

SOLUTIONS TO EXERCISES FOR CHAPTER 3

**Exercise 3.1.**

a. Random mating means that the parents or mates are genetically or genotypically independent of one another, i.e.,  $p(G_i^m \cup G_j^f) = p(G_i^m)p(G_j^f)$ .

b. In a single population composed of unisexual individuals, the frequencies of the genotypes are usually the same in the two sexes. Assuming that the frequencies are the same, the frequency of any homozygote is  $p(g_i)^2$  and the frequency of any heterozygote is  $2p(g_i)p(g_j)$ . This implies that for any single locus, the genotypic frequencies are Hardy-Weinberg frequencies. If the frequencies are different in the two sexes, then the frequency of any homozygote is  $p(g_i^m)p(g_i^f)$  and the frequency of any heterozygote is

$$p(g_i^m)p(g_j^f) + p(g_j^m)p(g_i^f).$$

On the other hand, for a single population composed of bisexual individuals, it is with certainty that the frequencies of the genotypes are the same in the sexes because every individual functions equally as both a male and a female. Hence, the frequency of any homozygote is  $p(g_i)^2$  and the frequency of any heterozygote is  $2p(g_i)p(g_j)$ .

c. In a cross between two populations the frequencies of the genotypes are usually different. Hence, the frequency of any homozygote in the cross is  $p(g_i^m)p(g_i^f)$  and the frequency of any heterozygote is

$p(g_i^m)p(g_j^f) + p(g_j^m)p(g_i^f)$ . This implies that for any single locus, the genotypic frequencies are not in H-W proportions.

d. If an individual is defined as one composed only of the loci represented in  $G$ , then the definition of random mating that any individual of one sex has an equal chance of mating with any other individual of the opposite sex is equivalent. For bisexual individuals this includes self-mating. On the other hand, if an individual is defined as one composed of more loci than are included in  $G$  than it is not proper to define random mating with respect to  $G$  by saying that every individual has an equal chance of mating with any other individual of the opposite sex. This is true because the individuals with any given  $G_i$  may not be mating at random, i.e., there may be positive phenotypic or genetic assortative mating with respect to loci not included in  $G$ . For example, if  $G$  denotes a single loci for seed color or blood type as described in my notes (see p. 3.8), random mating may occur with respect to the genotypes at this locus but yet with respect to the individuals (composed of loci for other traits or all remaining ones in the organism) there may be positive phenotypic assortative mating such as flowering time or intelligence.

**Exercise 3.2.**

The random mating theorem is important in that one can obtain the genotypic structure of the next generation under random mating by using only the gametic frequencies of the previous generation. Thus, one can avoid the laborious procedure of randomly pairing the parents, considering their individual gametics outputs, determining the genotypes of their offspring and their frequencies, and then tallying over all possible genotypic matings.

**Exercise 3.3.**

The allelic frequencies are [see equation (2.10)]

$$p_A = 0.25 + \frac{1}{2}(0.10) = 0.25 + 0.05 = 0.3$$

$$p_a = \frac{1}{2}(0.10) + 0.65 = 0.05 + 0.65 = 0.7$$

The genotypic frequencies after random mating are [equation (3.35)]

$$p_{AA} = (0.3)^2 = 0.09$$

$$p_{Aa} = 2(0.3)(0.7) = 0.42$$

$$p_{aa} = (0.7)^2 = 0.49$$

**Exercise 3.4.**

Population 1: Allelic and Hardy-Weinberg genotypic frequencies are

$$p_A = 0.25 + \frac{1}{2}(0.10) = 0.25 + 0.05 = 0.3$$

$$p_a = 0.65 + \frac{1}{2}(0.10) = 0.65 + 0.05 = 0.7$$

$$p_{AA} = p_A^2 = (0.3)^2 = 0.09; \quad 2p_{Aa} = 2(0.3)(0.7) = 0.42; \quad p_{aa} = p_a^2 = 0.49$$

Population 2: Allelic and Hardy-Weinberg genotypic frequencies are

$$p_A = 0.30 + 0 = 0.3$$

$$p_a = 0.70 + 0 = 0.7$$

$$p_{AA} = p_A^2 = (0.3)^2 = 0.09; \quad 2p_{Aa} = 2(0.3)(0.7) = 0.42; \quad p_{aa} = p_a^2 = 0.49$$

Population 3: Allelic and Hardy-Weinberg genotypic frequencies are

$$p_A = 0.00 + \frac{1}{2}(0.60) = 0.3$$

$$p_a = 0.40 + \frac{1}{2}(0.60) = 0.7$$

$$p_{AA} = p_A^2 = (0.3)^2 = 0.09; \quad 2p_{Aa} = 2(0.3)(0.7) = 0.42; \quad p_{aa} = p_a^2 = 0.49$$

**Exercise 3.5.**

A random-mating population showing 5% recessives has the following percentage of heterozygotes:

$$q^2 = 0.05, \text{ so } q = \sqrt{0.05} \text{ and } p = 1 - \sqrt{0.05}$$

$$2(1 - \sqrt{0.05})(\sqrt{0.05}) = 2(0.7764)(0.2236) = 0.3472 = \text{frequency of heterozygotes}$$

or 34.72% heterozygotes

**Exercise 3.6.**

Population 1:

$$p_{A-} = 0.02 = p^2 + 2pq; \quad p_{aa} = 0.98 = q^2; \quad q = \sqrt{0.98}$$

$$q = 0.9899 \quad p = 1 - q = 1 - 0.9899 = 0.0101$$

$$2pq = 2(0.0101)(0.9899) = 0.01999598$$

Population 2:

$$p_{aa} = 0.02 \quad q = \sqrt{0.02} = 0.1414 \quad p = 1 - 0.1414 = 0.8586$$

$$2pq = 2(0.8586)(0.1414) = 0.2428$$

**Exercise 3.7.**

$$\lim_{p \rightarrow 0} \frac{2p(1-p)}{p} = \lim_{p \rightarrow 0} 2(1-p) = 2,$$

namely, frequency of heterozygotes  $\cong 2p$  when  $p$  is very small.

An alternative way to solve this problem is to apply L'Hospital's Rule (given in any calculus book), which says that if both the numerator and denominator vanish as the variable  $p$  approaches the limit 0, so that the fraction assumes the indeterminate form  $0/0$ , then the limit of the original fraction is equal to the limit of the ratio of the derivatives of the numerator and denominator, namely,

$$\lim_{p \rightarrow 0} \frac{2p(1-p)}{p} = \lim_{p \rightarrow 0} \frac{d[2p(1-p)]/dp}{dp/dp} = \lim_{p \rightarrow 0} \frac{2p(-1) + (1-p)2}{1} = \lim_{p \rightarrow 0} \frac{-2p + 2(1-p)}{1} = \frac{-2(0) + 2(1-0)}{1} = 2$$

where we apply the derivative of a product to the numerator, namely,  $\frac{d(uv)}{dx} = u \frac{dv}{dx} + v \frac{du}{dx}$

**Exercise 3.8.**

First, assuming random mating, we solve this exercise by taking a more brute force, definitional approach as follows: We list only those matings (7 out of 9 possible matings) that give heterozygotes.

Male parent	Female parent	Mating frequency	Conditional probability			A <sub>1</sub> A <sub>2</sub> offspring from A <sub>1</sub> A <sub>2</sub> female	A <sub>1</sub> A <sub>2</sub> offspring from A <sub>1</sub> A <sub>1</sub> female	A <sub>1</sub> A <sub>2</sub> offspring from A <sub>2</sub> A <sub>2</sub> female
			A <sub>1</sub> A <sub>1</sub>	A <sub>1</sub> A <sub>2</sub>	A <sub>2</sub> A <sub>2</sub>			
A <sub>1</sub> A <sub>1</sub>	A <sub>1</sub> A <sub>2</sub>	p <sup>2</sup> 2pq	1/2	1/2		p <sup>3</sup> q		
A <sub>1</sub> A <sub>1</sub>	A <sub>2</sub> A <sub>2</sub>	p <sup>2</sup> q <sup>2</sup>		1				p <sup>2</sup> q <sup>2</sup>
A <sub>1</sub> A <sub>2</sub>	A <sub>1</sub> A <sub>1</sub>	2ppq <sup>2</sup>	1/2	1/2			p <sup>3</sup> q	
A <sub>1</sub> A <sub>2</sub>	A <sub>1</sub> A <sub>2</sub>	(2pq) <sup>2</sup>	1/4	1/2	1/4	2p <sup>2</sup> q <sup>2</sup>		
A <sub>1</sub> A <sub>2</sub>	A <sub>2</sub> A <sub>2</sub>	2ppq <sup>2</sup>		1/2	1/2			pq <sup>3</sup>
A <sub>2</sub> A <sub>2</sub>	A <sub>1</sub> A <sub>1</sub>	q <sup>2</sup> p <sup>2</sup>		1			p <sup>2</sup> q <sup>2</sup>	
A <sub>2</sub> A <sub>2</sub>	A <sub>1</sub> A <sub>2</sub>	q <sup>2</sup> 2pq		1/2	1/2	pq <sup>3</sup>		
						pq	p <sup>2</sup> q	pq <sup>2</sup>

Summing the column totals, we have

$pq + p^2q + pq^2 = 2pq =$  total frequency of heterozygotes in the random-mating population, as it should be.

Hence,  $\frac{\text{frequency of heterozygous offspring from heterozygous female parent}}{\text{frequency of heterozygous female parent}} = \frac{pq}{2pq} = \frac{1}{2}$

An alternative, more direct way to solve this exercise is the following:

Female parent	Random mating frequency	Nonrandom mating frequency	<u>Male allele</u>					
			$A_1A_1$	<u>Offspring</u> $A_1A_2$	$A_2A_2$	$A_1A_1$	<u>Offspring</u> $A_1A_2$	$A_2A_2$
$A_1A_1$	$p^2$	$P$	1				1	
$A_1A_2$	$2pq$	$2Q$	$\frac{1}{2}$	$\frac{1}{2}$			$\frac{1}{2}$	$\frac{1}{2}$
$A_2A_2$	$q^2$	$Q$		1				1

From the above it is easily seen that regardless of whether the male allele is  $A_1$  or  $A_2$  one-half of the offspring from the female parental genotype  $A_1A_2$  is  $A_1A_2$ .

Finally we present even a third nontabulating, purely probability way to solve the exercise. In set notation, let  $A$  = female parent  $A_1A_2$ ,

$B_1$  = offspring  $A_1A_1$ ,

$B_2$  = offspring  $A_1A_2$ ,

$B_3$  = offspring  $A_2A_2$ .

$$P(\text{female parent is } A_1A_2 | \text{offspring is } A_1A_2) = \frac{P(\text{female parent } A_1A_2 \text{ and offspring } A_1A_2)}{P(\text{offspring is } A_1A_2)}$$

$$P(A | B_2) = \frac{P(A \cap B_2)}{P(B_2)} = \frac{P(A)P(B_2 | A)}{P(B_2)} = \frac{2pq\left(\frac{1}{2}\right)}{2pq} = \frac{1}{2}$$

Or, for nonrandom-mating population

$$P(A | B_2) = \frac{2Q\left(\frac{1}{2}\right)}{2Q} = \frac{1}{2}$$

### Exercise 3.9.

a. The frequencies of the three alleles in the population are

	$A_1$	$A_2$	$A_3$
$\frac{1}{6} A_1A_2$	$\frac{1}{12}$	$\frac{1}{12}$	
$\frac{1}{3} A_1A_3$	$\frac{1}{6}$		$\frac{1}{6}$
$\frac{1}{2} A_2A_3$		$\frac{1}{4}$	$\frac{1}{4}$
Total	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{5}{12}$

Or, alternatively for  $A_1$  only:

$$\begin{aligned} P(A_1) &= P(A_1A_2)P(A_1 | A_1A_2) + P(A_1A_3)P(A_1 | A_1A_3) \\ &= \frac{1}{6} \frac{1}{2} + \frac{1}{3} \frac{1}{2} = \frac{3}{12} = \frac{1}{4} \end{aligned}$$

b. After random mating the genotypic frequencies from actual tallying is

$P(G_i \cdot G_j)$	$A_1A_1$	$A_1A_2$	$A_1A_3$	$A_2A_2$	$A_2A_3$	$A_3A_3$
$(A_1A_2)(A_1A_2) \left(\frac{1}{6}\right)^2 = \frac{1}{36}$	$\frac{1}{4}$	$\frac{2}{4}$		$\frac{1}{4}$		
$(A_1A_3)(A_1A_3) \left(\frac{1}{3}\right)^2 = \frac{1}{9} = \frac{4}{36}$	$\frac{1}{4}$		$\frac{2}{4}$			$\frac{1}{4}$
$(A_2A_3)(A_2A_3) \left(\frac{1}{2}\right)^2 = \frac{1}{4} = \frac{9}{36}$				$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$
$(A_1A_2)(A_1A_3) 2 \cdot \frac{1}{6} \frac{1}{3} = \frac{2}{18} = \frac{4}{36}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$		$\frac{1}{4}$	
$(A_1A_2)(A_2A_3) 2 \cdot \frac{1}{6} \frac{1}{2} = \frac{2}{12} = \frac{6}{36}$		$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	
$(A_1A_3)(A_2A_3) 2 \cdot \frac{1}{3} \frac{1}{2} = \frac{2}{6} = \frac{12}{36}$		$\frac{1}{4}$	$\frac{1}{4}$		$\frac{1}{4}$	$\frac{1}{4}$

Multiplying the mating frequencies by the conditional probability for each offspring genotype gives the following frequencies of the offspring genotypes:

$$P(A_1A_1) = \frac{1+4+4}{144} = \frac{9}{144} = \frac{1}{16}; \quad P(A_1A_2) = \frac{2+4+6+12}{144} = \frac{24}{144} = \frac{1}{6}; \quad P(A_1A_3) = \frac{8+4+6+12}{144} = \frac{30}{144} = \frac{5}{24};$$

$$P(A_2A_2) = \frac{1+9+6}{144} = \frac{16}{144} = \frac{1}{9}; \quad P(A_2A_3) = \frac{18+4+6+12}{144} = \frac{40}{144} = \frac{5}{18}; \quad P(A_3A_3) = \frac{4+9+12}{144} = \frac{25}{144};$$

c. Yes, the offspring genotypic frequencies are in Hardy-Weinberg equilibrium. If the male and female gametic arrays are equal, Hardy-Weinberg proportions are obtained immediately in the first generation after random mating.

$$P(A_1A_1) = \frac{1}{4} \frac{1}{4} = \frac{1}{16}; \quad P(A_1A_2) = 2 \cdot \frac{1}{4} \frac{1}{3} = \frac{1}{6}; \quad P(A_1A_3) = 2 \cdot \frac{1}{4} \frac{5}{12} = \frac{5}{24};$$

$$P(A_2A_2) = \frac{1}{3} \frac{1}{3} = \frac{1}{9}; \quad P(A_2A_3) = 2 \cdot \frac{1}{3} \frac{5}{12} = \frac{5}{18}; \quad P(A_3A_3) = \frac{5}{12} \frac{5}{12} = \frac{25}{144};$$

### Exercise 3.10.

Three-fourths of the AO individuals would be expected to be  $A_1O$ , and one-fourths of the AO individuals would be expected to be  $A_2O$ . The expected proportions of  $A_1A_1$ ,  $A_1A_2$ , and  $A_2A_2$  among AA individuals would be

$A_1A_1$	$A_1A_2$	$A_2A_2$
$9/16$	$6/16$	$1/16$

### Exercise 3.11.

The maximum proportions of heterozygotes in a random-mating population are  $\frac{1}{2}$ ,  $\frac{2}{3}$ , and  $\frac{m-1}{m}$  for two, three, and  $m$  alleles, respectively. See equation (3.59).

### Exercise 3.12.

Let  $p_A$  be the frequency of the recessive allele.

$$8 = \frac{2p_A(1-p_A)}{p_A^2}$$

$$8p_A = 2(1-p_A)$$

$$= 2 - 2p_A$$

$$10p_A = 2$$

$$p_A = \frac{1}{5} = 0.20 \quad \text{--see equation (3.35)}$$

$$98 = \frac{2(1-p_A)}{p_A}$$

$$98p_A = 2(1-p_A)$$

$$= 2 - 2p_A$$

$$100p_A = 2$$

$$p_A = \frac{2}{100} = \frac{1}{50} = 0.02$$

**Exercise 3.13.**

$$\hat{p}_A = \frac{2(13,655) + 678}{28,690} = \frac{27,988}{28690} = 0.97553154$$

$$\hat{p}_a = 1 - 0.97553154 = 0.02446846$$

	Expected number (E)	O - E	$\chi^2$
AA	$\hat{p}^2(14,345) = 13,651.588$	-3.412	0.00085277
Aa	$2\hat{p}\hat{q}(14,345) = 684.823$	+6.823	0.06797862
aa	$\hat{q}^2(14,345) = 8.588$	-3.412	1.35558267

$$1.42440 < \chi_{\nu=1, \alpha=0.05}^2 = 3.84, \text{ nonsignificant}$$

Degrees of freedom equal the number of cells minus 1 minus the number of independent parameters which are estimated from the data and are used to calculate the expected numbers. In this case, the degrees of freedom equal one.

**Exercise 3.14.**

Two generations of random mating are required to reach Hardy-Weinberg proportions when the allelic frequencies at one or more loci are different in the two sexes. In first generation of random mating, the genotypic frequencies are not in H-W proportions, but the allelic frequencies are equalized between the two sexes due to independent assortment of the sex chromosomes in gamete formation in every individual or for each allele. There are equal frequencies of males and females for every genotype. Then in the second generation of random mating with equal allelic frequencies in both males and females, H-W proportions are established.

**Exercise 3.15.**

The allelic frequencies in the two populations are

population 1:	$p_A = 0.25 + \frac{0.50}{2} = 0.5$	population 2:	$p_A = 0.16 + \frac{0.48}{2} = 0.4$
	$p_a = \frac{0.50}{2} + 0.25 = 0.5$		$p_a = \frac{0.48}{2} + 0.36 = 0.6$

In the first generation of the cross:

		Females		
		0.4 A	0.6 a	
Males	0.5 A	0.20 AA	0.30 Aa	0.20 AA + 0.50 Aa + 0.30 aa
	0.5 a	0.20 Aa	0.30 aa	

In the second generation of the cross after random mating:

	0.45 A	0.55 a	
0.45 A	0.2025 AA	0.2475 Aa	0.2025 AA + 0.4950 Aa + 0.3025 aa
0.55 a	0.2475 Aa	0.3025 aa	

Equilibrium proportions would be achieved after two generations of random mating or one generation after the cross itself.

**Exercise 3.16.**

- a. The allelic frequencies in each of the two populations are:

Population X

$$p_1^X = p_{11}^X + 2p_{12}^X\left(\frac{1}{2}\right) = p_{11}^X + p_{12}^X$$

$$p_2^X = p_{22}^X + 2p_{12}^X\left(\frac{1}{2}\right) = p_{22}^X + p_{12}^X$$

Population Y

$$p_1^Y = p_{11}^Y + 2p_{12}^Y\left(\frac{1}{2}\right) = p_{11}^Y + p_{12}^Y$$

$$p_2^Y = p_{22}^Y + 2p_{12}^Y\left(\frac{1}{2}\right) = p_{22}^Y + p_{12}^Y$$

- b. The genotypic structure in the cross population is

$$\begin{array}{ccc} A_1A_1 & A_1A_2 & A_2A_2 \\ \left(p_1^X\right)\left(p_1^Y\right) & \left(p_1^X\right)\left(p_2^Y\right) + \left(p_2^X\right)\left(p_1^Y\right) & \left(p_2^X\right)\left(p_2^Y\right) \end{array}$$

The allelic frequencies in the cross population are [equation (3.100)]

$$p_1^c = \frac{p_1^X + p_1^Y}{2}$$

$$p_2^c = \frac{p_2^X + p_2^Y}{2}$$

- c. If individuals of the cross generation mate at random, Hardy-Weinberg equilibrium will be obtained in the next generation, because with equal gametic frequencies on the male and female sides and random mating Hardy-Weinberg proportions are obtained. The genotypic structure of the cross generation after random mating is

$$\begin{array}{ccc} A_1A_1 & A_1A_2 & A_2A_2 \\ \left(p_1^c\right)^2 & 2\left(p_1^c\right)\left(p_2^c\right) & \left(p_2^c\right)^2 \\ = \left(\frac{p_1^X + p_1^Y}{2}\right)^2 & = 2\left(\frac{p_1^X + p_1^Y}{2}\right)\left(\frac{p_2^X + p_2^Y}{2}\right) & = \left(\frac{p_2^X + p_2^Y}{2}\right)^2 \end{array}$$

- d. The numerical values for the allelic frequencies in each of the two populations are

Population X

$$p_1^X = 0.20 + 0.25 = 0.45$$

$$p_2^X = 0.30 + 0.25 = 0.55$$

Population Y

$$p_1^Y = 0.40 + 0.20 = 0.60$$

$$p_2^Y = 0.20 + 0.20 = 0.40$$

The genotypic structure for the cross population is

		Population Y			
		$A_1$	$A_2$		
		0.60	0.40		
Population X	$A_1$ 0.45	$A_1A_1$ $(0.45)(0.60)$ $= 0.27$	$A_1A_2$ $(0.45)(0.40)$ $= 0.18$	0.45	
	$A_2$ 0.55	$A_2A_1$ $(0.55)(0.60)$ $= 0.33$	$A_2A_2$ $(0.55)(0.40)$ $= 0.22$	0.55	
		0.60	0.40		

The allelic frequencies in the cross population are

$$p_1^c = \frac{0.45 + 0.60}{2} = 0.525$$

$$p_2^c = \frac{0.55 + 0.40}{2} = 0.475$$

The genotypic structure of the cross population after random mating is

$A_1A_1$	$A_1A_2$	$A_2A_2$
$(p_1^c)^2$	$2(p_1^c)(p_2^c)$	$(p_2^c)^2$
$= (0.525)^2 = 0.275625$	$= 2(0.525)(0.475) = 0.49875$	$= (0.475)^2 = 0.225625$

### Exercise 3.17.

No, a stable allelic structure at every locus does not imply a stable multilocus gametic structure or a stable genotypic structure from generation to generation. If the population is not in linkage equilibrium, the gametic structure will change from generation to generation.

### Exercise 3.18.

a. Let  $p_i$  = frequency of  $i$ th allele at  $A$  locus, and  $r_j$  = frequency of  $j$ th allele at  $B$  locus. With linkage equilibrium the frequency of  $A_iB_j$  gamete is  $p_i r_j$ . Using the random-mating theorem, we know that the frequency of any genotype is the product of the gametic frequency, i.e.,

$$p(A_i A_{i'} B_j B_{j'}) = (p_i r_j)(p_{i'} r_{j'}) = p_i p_{i'} r_j r_{j'}$$

Thus, for any fixed combination  $i$  and  $i'$ , we sum over all combinations of  $j$  and  $j'$

$$\begin{aligned}
\sum_{j=1}^{m_B} \sum_{j'=1}^{m_B} p(A_i A_{i'} B_j B_{j'}) &= \sum_{j=1}^{m_B} \sum_{j'=1}^{m_B} p_i p_{i'} r_j r_{j'} \\
&= p_i p_{i'} \sum_{j=1}^{m_B} \sum_{j'=1}^{m_B} r_j r_{j'} \\
&= p_i p_{i'} (1) \\
&= p_i p_{i'} \quad \text{for every } i \text{ and } i'
\end{aligned}$$

This proves that H-W proportions exist for the  $A$  locus in that every genotype at the  $A$  locus is the product of the corresponding allelic frequencies. In like manner we can prove that H-W proportions exist for the  $B$  locus.

We can, of course, double sum with respect to  $i$  and  $i'$ , in addition to the double sum with respect to  $j$  and  $j'$ ,

$$= \sum_{i=1}^{m_A} \sum_{i'=1}^{m_A} p_i p_{i'} \sum_{j=1}^{m_B} \sum_{j'=1}^{m_B} r_j r_{j'}$$

and, in like manner to that above, establish H-W proportions for both the  $A$  and  $B$  loci.

b. Using the random-mating theorem, we know that the frequency of any genotype is the product of the gametic frequencies, i.e.,

$$p(A_i A_{i'} B_j B_{j'})_t = (p_i r_j + \Delta_{ijt-1})(p_{i'} r_{j'} + \Delta_{i'j't-1})$$

Furthermore to obtain the frequency of any genotype, say  $A_i A_{i'}$ , we must sum over all genotypes at the  $B$  locus

$$\begin{aligned}
\sum_{j=1}^{m_B} \sum_{j'=1}^{m_B} p(A_i A_{i'} B_j B_{j'})_t &= \sum_{j=1}^{m_B} \sum_{j'=1}^{m_B} (p_i r_j + \Delta_{ijt-1})(p_{i'} r_{j'} + \Delta_{i'j't-1}) \\
&= \sum_{j=1}^{m_B} (p_i r_j + \Delta_{ijt-1}) \sum_{j'=1}^{m_B} (p_{i'} r_{j'} + \Delta_{i'j't-1}) \\
&= \left( p_i \sum_{j=1}^{m_B} r_j + \sum_{j=1}^{m_B} \Delta_{ijt-1} \right) \left( p_{i'} \sum_{j'=1}^{m_B} r_{j'} + \sum_{j'=1}^{m_B} \Delta_{i'j't-1} \right) \\
&= [p_i (1) + 0][p_{i'} (1) + 0] \quad \text{[see equations (3.62A) and (3.62B)]} \\
&= p_i p_{i'}
\end{aligned}$$

which proves that H-W proportions or random mating exists at the  $A$  locus. Likewise we can show that H-W proportions or random mating exists at the  $B$  locus.

### Exercise 3.19.

Yes, with multiple alleles linkage equilibrium may exist for some combination of  $i$  and  $j$  in that  $p_{A_i B_j} = p_{A_i} p_{B_j}$ , but for other combinations of  $i$  and  $j$ ,  $p_{A_i B_j} \neq p_{A_i} p_{B_j}$ . All that is required for the existence of linkage disequilibrium in the population as a whole is to have  $p_{A_i B_j} \neq p_{A_i} p_{B_j}$  for at least one combination of  $i$  and  $j$ .

### Exercise 3.20.

a. From equation (3.63) the frequency of the  $AB$  gamete is

$$\begin{aligned}
p_{AB} &= p_A p_B + \Delta_{AB} \\
&= (0.4)(0.7) + 0.07 \\
&= 0.35
\end{aligned}$$

The frequencies of all gametes are

	B	b	
A	$p_{AB} = 0.35$	$p_{Ab} = 0.05$	0.4
a	$p_{aB} = 0.35$	$p_{ab} = 0.25$	0.6
	0.7	0.3	1.0

The frequencies of the following genotypes in the population are

$$p\left(\frac{AB}{AB}\right) = (0.35)(0.35) = 0.1225$$

$$2p\left(\frac{AB}{ab}\right) = 2(0.35)(0.25) = 0.1750$$

$$2p\left(\frac{Ab}{aB}\right) = 2(0.05)(0.35) = 0.0350$$

b. When the loci are linked ( $\rho_1 = \frac{1}{4}$ ), the frequency of the gametes from the population are obtained

directly as follows:

Genotype	Frequency	AB	Ab	aB	ab
AB/AB	$(0.35)(0.35) = 0.1225$	1			
AB/Ab	$2(0.35)(0.05) = 0.0350$	0.5	0.5		
AB/aB	$2(0.35)(0.35) = 0.2450$	0.5		0.5	
AB/ab	$2(0.35)(0.25) = 0.1750$	0.375	0.125	0.125	0.375
Ab/Ab	$(0.05)(0.05) = 0.0025$		1		
Ab/aB	$2(0.05)(0.35) = 0.0350$	0.125	0.375	0.375	0.125
Ab/ab	$2(0.05)(0.25) = 0.0250$		0.5		0.5
aB/aB	$(0.35)(0.35) = 0.1225$			1	
aB/ab	$2(0.35)(0.25) = 0.1750$			0.5	0.5
ab/ab	$(0.25)(0.25) = 0.0625$				1
Total		0.3325	0.0675	0.3675	0.2325

By use of equation (3.82), we obtain each gametic frequency directly as a function of the gametic frequency of the previous generation:

$$\begin{aligned} P_{ABt} &= \rho_0 P_{ABt-1} + \rho_1 P_A P_B \\ &= 0.75(0.35) + 0.25(0.4)(0.7) \\ &= 0.3325 \end{aligned}$$

$$\begin{aligned} P_{Abt} &= \rho_0 P_{Abt-1} + \rho_1 P_A P_b \\ &= 0.75(0.05) + 0.25(0.4)(0.3) \\ &= 0.0675 \end{aligned}$$

$$\begin{aligned} P_{aBt} &= \rho_0 P_{aBt-1} + \rho_1 P_a P_B \\ &= 0.75(0.35) + 0.25(0.6)(0.7) \\ &= 0.3675 \end{aligned}$$

$$\begin{aligned} P_{abt} &= \rho_0 P_{abt-1} + \rho_1 P_a P_b \\ &= 0.75(0.25) + 0.25(0.6)(0.3) \\ &= 0.2325 \end{aligned}$$

c..After random mating the linkage disequilibrium for gamete  $AB$  by definition is [see equation (3.61)]

$$\begin{aligned}\Delta_{ABt} &= p_{ABt} - p_A p_B \\ &= 0.3325 - (0.4)(0.7) \\ &= 0.0525\end{aligned}$$

The linkage disequilibrium calculated from that of the previous generation by equation (3.85) is

$$\begin{aligned}\Delta_{ABt} &= \rho_0 \Delta_{ABt-1} \\ &= 0.75(0.07) \\ &= 0.0525\end{aligned}$$

### Exercise 3.21.

a.

Genotype	Frequency	Conditional probability			
		AB	Ab	aB	ab
$\frac{AB}{AB}$	$\frac{1}{4}$	1			
$\frac{AB}{ab}$	$\frac{1}{2}$	$\frac{2}{5}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{5}$
$\frac{ab}{ab}$	$\frac{1}{4}$				1
Total		$\frac{9}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{9}{20}$

$$b. f\left(\frac{AB}{aB}\right) = 2\left(\frac{9}{20}\right)\left(\frac{1}{20}\right) = \left(\frac{9}{10}\right)\left(\frac{1}{20}\right) = \frac{9}{200} = 0.045$$

$$c. p_{AB} = p_{Ab} = p_{aB} = p_{ab} = 0.5(0.5) = 0.25$$

$$d. f\left(\frac{AB}{aB}\right) = 2\left(\frac{1}{4}\right)\left(\frac{1}{4}\right) = \frac{1}{8}$$

$$e. \Delta_{AB0} = \frac{9}{20} - \frac{1}{2}\left(\frac{1}{2}\right) = 0.45 - 0.25 = 0.20 \quad [\text{see equation (3.94)}]$$

$$\begin{aligned}\Delta_{ABt} &= 0.10 = \rho_0^t \Delta_{AB0} \\ &= (0.8)^t 0.20\end{aligned}$$

$$\frac{0.10}{0.20} = (0.8)^t$$

$$t \log 0.8 = \log(0.5)$$

$$t(1 - 0.90309) = 1 - 0.69897$$

$$t(0.09691) = 0.30103$$

$t = 3.10628$  generations required for the linkage disequilibrium to decrease halfway to equilibrium

**Exercise 3.22.**

The proof that  $\sum_{j=1}^{m_B} \Delta_{ij} = 0$  and  $\sum_{i=1}^{m_A} \Delta_{ij} = 0$  for  $m_A$  and  $m_B$  greater than two is derived in equations (3.62A)

and (3.62B). For the first equality we have

$$\sum_{j=1}^{m_B} \Delta_{ij} = 0$$

$$\sum_{j=1}^{m_B} (p_{ij} - p_i p_j) = 0$$

$$\sum_j p_{ij} - p_i \sum_j p_j = 0$$

$$p_i - p_i(1) = 0$$

$$p_i - p_i = 0$$

$$0 = 0$$

Likewise for the second equality we have

$$\sum_{i=1}^{m_A} \Delta_{ij} = 0$$

$$\sum_{i=1}^{m_A} (p_{ij} - p_i p_j) = 0$$

$$\sum_i p_{ij} - p_j \sum_i p_i = 0$$

$$p_j - p_j(1) = 0$$

$$p_j - p_j = 0$$

$$0 = 0$$

**Exercise 3.23.**

a. The Hardy-Weinberg condition for an autosomal locus is that condition when every genotypic frequency at any single locus in a single population is equal to the product of the corresponding allelic frequencies.

It is achieved after one generation of random mating when the male and female gametic arrays are equal.

b. Linkage disequilibrium exists for any population when the frequency of any kind of gamete differs from the product of the frequencies of the corresponding genes. For example, linkage disequilibrium exists at two loci if the linkage disequilibrium value  $\Delta_{A_i B_j} = p_{A_i B_j} - p_{A_i} p_{B_j}$  differs from zero.

Linkage disequilibrium is brought about in the initial population or amongst the founder individuals from selection, or migration, or hybridization, or by chance in a small population, when the particular genes are associated more or less frequently than the product of the frequencies of the corresponding genes.

No, association of genes on the same chromosome is not essential for linkage disequilibrium; it simply depends upon how they are associated in the individuals in the initial population.

The only role of chromosomal association upon linkage disequilibrium is its influence upon the loss of linkage disequilibrium from generation to generation. The rate of loss is  $\rho_1$ .

Linkage disequilibrium affects the frequencies of all 10 possible two-locus genotypes from generation to generation until  $t \rightarrow \infty$ .

c. Hardy-Weinberg is a one-locus genotypic phenomenon whereas linkage equilibrium is a two or more locus gametic and genotypic phenomenon.

**Exercise 3.24.**

Probably the simplest approach is to construct a  $2 \times 2$  table, incorporating the restrictions that the sum of the disequilibrium values, weighting each cell equally in every row and column, equals zero and the further restriction that all gametic frequencies must be zero or greater (i.e., no negative gametic frequencies can exist). Focusing on the cell with the lowest equilibrium frequency (cell for  $A_1B_1$ ), the maximum absolute disequilibrium would occur for  $p_{A_1B_1} = 0.00$  or  $p_{A_1B_1} = 0.20$ . Hence, we present two  $2 \times 2$  tables: 1) when the gametic frequency  $p_{A_1B_1} = 0.00$ , and 2) when the gametic frequency  $p_{A_1B_1} = 0.20$ .

1) When the gametic frequency  $p_{A_1B_1} = 0.00$ :

	$B_1$	$B_2$	Total
$A_1$	$A_1B_1$ $p_1.p_1 = 0.06$ $p_{A_1B_1} = 0.00$ $\Delta_{A_1B_1} = -0.06$	$A_1B_2$ $p_1.p_2 = 0.24$ $p_{A_1B_2} = 0.30$ $\Delta_{A_1B_2} = 0.06$	$p_{1.} = 0.3$
$A_2$	$A_2B_1$ $p_2.p_1 = 0.14$ $p_{A_2B_1} = 0.20$ $\Delta_{A_2B_1} = 0.06$	$A_2B_2$ $p_2.p_2 = 0.56$ $p_{A_2B_2} = 0.50$ $\Delta_{A_2B_2} = -0.06$	$p_{2.} = 0.7$
Total	$p_{.1} = 0.2$	$p_{.2} = 0.8$	1

2) When the gametic frequency  $p_{A_1B_1} = 0.20$ :

	$B_1$	$B_2$	Total
$A_1$	$A_1B_1$ $p_1.p_1 = 0.06$ $p_{A_1B_1} = 0.20$ $\Delta_{A_1B_1} = 0.14$	$A_1B_2$ $p_1.p_2 = 0.24$ $p_{A_1B_2} = 0.10$ $\Delta_{A_1B_2} = -0.14$	$p_{1.} = 0.3$
	$A_2B_1$ $p_2.p_1 = 0.14$ $p_{A_2B_1} = 0.00$ $\Delta_{A_2B_1} = -0.14$	$A_2B_2$ $p_2.p_2 = 0.56$ $p_{A_2B_2} = 0.70$ $\Delta_{A_2B_2} = 0.14$	$p_{2.} = 0.7$
	Total $p_{.1} = 0.2$	$p_{.2} = 0.8$	1

The absolute linkage disequilibrium value is the greatest when  $p_{A_1B_1} = 0.20$ , namely,  $\Delta_{A_1B_1} = 0.14$ . The linkage disequilibrium values for the other gametes are:  $\Delta_{A_1B_2} = -0.14$ ,  $\Delta_{A_2B_1} = -0.14$ , and  $\Delta_{A_2B_2} = 0.14$ .

More formally, one can write  $\Delta_{A_1B_1} = p_{A_1B_1} - p_{A_1}p_{B_1}$  where  $p_{A_1B_1} \leq \min(p_{A_1} = 0.3, p_{B_1} = 0.2) \leq 0.20$ , so  $0.00 \leq p_{A_1B_1} \leq 0.20$ . When  $p_{A_1B_1} = 0.00$ ,  $\Delta_{A_1B_1} = -0.06$ ; and when  $p_{A_1B_1} = 0.20$ ,  $\Delta_{A_1B_1} = 0.14$ . Thus, the maximum absolute linkage disequilibrium exists when  $p_{A_1B_1} = 0.20$ . All of the gametic frequencies and linkage disequilibrium values for that case of  $p_{A_1B_1} = 0.20$  are given above.

### Exercise 3.25.

Way 1:  $p_{A_iB_j,t} = p_{A_i}p_{B_j} + \Delta_{A_iB_j,t} = p_{A_i}p_{B_j} + \rho_0^t \Delta_{A_iB_j,0}$  see equations (3.63) and (3.86)

Way 2:  $p_{A_iB_j,t} = p_{A_iB_j,0} - (1 - \rho_0^t) \Delta_{A_iB_j,0}$  see equation (3.95B)

In the first way, we add the linkage disequilibrium in generation  $t$  to the gametic frequency with independence of allelic frequencies. In the second way, we subtract the loss in linkage disequilibrium over  $t$  generations from the initial gametic frequency.

**Exercise 3.26.**

$$\Delta_{A_i B_j 0} = p_{A_i} p_{B_j} - p_{A_i} p_{B_j} \quad [\text{see equation (3.61)}]$$

$$= 0.28 - (0.3)(0.8)$$

$$= 0.28 - 0.24$$

$$= 0.04$$

$$\Delta_{A_i B_j t} = \rho_0^t \Delta_{A_i B_j 0}$$

$$\Delta_{A_i B_j 10} = (0.9)^{10} \Delta_{A_i B_j 0}$$

$$p_{A_i B_j 10} = p_{A_i} p_{B_j} + \Delta_{A_i B_j 10} \quad [\text{see equation (3.63)}]$$

$$= (0.3)(0.8) + (0.9)^{10}(0.04)$$

**Exercise 3.27.** [3.22]

The fraction of the linkage disequilibrium lost in each generation is  $\rho_1$  [see equation (3.91)], and that lost after  $t$  generations is  $1 - \rho_0^t$  [see equation (3.93)].

**Exercise 3.28.**

The loss in linkage disequilibrium after the first generation of random mating is  $\rho_1 \Delta_{A_i B_j 0}$ . The amount of loss in the second generation is  $\rho_1$  of that retained after the first generation, namely,  $\rho_1 (\rho_0 \Delta_{A_i B_j 0})$ . The amount of loss in the third generation is  $\rho_1$  of that retained after the second generation, namely,  $\rho_1 (\rho_0^2 \Delta_{A_i B_j 0})$ , and so on for the  $t$ th generation. Hence, the total loss is

$$\begin{aligned} \text{total loss} &= \rho_1 \Delta_0 + \rho_1 \rho_0 \Delta_0 + \rho_1 \rho_0^2 \Delta_0 + \cdots + \rho_1 \rho_0^{t-1} \Delta_0 \\ &= (\rho_1 + \rho_1 \rho_0 + \rho_1 \rho_0^2 + \cdots + \rho_1 \rho_0^{t-1}) \Delta_0 \end{aligned}$$

The quantity  $(\rho_1 + \rho_1 \rho_0 + \rho_1 \rho_0^2 + \cdots + \rho_1 \rho_0^{t-1})$  is the sum of a geometric series where the first term  $a$  is  $\rho_1$ , the

$$\text{common ratio } r \text{ is } \rho_0, \text{ and the number of terms } n \text{ is } t, \text{ so } S = \frac{\rho_1 - \rho_1 \rho_0^t}{1 - \rho_0} = \frac{\rho_1 (1 - \rho_0^t)}{\rho_1} = (1 - \rho_0^t).$$

A more direct solution is the following. We desire to prove that

$$1 - \rho_0^t = S = \frac{a - ar^n}{1 - r}$$

Starting with  $1 - \rho_0^t$ , we multiple both the numerator and denominator by  $(1 - \rho_0)$

$$1 - \rho_0^t = \frac{(1 - \rho_0)(1 - \rho_0^t)}{1 - \rho_0} = \frac{(1 - \rho_0) - (1 - \rho_0)\rho_0^t}{1 - \rho_0}$$

Letting  $1 - \rho_0 = a$ ,  $\rho_0 = r$ , and  $t = n$ , we have

$$\frac{(1 - \rho_0) - (1 - \rho_0)\rho_0^t}{1 - \rho_0} = \frac{a - ar^n}{1 - r} \text{ which is the sum of a geometric series.}$$

In this particular geometric series  $a$  and  $r$  are functionally related, i.e.,  $a = 1 - r$ .

**Exercise 3.29.**

I would expect inbreeding to slow the rate of approach to linkage equilibrium in a large inbred population. Inbreeding leads to an increased probability of union of like gametes. Conversely, inbreeding leads to a decrease in the probability of union of unlike gametes, and thus a decrease in the probability of recombination between unlike gametes. Since the recombination between unlike gametes is what leads to linkage equilibrium, inbreeding would reduce the rate of approach to linkage equilibrium.

**Exercise 3.30.**

a. See Section 3.4.5, pp. 3.74 to 3.77. The initial linkage disequilibria in males and females for linked loci are

$$\Delta_0^m = p_{A_i B_j}^m - p_{A_i} p_{B_j} = 0.28 - (0.3)(0.8) = 0.04$$

$$\Delta_0^f = p_{A_i B_j}^f - p_{A_i} p_{B_j} = 0.26 - (0.3)(0.8) = 0.02$$

The linkage disequilibria in males and females for linked loci in generation 1 are

$$\Delta_1^m = \rho_0^m \frac{1}{2} (\Delta_0^m + \Delta_0^f) = (1) \frac{1}{2} (0.04 + 0.02) = 0.03$$

$$\Delta_1^f = \rho_0^f \frac{1}{2} (\Delta_0^m + \Delta_0^f) = (0.9) \frac{1}{2} (0.04 + 0.02) = 0.027$$

The linkage disequilibria in males and females for linked loci in generation 2 are

$$\Delta_2^m = \rho_0^m \frac{1}{2} (\Delta_1^m + \Delta_1^f) = (1) \frac{1}{2} (0.03 + 0.027) = 0.0285$$

$$\Delta_2^f = \rho_0^f \frac{1}{2} (\Delta_1^m + \Delta_1^f) = (0.9) \frac{1}{2} (0.03 + 0.027) = 0.02565$$

The general expressions for linkage disequilibrium in the males and females as a function of initial linkage disequilibrium are

$$\Delta_1^m = \rho_0^m \bar{\Delta}_0 \qquad \Delta_1^f = \rho_0^f \bar{\Delta}_0$$

$$\begin{aligned} \Delta_2^m &= \rho_0^m \frac{1}{2} (\Delta_1^m + \Delta_1^f) & \Delta_2^f &= \rho_0^f \frac{1}{2} (\Delta_1^m + \Delta_1^f) \\ &= \rho_0^m \frac{1}{2} (\rho_0^m \bar{\Delta}_0 + \rho_0^f \bar{\Delta}_0) & &= \rho_0^f \frac{1}{2} (\rho_0^m \bar{\Delta}_0 + \rho_0^f \bar{\Delta}_0) \\ &= \rho_0^m \frac{\rho_0^m + \rho_0^f}{2} \bar{\Delta}_0 & &= \rho_0^f \frac{\rho_0^m + \rho_0^f}{2} \bar{\Delta}_0 \\ &= \rho_0^m \bar{\rho}_0 \bar{\Delta}_0 & &= \rho_0^f \bar{\rho}_0 \bar{\Delta}_0 \end{aligned}$$

$$\begin{aligned} \Delta_3^m &= \rho_0^m \frac{1}{2} (\Delta_2^m + \Delta_2^f) & \Delta_3^f &= \rho_0^f \frac{1}{2} (\Delta_2^m + \Delta_2^f) \\ &= \rho_0^m \frac{1}{2} (\rho_0^m \bar{\rho}_0 \bar{\Delta}_0 + \rho_0^f \bar{\rho}_0 \bar{\Delta}_0) & &= \rho_0^f \frac{1}{2} (\rho_0^m \bar{\rho}_0 \bar{\Delta}_0 + \rho_0^f \bar{\rho}_0 \bar{\Delta}_0) \\ &= \rho_0^m \frac{\rho_0^m + \rho_0^f}{2} \bar{\rho}_0 \bar{\Delta}_0 & &= \rho_0^f \frac{\rho_0^m + \rho_0^f}{2} \bar{\rho}_0 \bar{\Delta}_0 \\ &= \rho_0^m \bar{\rho}_0^2 \bar{\Delta}_0 & &= \rho_0^f \bar{\rho}_0^2 \bar{\Delta}_0 \end{aligned}$$

$$\Delta_t^m = \rho_0^m \bar{\rho}_0^{(t-1)} \bar{\Delta}_0 \qquad \Delta_t^f = \rho_0^f \bar{\rho}_0^{(t-1)} \bar{\Delta}_0$$

By the use of these general expressions the linkage disequilibrium values in generation 3 are

$$\Delta_{t=3}^m = \rho_0^m \bar{\rho}_0^{(t-1)} \bar{\Delta}_0 = (1)(0.95)^2 (0.03) = 0.027075$$

$$\Delta_{t=3}^f = \rho_0^f \bar{\rho}_0^{(t-1)} \bar{\Delta}_0 = (0.9)(0.95)^2 (0.03) = 0.0243675$$

We calculate the mean linkage disequilibria for generation 1, 2, and 3:

$$\bar{\Delta}_1 = \bar{\rho}_0 \bar{\Delta}_0 = 0.95(0.03) = 0.0285$$

$$\bar{\Delta}_2 = \bar{\rho}_0^2 \bar{\Delta}_0 = (0.95)^2 (0.03) = 0.027075$$

$$\bar{\Delta}_3 = \bar{\rho}_0^3 \bar{\Delta}_0 = (0.95)^3 (0.03) = 0.02572125$$

which agrees with the mean of the disequilibrium values for the males and females:

$$\frac{\Delta_1^m + \Delta_1^f}{2} = \frac{0.03 + 0.027}{2} = 0.0285$$

$$\frac{\Delta_2^m + \Delta_2^f}{2} = \frac{0.0285 + 0.02565}{2} = 0.027075$$

$$\frac{\Delta_3^m + \Delta_3^f}{2} = \frac{0.027075 + 0.0243675}{2} = 0.02572125$$

b. For loci on separate chromosomes (so recombination does occur) we have

$$\Delta_1^m = \rho_0^m \frac{1}{2} (\Delta_0^m + \Delta_0^f) = 0.5 \left( \frac{1}{2} \right) (0.04 + 0.02) = \frac{0.03}{2} = 0.015$$

$$\Delta_1^f = \rho_0^f \frac{1}{2} (\Delta_0^m + \Delta_0^f) = 0.5 \left( \frac{1}{2} \right) (0.04 + 0.02) = \frac{0.03}{2} = 0.015$$

The linkage disequilibrium values become the same for both sexes.

The easiest way to calculate the linkage disequilibrium is to calculate the mean of the initial linkage disequilibria and to apply a formula analogous to (3.86).

$$\Delta_t = \rho_0^t \bar{\Delta}_0$$

$$\Delta_3 = \left( \frac{1}{2} \right)^3 (0.03) = \frac{0.03}{8} = 0.00375$$

It is dissipated in the same manner as occurs in a single random-mating population.

$$\Delta_{A_i B_j t} = \rho_0^t \Delta_{A_i B_j 0}$$

### Exercise 3.31.

Equation (3.102) embodies the conditions when either linkage equilibrium or linkage disequilibrium exists in the gametes produced from the cross population itself. The equation is repeated here.

$$\Delta_{A_i B_j}^c = \frac{1}{2} \rho_0 \left( \Delta_{A_i B_j}^m + \Delta_{A_i B_j}^f \right) + \frac{1}{4} (2\rho_0 - 1) \left( p_{A_i}^m - p_{A_i}^f \right) \left( p_{B_j}^m - p_{B_j}^f \right)$$

It is observed that there are two major terms on the right-hand side of the equation, term 1 and term 2.

Linkage equilibrium exists in the gametes of the cross population when equation (3.102) is equal to zero. This occurs when both terms equal zero, or when these two terms are equal in magnitude but opposite in sign.

The first term is equal to zero when  $\Delta_{A_i B_j}^m$  and  $\Delta_{A_i B_j}^f$  are both equal to zero or when they are equal in magnitude

but opposite in sign, i.e.,  $\Delta_{A_i B_j}^m = -\Delta_{A_i B_j}^f$ . The second term is equal to zero when  $\rho_0 = \frac{1}{2}$  or

$$p_{A_i}^m = p_{A_i}^f \text{ or } p_{B_j}^m = p_{B_j}^f.$$

Linkage disequilibrium exists in the gametes of the cross population when equation (3.102) is not equal to zero. This occurs when one or both terms are not equal to zero and do not cancel. This occurs when

1)  $\Delta_{A_i B_j}^m$  and/or  $\Delta_{A_i B_j}^f$  are not equal to zero, i.e., one or both populations are in linkage disequilibrium, and do not cancel, i.e.,  $\Delta_{A_i B_j}^m \neq -\Delta_{A_i B_j}^f$ , and

2) term 2 is not equal to zero when the genes are linked ( $\rho_0 > \frac{1}{2}$ ),  $p_{A_i}$  is not equivalent in both populations ( $p_{A_i}^m \neq p_{A_i}^f$ ), and  $p_{B_j}$  is not equivalent in both populations ( $p_{B_j}^m \neq p_{B_j}^f$ ).

### Exercise 3.32.

The expression for linkage disequilibrium in a cross is equation (3.102):

$$\Delta_{A_i B_j}^c = \frac{1}{2} \rho_0 \left( \Delta_{A_i B_j}^m + \Delta_{A_i B_j}^f \right) + \frac{1}{4} (2\rho_0 - 1) \left( p_{A_i}^m - p_{A_i}^f \right) \left( p_{B_j}^m - p_{B_j}^f \right)$$

If that expression sums to zero for every combination, we have linkage equilibrium at loci  $A$  and  $B$ . There can be specific cancelling effect within the first major term and/or between the two major terms. If the first major term sums to zero for all combinations of  $i$  and  $j$  and there is independence (no linkage) or the same frequency of every  $A_i$  allele exists in the two populations or the same frequency of every  $B_j$  allele exists in the two population, there is linkage equilibrium. If any one of the three factors is zero, the second term equals zero.

### Exercise 3.33.

In Example 3.2, the observed departures are all due to linkage disequilibrium. For example, in generation 2 the frequency of any genotype is a function of linkage disequilibrium, i.e., for  $AB/AB$ , we have

$$\begin{aligned} p\left(\frac{AB}{AB}\right)_2 &= (p_{AB,1})^2 = (p_A p_B + \Delta_{AB,1})^2 \\ &= \left[ \frac{1}{2} \left( \frac{1}{2} \right) + \frac{1}{8} \right]^2 \\ &= \left( \frac{3}{8} \right)^2 \\ &= \frac{9}{64} \end{aligned}$$

Likewise for the double heterozygotes,  $AB/ab$  and  $Ab/aB$ , the sum of the frequencies of coupling and repulsion phase, heterozygotes is

$$\begin{aligned} p\left(\frac{AB}{ab}\right)_2 + p\left(\frac{Ab}{aB}\right)_2 &= 2(p_{AB,1} p_{ab,1}) + 2(p_{Ab,1} p_{aB,1}) \\ &= 2(p_A p_B + \Delta_{AB,1})(p_a p_b + \Delta_{ab,1}) + 2(p_A p_b + \Delta_{Ab,1})(p_a p_B + \Delta_{aB,1}) \\ &= 2 \left[ \frac{1}{2} \frac{1}{2} + \frac{1}{8} \right] \left[ \frac{1}{2} \frac{1}{2} + \frac{1}{8} \right] + 2 \left[ \frac{1}{2} \frac{1}{2} - \frac{1}{8} \right] \left[ \frac{1}{2} \frac{1}{2} - \frac{1}{8} \right] \\ &= 2 \left( \frac{3}{8} \right)^2 + 2 \left( \frac{1}{8} \right)^2 \\ &= 2 \left( \frac{9}{64} \right) + 2 \left( \frac{1}{64} \right) = \frac{9}{32} + \frac{1}{32} = \frac{5}{16} \end{aligned}$$

The difference between the frequencies of the coupling and repulsion phase, double heterozygotes is

$$p\left(\frac{AB}{ab}\right)_2 - p\left(\frac{Ab}{aB}\right)_2 = \frac{9}{32} - \frac{1}{32} = \frac{1}{4}$$

This difference in genotypic frequencies between coupling and repulsion phase, double heterozygotes is an equivalent measure of linkage disequilibrium [see equation (3.66)]; the difference in generation 2 is twice the linkage disequilibrium in generation 1 as we have designated the generations [see equations (3.64) and (3.72)],

namely,  $2\Delta_{AB1} = 2\left(\frac{1}{8}\right) = \frac{1}{4}$ .

In Example 3.4, the frequency of all nine genotypes (coupling and repulsion added together) is equal to the product of the marginal frequencies of the corresponding single-locus genotypic values. This is true because in this example both parental populations were in linkage equilibrium. It is this fact that makes the following true (along with random mating):

$$\begin{aligned} p\left(\frac{AB}{AB}\right)_c &= p(AA)_c p(BB)_c \\ &= p_A^2 p_B^2 \\ &= (p_A p_B)^2 \quad \text{-- the product } p_A \text{ times } p_B \text{ implies linkage equilibrium, and} \\ &\quad \text{and the product } p_A p_B \text{ times } p_A p_B \text{ implies random mating} \\ &= (p_{AB})^2 \end{aligned}$$

This last expression above implies that the linkage disequilibrium is zero because

$(p_{AB})^2 = (p_{APB} + \Delta_{AB})^2$  where  $\Delta_{AB} = 0$ . This is also true for the sum of the coupling and repulsion double heterozygotes

$p\left(\frac{AB}{ab}\right) + p\left(\frac{Ab}{aB}\right) = 0.075 + 0.095 = 0.170$  which is equal to the product of the two single-locus frequencies,  $p_{Aa} p_{Bb} = 0.340(0.50) = 0.170$ , implying linkage equilibrium.

Now why do the frequencies of the coupling and repulsion double heterozygotes differ? It is simply because the frequency of the four kinds of gametes in the two parental populations are different,

$$\begin{aligned} p\left(\frac{AB}{ab}\right) - p\left(\frac{Ab}{aB}\right) &= p_{AB}^m p_{ab}^f + p_{ab}^m p_{AB}^f - \left[ p_{Ab}^m p_{aB}^f + p_{aB}^m p_{Ab}^f \right] \\ &= 0.42(0.05) + 0.12(0.45) - \left[ 0.28(0.05) + 0.18(0.45) \right] \\ &= 0.021 + 0.054 - (0.014 + 0.081) \\ &= 0.075 - 0.095 \end{aligned}$$

If the corresponding frequencies of the four kinds of gametes were the same in the two parental populations, then the difference between the frequencies of the coupling and repulsion double heterozygotes would be zero. For example, if the  $f$  population would have the same gametic frequencies as the  $m$  population, then we would have  $0.42(0.12) + 0.12(0.42) - [0.28(0.18) + 0.18(0.28)] = 0.504 + 0.504 - [0.504 + 0.504] = 0$ .

### Exercise 3.34.

a. The linkage disequilibrium for the  $AB$  gamete from the  $m$  population is

$$\begin{aligned} \Delta_{AB}^m &= p_{AB}^m - p_A^c(p_B^c) \\ &= 0.42 - (0.8)(0.55) = 0.42 - 0.44 = -0.02 \end{aligned}$$

The linkage disequilibrium for the  $AB$  gamete from the  $f$  population is

$$\begin{aligned} \Delta_{AB}^f &= p_{AB}^f - p_A^c(p_B^c) \\ &= 0.45 - (0.8)(0.55) = 0.45 - 0.44 = 0.01 \end{aligned}$$

Notice that the linkage disequilibrium value of  $-0.02$  for the  $m$  population and that of  $0.01$  for the  $f$  population are opposite in sign. The mean of the two values is

$$\bar{\Delta}_{AB} = \frac{-0.02 + 0.01}{2} = -0.005$$

An alternative way to calculate  $\bar{\Delta}_{AB}$  is by use of the mean of the gametic frequencies:

$$\bar{\Delta}_{AB} = \frac{0.42 + 0.45}{2} - (0.8)(0.55) = 0.435 - 0.44 = -0.005$$

b. The corresponding values for the other three gametes are:

$$\text{Ab} \quad \Delta_{Ab}^m = p_{Ab}^m - p_A^c(p_b^c) = 0.28 - (0.8)(0.45) = 0.28 - 0.36 = -0.08$$

$$\Delta_{Ab}^f = 0.45 - 0.36 = 0.09$$

$$\bar{\Delta}_{Ab} = \frac{-0.08+0.09}{2} = \frac{0.01}{2} = 0.005$$

$$\text{aB} \quad \Delta_{aB}^m = p_{aB}^m - p_a^c(p_B^c) = 0.18 - (0.2)(0.55) = 0.18 - 0.11 = 0.07$$

$$\Delta_{aB}^f = 0.05 - 0.11 = -0.06$$

$$\bar{\Delta}_{aB} = \frac{0.07-0.06}{2} = \frac{0.01}{2} = 0.005$$

$$\text{ab} \quad \Delta_{ab}^m = p_{ab}^m - p_a^c(p_b^c) = 0.12 - (0.2)(0.45) = 0.12 - 0.09 = 0.03$$

$$\Delta_{ab}^f = 0.05 - 0.09 = -0.04$$

$$\bar{\Delta}_{ab} = \frac{0.03+(-0.04)}{2} = \frac{-0.01}{2} = -0.005$$

c. The mean of the gametic frequencies from population  $m$  and population  $f$  are

$$\bar{p}_{AB} = \frac{0.42+0.45}{2} = 0.435$$

$$\bar{p}_{Ab} = \frac{0.28+0.45}{2} = 0.365$$

$$\bar{p}_{aB} = \frac{0.18+0.05}{2} = 0.115$$

$$\bar{p}_{ab} = \frac{0.12+0.05}{2} = 0.085$$

The linkage disequilibrium is [see equation (3.64)]

$$\bar{\Delta}_{AB} = \bar{p}_{AB}\bar{p}_{ab} - \bar{p}_{Ab}\bar{p}_{aB} = 0.435(0.085) - 0.365(0.115) = 0.036975 - 0.041975 = -0.005$$

No, it is not equal to  $-0.010$ , the value given in the Exercise itself; it is one half of that value. It is the same as that in (a) and (b).

Why? I am not too clear on the possible reason, but that measure of  $-0.010$  is the result of union of individual gametes from the  $m$  population and from the  $f$  population. There is no averaging effect to reduce the deviation. When we define linkage equilibrium for each gamete from the  $m$  and  $f$  populations relative to the cross itself, we observe that they are opposite in sign and show less deviation (like a mean). I would suggest that the reason for the relation is that we are using the mean of the gametic frequencies from the two populations, i.e.,

$$\bar{\Delta}_{AB} = \frac{0.42+0.45}{2} - (0.8)(0.55) = 0.435 - 0.440 = -0.005$$

The original measure of  $-0.010$  is the union of gametes from the  $m$  and  $f$  population and there is no averaging involved.

### Exercise 3.35.

Similarities between two-locus and three-locus disequilibria are:

(1) Two-locus and three-locus linkage disequilibria are defined in the same way, namely, as a departure of their gametic frequencies from that of independence of the corresponding genes

$$\Delta_{A_i B_j} = p_{A_i B_j} - p_{A_i} p_{B_j} \quad \text{equation (3.61)}$$

$$\Delta_{A_i B_j C_k} = p_{A_i B_j C_k} - p_{A_i} p_{B_j} p_{C_k} \quad \text{equation (3.112)}$$

(2) Gametic frequencies [and hence their linkage disequilibria, see equation (3.63)] for both can be expressed in generation  $t$  in terms of complementary gametic group probabilities

$$P_{A_i B_j t} = \rho_0 P_{A_i B_j t-1} + \rho_1 P_{A_i} P_{B_j} \quad \text{equation (3.82)}$$

$$P_{A_i B_j C_k t} = \rho_{00} P_{A_i B_j C_k t-1} + \rho_{01} P_{C_k} P_{A_i B_j t-1} + \rho_{10} P_{A_i} P_{B_j C_k t-1} + \rho_{11} P_{B_k} P_{A_i C_k t-1} \quad \text{equation (3.114)}$$

(3) The higher-order, three-locus linkage disequilibrium measure or the adjusted three-locus linkage disequilibrium value behaves like the two-locus linkage disequilibrium measure, i.e.,

$$\delta_{A_i B_j C_k t} = \rho_{00} \delta_{A_i B_j C_k t-1} \quad \text{vs.} \quad \Delta_{A_i B_j t} = \rho_0 P_{A_i B_j t-1} \quad \text{equation (3.119) vs. equation (3.86)}$$

$$\delta_{A_i B_j C_k t} = \rho_{00}^t \delta_{A_i B_j C_k 0} \quad \text{vs.} \quad \Delta_{A_i B_j t} = \rho_0^t P_{A_i B_j 0} \quad \text{equation (3.119) vs. equation (3.86)}$$

We regard the three-locus linkage disequilibrium as less important because the adjusted three-locus linkage disequilibrium value is dissipated more rapidly than the two-locus linkage disequilibrium value, i.e., rate of loss is  $3\bar{\rho}_1/2$  [see equation (3.125)] which is 50% greater than that for two loci for an equivalent recombination value.

**Exercise 3.36.**

We assume the following allelic frequencies:

$$p_A = 0.6, p_a = 0.4$$

$$p_B = 0.7, p_b = 0.3$$

$$p_C = 0.5, p_c = 0.5$$

Then the equilibrium gametic frequencies are:

$$ABC = 0.6(0.7)(0.5) = 0.21$$

$$ABc = 0.6(0.7)(0.5) = 0.21$$

$$AbC = 0.6(0.3)(0.5) = 0.09$$

$$Abc = 0.6(0.3)(0.5) = 0.09$$

$$aBC = 0.4(0.7)(0.5) = 0.14$$

$$aBc = 0.4(0.7)(0.5) = 0.14$$

$$abC = 0.4(0.3)(0.5) = 0.06$$

$$abc = 0.4(0.3)(0.5) = 0.06$$

We create the three-locus linkage disequilibrium by adding or subtracting an arbitrary amount (0.01) from the three-locus gametic equilibrium values. Since it is a three-way table, we construct two two-way tables—one for gametes which carry the  $C$  allele and one which carries the  $c$  allele.

		Gametes with C allele					Gametes with c allele																																							
		B	b				B	b																																						
A	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 5px;">ABC</td> <td style="padding: 5px;">AbC</td> <td style="padding: 5px;">AC</td> </tr> <tr> <td style="padding: 5px;">0.22</td> <td style="padding: 5px;">0.08</td> <td style="padding: 5px;">0.30</td> </tr> <tr> <td style="padding: 5px;">= 0.21 + 0.01</td> <td style="padding: 5px;">= 0.09 - 0.01</td> <td style="padding: 5px;">=(0.6)(0.5)</td> </tr> </table>	ABC	AbC	AC	0.22	0.08	0.30	= 0.21 + 0.01	= 0.09 - 0.01	=(0.6)(0.5)	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 5px;">ABc</td> <td style="padding: 5px;">Abc</td> <td style="padding: 5px;">Ac</td> </tr> <tr> <td style="padding: 5px;">0.20</td> <td style="padding: 5px;">0.10</td> <td style="padding: 5px;">0.30</td> </tr> <tr> <td style="padding: 5px;">= 0.21 - 0.01</td> <td style="padding: 5px;">= 0.09 + 0.01</td> <td style="padding: 5px;">=(0.6)(0.5)</td> </tr> </table>	ABc	Abc	Ac	0.20	0.10	0.30	= 0.21 - 0.01	= 0.09 + 0.01	=(0.6)(0.5)	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 5px;">aBC</td> <td style="padding: 5px;">abC</td> <td style="padding: 5px;">aC</td> </tr> <tr> <td style="padding: 5px;">0.13</td> <td style="padding: 5px;">0.07</td> <td style="padding: 5px;">0.20</td> </tr> <tr> <td style="padding: 5px;">= 0.14 - 0.01</td> <td style="padding: 5px;">= 0.06 + 0.01</td> <td style="padding: 5px;">=(0.4)(0.5)</td> </tr> </table>	aBC	abC	aC	0.13	0.07	0.20	= 0.14 - 0.01	= 0.06 + 0.01	=(0.4)(0.5)	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 5px;">aBc</td> <td style="padding: 5px;">abc</td> <td style="padding: 5px;">ac</td> </tr> <tr> <td style="padding: 5px;">0.15</td> <td style="padding: 5px;">0.05</td> <td style="padding: 5px;">0.20</td> </tr> <tr> <td style="padding: 5px;">= 0.14 + 0.01</td> <td style="padding: 5px;">= 0.06 - 0.01</td> <td style="padding: 5px;">=(0.4)(0.5)</td> </tr> </table>	aBc	abc	ac	0.15	0.05	0.20	= 0.14 + 0.01	= 0.06 - 0.01	=(0.4)(0.5)	BC	bC	C	Bc	bc	c
ABC	AbC	AC																																												
0.22	0.08	0.30																																												
= 0.21 + 0.01	= 0.09 - 0.01	=(0.6)(0.5)																																												
ABc	Abc	Ac																																												
0.20	0.10	0.30																																												
= 0.21 - 0.01	= 0.09 + 0.01	=(0.6)(0.5)																																												
aBC	abC	aC																																												
0.13	0.07	0.20																																												
= 0.14 - 0.01	= 0.06 + 0.01	=(0.4)(0.5)																																												
aBc	abc	ac																																												
0.15	0.05	0.20																																												
= 0.14 + 0.01	= 0.06 - 0.01	=(0.4)(0.5)																																												
	0.35	0.15	0.5	0.35	0.15	0.5		0.35	0.15	0.5																																				
	=(0.7)(0.5)	=(0.3)(0.5)		=(0.7)(0.5)	=(0.3)(0.5)			=(0.7)(0.5)	=(0.3)(0.5)																																					

From the marginal frequencies of the two above tables it is observed that linkage equilibrium exists for loci  $A$  and  $C$  and for loci  $B$  and  $C$ . By adding the corresponding cells in the two above tables one obtains the gametic frequencies for loci  $A$  and  $B$  which is also in linkage equilibrium:

$$\text{gamete } AB = 0.22 + 0.20 = 0.42 = (0.6)(0.7)$$

$$\text{gamete } Ab = 0.08 + 0.10 = 0.18 = (0.6)(0.3)$$

$$\text{gamete } aB = 0.13 + 0.15 = 0.28 = (0.4)(0.7)$$

$$\text{gamete } ab = 0.07 + 0.05 = 0.12 = (0.4)(0.3)$$

The three-locus gametic frequencies, of course, exhibit linkage disequilibrium:

$$\Delta_{ABC} = \Delta_{abC} = \Delta_{Abc} = \Delta_{aBc} = 0.01$$

$$\Delta_{AbC} = \Delta_{aBC} = \Delta_{ABC} = \Delta_{abc} = -0.01$$

### Exercise 3.37.

Mating type	Offspring							
	Male	Female	Frequency	A	a	AA	Aa	aa
$A$	$AA$		$pp^2$	1		1		
	$Aa$		$p \cdot 2pq$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	
	$aa$		$p \cdot q^2$		1		1	
$a$	$AA$		$q \cdot p^2$	1			1	
	$Aa$		$q \cdot 2pq$	$\frac{1}{2}$	$\frac{1}{2}$		$\frac{1}{2}$	$\frac{1}{2}$
	$aa$		$q \cdot q^2$		1			1

Male:

$$\begin{aligned} p_A : & pp^2(1) + p \cdot 2pq \frac{1}{2} + q \cdot p^2(1) + q \cdot 2pq \frac{1}{2} \\ &= p^3 + p^2q + p^2q + pq^3 \\ &= p^2(p+q) + pq(p+q) \\ &= p^2 + pq \\ &= p(p+q) \\ &= p \end{aligned}$$

$$\begin{aligned} p_a : & p \cdot 2pq \frac{1}{2} + p \cdot q^2(1) + q \cdot 2pq \frac{1}{2} + q \cdot q^2(1) \\ &= p^2q + pq^2 + pq^2 + q^3 \\ &= pq(p+q) + q^2(p+q) \\ &= pq + q^2 \\ &= q(p+q) \\ &= q \end{aligned}$$

Female:

$$p_{AA} : pp^2(1) + p \cdot 2pq \frac{1}{2} = p^3 + p^2q = p^2(p+q) = p^2$$

$$\begin{aligned} p_{Aa} : p \cdot 2pq \frac{1}{2} + p \cdot q^2(1) + q \cdot p^2(1) + q \cdot 2pq \frac{1}{2} \\ = p^2q + pq^2 + qp^2 + pq^2 \\ = pq(p+q+p+q) = 2pq \end{aligned}$$

$$p_{aa} : q \cdot 2pq \frac{1}{2} + q \cdot q^2(1) = pq^2 + q^3 = q^2(p+q) = q^2$$

**Exercise 3.38.**

The percentage of males showing the trait is

$$\sqrt{0.2025} = 0.45 \text{ or } 45\%$$

That proportion is the same for a nonrandom-mating population as for a random-mating population.

**Exercise 3.39.**

The probability that the woman is a carrier is one-half, or  $P(\text{woman is carrier}) = \frac{1}{2}$

If she is a carrier, the probability that she will have an affected son is one-half, or

$$P(\text{son is affected} | \text{woman is carrier}) = \frac{1}{2},$$

$$\text{or } P(\text{son is affected}) = P(\text{woman is carrier})P(\text{son is affected} | \text{woman is carrier}) = \frac{1}{4}$$

**Exercise 3.40.**

$$\text{a. } p^2q + pq^2 = pq(p+q) = pq = (1-q)q = q - q^2$$

$$\text{b. } q^3 + pq^2 = q^2(q+p) = (1-p)^2$$

$$\text{c. } (1-2q)^2 = [1-2(1-p)]^2 = [1-2+2p]^2 = (2p-1)^2$$

$$\text{d. } p^2 - q^2 = (p+q)(p-q) = p - q$$

$$\begin{aligned} \text{e. } 1 - 2pq - 2pq &= p^2 + q^2 - 2pq = (q-p)^2 \text{ or} \\ p^2 + 2pq + q^2 - 4pq &= q^2 - 2pq + p^2 = (q-p)^2 \end{aligned}$$

$$\text{f. } \frac{1-2p}{pq} = \frac{p+q-2p}{pq} = \frac{q-p}{pq} = \frac{q}{pq} - \frac{p}{pq} = \frac{1}{p} - \frac{1}{q}$$