lambda^2) \sigma_{AD}^2 + \frac{1}{16} (1 + \lambda^2)^2 \sigma_{DD}^2$$
 [see equation (9.171)]  
 $= \frac{1}{2}(0.36107904) + \frac{1}{4}(0.31526208)$   
 $+ \frac{1}{4} \left( 1 + \frac{1}{2} (0.6)^2 \right) (0.01161216) + \frac{1}{8} (1 + (0.6)^2) (0.02225664) + \frac{1}{16} (1 + (0.6)^2)^2 (0.01016064)$   
 $= 0.267738826$

$$\text{Change in } \sigma_{AA}^2 = \frac{1 + \frac{1}{2} \lambda^2}{1} = 1 + \frac{1}{2} (0.6)^2 = 1.18 \quad \text{or 18\% increase [see equations (9.171) and (9.171A)]}$$

$$\text{Change in } \sigma_{AD}^2 = \frac{1 + \lambda^2}{1} = 1 + (0.6)^2 = 1.36 \quad \text{or 36\% increase}$$

$$\text{Change in } \sigma_{DD}^2 = \frac{(1 + \lambda^2)^2}{1} = \left[ 1 + (0.6)^2 \right]^2 = 1.8496 \quad \text{or 85\% increase}$$

$$\text{Change in full-sib covariance} = \frac{0.26773882}{0.2656752} = 1.007767477 = 0.7767\% \text{ increase}$$

e.  $\text{Cov}(P.,O.) = \frac{1}{2}\sigma_A^2 + \frac{1}{4}\sigma_{AA}^2 = \frac{1}{2}(0.36107904) + \frac{1}{4}(0.01161216) = 0.18344256$  [see equation (9.145)]

f. Linkage has no effect on the parent-offspring covariance. (See Section 9.5.5, pp. 9.90 and 9.91.)

g.  $\text{Cov}(GP.,GO.) = \frac{1}{4}\sigma_A^2 + \frac{1}{16}\sigma_{AA}^2 = \frac{1}{4}(0.36107904) + \frac{1}{16}(0.01161216) = 0.09099552$

[see equation (9.146)]

h.  $\text{Cov}(GP.,GO.) = \frac{1}{4}\sigma_A^2 + \frac{1}{16}(1+\lambda)\sigma_{AA}^2 = \frac{1}{4}(0.36107904) + \frac{1}{16}(1+0.6)(0.01161216) = 0.091430976$

[see equations (9.205) and (9.206)]

There is no change in the additive variance component. The change in the additive-by-additive variance component is:

$$\text{Change in } \sigma_{AA}^2 = 1 + \lambda = 1 + (0.6) = 1.6 \text{ or } 60\% \text{ increase}$$

$$\text{Change in grandparent-grandoffspring covariance} = \frac{0.091430976}{0.09099552} = 1.004785466 = 0.7767\% \text{ increase}$$

### Exercise 9.11.

Assumptions are listed on pp. 9.68 and 9.69 (also see pp. 10.18 and 10.19)

(1) Parents unrelated ( $\theta = 0$ )

Reason: So that all offspring are noninbred, random members (also implies independence between different parents within each parental set) of a random-mating population. This relates to (6) below.

(2) Parents are random members either from a random-mating population (if the parents are noninbred or from random inbred derivatives from a random-mating population.

Reason: So that an unbiased estimate of the variance component or the covariance between relatives is obtained. This relates to (7) below.

(3) Disomic inheritance.

Reason: If tetrasomic or higher level of inheritance existed, one would have a different genic factorial model with additional kinds of genetic variances such as digenic, trigenic, etc.

(4) No environmental covariance.

Reason: If nonzero environmental covariance exists between relatives, it will increase or decrease the covariance between relatives (or variance components of mating design). See equation (10.9)

(5) No maternal effects.

Reason: If a maternal effect exists, it is nongenetic and will increase the covariances or variance components of the mating design.

(6) Relatives are noninbred.

Reason: So that all offspring or relatives are members of a noninbred population. This is a consequence of (1) above and a prerequisite for (7). See equation (9.52) and top of p. 9.33 on reduced number of states of IBD for four gene functions.

(7) Relatives are random members of a population in Hardy-Weinberg equilibrium.

Reason: So that the frequency of both X and Y may be expressed as a product of the allelic frequencies with respect to a single locus. This has many consequences as illustrated in (9.40), for example. Also the free subscript idea is also involved. This is a consequence of (2) above.

(8) Linkage equilibrium.

Reason: This is extending the nicety of random mating with respect to a single locus to two or more loci. It permits writing equation (9.115) with the frequency of each  $(\alpha\alpha)^{mm}$  as  $p_{a_i}p_{b_k}$ , for example, and even more important, gives so-called “free subscripts” [see equation (9.116) and p. 9.58]

(9) No linkage.

Reason: This permits writing such probability statements as  $P(a_X^m \equiv a_Y^m; b_X^m \equiv b_Y^m) = P(a_X^m \equiv a_Y^m)P(b_X^m \equiv b_Y^m)$ .

See equation (9.119). Also see pp. 9.70 to 9.95.