EXERCISES FOR CHAPTER 11

Exercise 11.1. In equation (11.4) I give an expression for the genotypic variance of a population inbred to the level of the coefficient of inbreeding \( F \) for a single locus. It agrees with an equation given by Sewall Wright (1951, The genetical structure of populations, Annals of Eugenics 15:323-354) on page 343, line 15 below “Properties of populations as related to \( F \)”. (For your information I present below a portion of the article on p. 324 which refers to Appendix C and Appendix C itself on pp. 343 and 344.) Wright seems to imply that the equation for a single locus is the same as that for multiple loci with no epistasis.

Statistical properties of populations

The statistical properties of a population change with changes in zygotic composition even though gene frequencies are not changed (Appendix C). Consider first the case of exact semi-dominance. The mean is not affected, but the genetic variance of the population as a whole increases from its value under random mating \( (\sigma^2_{\pi \sigma}) \) by the proportion \( F \) becoming \( (1 + F)\sigma^2_{\pi \sigma} \). If the system of mating is one of subdivision into strains with internal random mating, the average variance within these falls off by the proportion \( F \), becoming \( (1 - F)\sigma^2_{\pi \sigma} \) and the variance of strain means becomes \( 2F\sigma^2_{\pi \sigma} \).

If, on the other hand, there is any departure from semi-dominance, the mean shifts from its value, \( m_{\pi \sigma} \), under random mating toward a value \( m_{\pi \mu} \) characteristic of an array of completely fixed lines and the amount of change is proportional to \( F \). The decrease in size, fecundity and viability that are usually observed on inbreeding seem to have their basis in an association of recessiveness with deleterious effect.

APPENDIX C

Properties of populations as related to \( F \)

The relation of the mean and variance of characters to the inbreeding coefficient, noted in the text, apply to multiple alleles as well as to pairs of alleles. We assume here that contributions of loci and of environmental factors are additive. The array of zygotic frequencies of \( n \) alleles of type \( A_i \) gene frequency \( q_i \) is

\[
(1 - F) [\sum q_i A_i]^2 + F[\sum q_i A_i A_i].
\]

Let \( c_{ij} \) represent the contribution of \( A_i A_j \) to the character. Let \( m_{T(F)} \) be the mean and \( \sigma^2_{T(F)} \) the genetic component of the variance under a system of mating characterized by coefficient \( F \). Then \( m_{T(0)} \) and \( \sigma^2_{T(0)} \) are the corresponding statistics under random mating \( (F = 0) \), and \( m_{T(0)} \) and \( \sigma^2_{T(0)} \) those for an array of completely fixed lines \( (F = 1) \), arrived at without selection:

\[
m_{T(F)} = (1 - F) \sum_{j=1}^{n} \sum_{i=1}^{n} (c_{ij} q_i q_j) + F \sum_{i=1}^{n} (c_{ii} q_i^2) = (1 - F) m_{T(0)} + F m_{T(0)}
\]

\[
\sigma^2_{T(F)} = (1 - F) \sum (c_{ij} q_i q_j) + F \sum (c_{ii} q_i^2) - m_{T(F)}^2
\]

\[
= (1 - F) (\sigma^2_{T(0)} + m_{T(0)}^2) + F (\sigma^2_{T(0)} + m_{T(0)}^2) - [(1 - F) m_{T(0)} + F m_{T(0)}]^2
\]

Thus the change in mean due to the system of mating is proportional to \( F \), but that in the total variance is a quadratic fraction of \( F \).

The case of semi-dominance at all pertinent loci of interest. Let \( c_i \) be contribution of allele \( A_i \) and \( c_{ij} = c_i + c_j \). The mean is unaffected by the system of mating in this case:
However, in equation (11.27) I present an equation for multiple loci which seems to differ from Wright’s equation on p. 343, line 15. My equation (11.27) does reduce to equation (11.4) for \( n = 1 \) in that the variance \( \sigma^2_{TH} = 2 \sum c_i q_i \) for \( n = 1 \) in that the variance

\[
\sigma^2_{TH} = \sum_j \sum_{i} (c_i + c_j)^2 q_i q_j - 4(\sum c_i q_i)^2
\]

\[
= 2 \sum c_i^2 q_i - 2(\sum c_i q_i)^2 = 2\sigma^2_{c_i}
\]

\[
\sigma^2_{TF} = 4 \sum c_i^2 q_i - 4(\sum c_i q_i)^2 = 4\sigma^2_{c_i}
\]

\[
\sigma^2_{TF} = (1 - F) \sigma^2_{TH} + F \sigma^2_{TF}
\]

If the inbreeding is due to isolation of a number \( (n) \) of strains \( (s) \) breeding at random within themselves, the variance within one is as follows, letting \( q_{is} \) be the frequency of \( A_i \) in this strain:

\[
\sigma^2_{S(0)} = 2n \sum c_i^2 q_{is} - (\sum c_i q_{is})^2
\]

\[
= 2n \sum c_i^2 q_{is}(1 - q_{is}) - 2n c_i c_j q_{is} q_{js}
\]

But \( 2q_{is}(1 - q_{is}) \) is the amount of heterozygosis of \( A_i \) with all other alleles collectively, within the strain. The average for all strains is \( 2q_i(1 - q_i)(1 - F) \). Similarly, \( 2q_{is} q_{js} \) is the proportion of the specific heterozygote \( A_i A_j \) within the strain, the average of which for all strains is \( 2q_i q_j(1 - F) \). Thus, the average intra-strain variance is

\[
\sigma^2_{S(0)} = 2n \sum c_i^2 q_{is}(1 - q_{is}) - 2n c_i c_j q_{is} q_{js}
\]

\[
= 2n \sum c_i^2 q_{is} - (\sum c_i q_{is})^2
\]

\[
= (1 - F) \sigma^2_{TF}
\]

The variance of strain means is

\[
\sigma^2_{mS} = \sigma^2_{S(0)} - \sigma^2_{S(0)}
\]

\[
\sigma^2_{mS} = 2F \sigma^2_{S(0)}
\]

The genetic correlation between any relatives is readily found by the method of path coefficients assuming additive effects of genes.

In the case of autosomal diploids, the correlation between propositi \( Z_1 \) and \( Z_2 \) is given by

\[
r_{Z_1, Z_2} = \frac{\sum[(\frac{1}{2})^n - (1 + F_{2Z_1})]}{\sqrt{[1 + F_{Z_1}][1 + F_{Z_2}]}} \quad \text{(Wright, 1922 a)}
\]

For sex-linked loci

\[
r_{Z_1, Z_2} = \frac{\sum[(\frac{1}{2})^n - n Z(1 + F_{2Z_1})]}{\sqrt{[1 + F_{Z_1}][1 + F_{Z_2}]]}
\]

where \( n_{Zf} \) is the number of female propositi \( (0, 1 \) or \( 2 \). These formulae may be modified in the usual ways to take care of dominance, types of factor interaction and environmental effects.

However, in equation (11.27) I present an equation for multiple loci which seems to differ from Wright’s equation on p. 343, line 15. My equation (11.27) does reduce to equation (11.4) for \( n = 1 \) in that the variance \( \sigma^2_{H} = 0 \). If one assumes multiple loci and equal inbreeding depressions for all loci so that \( \sigma^2_{H} = 0 \), it would appear that the \( (\mu_1 - \mu_0)^2 \) term would still have a divisor \( n \), i.e., the complete equation for the genotypic variance for the inbred population would be

\[
\sigma^2_{G(F)} = (1 - F) \sigma^2_{G(0)} + F \sigma^2_{G(1)} + F(1 - F) \frac{(\mu_1 - \mu_0)^2}{n}
\]

Have I made a mistake or failed to recognize one or more implied assumptions made by Wright? Or, did Wright make a mistake? Who is correct? Discuss.
Exercise 11.2. The genotypic variance at a single locus with multiple alleles for any arbitrarily inbred population is expressed in terms of five genetic parameters defined in terms of effects in the original random-mating, noninbred reference population. Define these five parameters symbolically in terms of effects of the noninbred reference population and state, in words, what each is equal to.

Exercise 11.3. The dominance effects associated with the homozygous genotypes in a random-mating population are of what particular value in the application of statistical genetics?

Exercise 11.4. The concept of identity by descent has proven to be a very useful one in genetics. We have used it in the derivation of a number of equations or general expressions. Enumerate as many of these different situations, as you can, where the concept of identity by descent has been used. You may reference chapter numbers, section numbers, and/or equation numbers in my notes to identify the use of identity by descent.

Exercise 11.5. In Johannsen’s classical pure-line experiment in which he had 19 different pure lines of beans, he observed the mean seed weight of a single individual within each line in 1901, and seed weight of many offspring individuals from each of the parental individuals in 1902. The seed weights observed in 1901 and 1902, and the standard deviations of the individuals in each line in 1902 are (weights in mg):

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1901 (parents):</td>
<td>600</td>
<td>520</td>
<td>570</td>
<td>600</td>
<td>512</td>
<td>395</td>
<td>440</td>
<td>405</td>
<td>395</td>
<td>400</td>
<td>380</td>
<td>410</td>
<td>400</td>
<td>390</td>
<td>350</td>
<td>312</td>
<td>310</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1902 (offspring mean):</td>
<td>642</td>
<td>558</td>
<td>554</td>
<td>547</td>
<td>512</td>
<td>506</td>
<td>492</td>
<td>488</td>
<td>482</td>
<td>465</td>
<td>455</td>
<td>455</td>
<td>454</td>
<td>453</td>
<td>450</td>
<td>446</td>
<td>428</td>
<td>407</td>
<td>351</td>
</tr>
<tr>
<td>1902 (s):</td>
<td>109</td>
<td>93</td>
<td>76</td>
<td>84</td>
<td>76</td>
<td>64</td>
<td>69</td>
<td>72</td>
<td>76</td>
<td>79</td>
<td>70</td>
<td>66</td>
<td>74</td>
<td>75</td>
<td>66</td>
<td>69</td>
<td>72</td>
<td>78</td>
<td>65</td>
</tr>
<tr>
<td>1901 (parents):</td>
<td>380</td>
<td>410</td>
<td>400</td>
<td>390</td>
<td>510</td>
<td>360</td>
<td>340</td>
<td>312</td>
<td>310</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1902 (offspring mean):</td>
<td>455</td>
<td>455</td>
<td>454</td>
<td>453</td>
<td>450</td>
<td>446</td>
<td>428</td>
<td>407</td>
<td>351</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1902 (s):</td>
<td>70</td>
<td>66</td>
<td>74</td>
<td>75</td>
<td>66</td>
<td>69</td>
<td>72</td>
<td>78</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Find the linear regression coefficient of 1902 weights on 1901 weights. What explanation(s) are possible for the regression toward the grand mean, i.e., for the regression coefficient being less than one? Discuss.

Exercise 11.6. In Chapter 22, entitled, Quantitative Genetics, in a general genetics textbook by Peter J. Russell (1992, Genetics, third edition, Harper-Collins), he discussed the concept of heritability. Page 692 in that chapter is presented below.

a. Criticize particularly Figure 22.14(a). How should it be changed?

b. Criticize the statement: “If the slope is less than 1 but greater than zero, as in Figure 22.14b, both additive genes and nonadditive factors (genes with dominance, genes with epistasis, and environmental factors) affect the phenotypic variation.” How would you change the statement?

c. What must be the meaning of the label for the X axis? Suggest a new wording to make the label clearer.
measure the phenotypes of parents and offspring in a series of families, and then statistically analyze the relationship between their phenotypes. Correlation and regression are appropriate for this type of problem.

We can represent the relationship between offspring phenotype and parental phenotype by plotting the mean phenotype of the parents against the mean phenotype of the offspring, as shown in Figure 22.14. Each point on the graph represents one family. If the points are randomly scattered across the plot, as in Figure 22.14c, then no relationship exists between the traits of parents and offspring. We would conclude that additive genetic differences are not important in determining the phenotypic variability and that the heritability is low. On the other hand, if a definite relationship exists between the phenotypes of parents and offspring, as shown in Figure 22.14a and 22.14b, then additive genetic variance is more important and heritability is high (assuming that no common environmental effects between parents and offspring influence the trait).

In a plot of the parent and offspring phenotypes, the slope of the regression line can provide us with information about the magnitude of the heritability (see section of this chapter on statistics). If the slope is 0, as shown in Figure 22.14c, then the narrow-sense heritability \( h^2 \) is zero. When the slope of the parent-offspring regression is 1, as in Figure 22.14a, the mean offspring phenotype is exactly intermediate to the phenotype of the two parents, and genes with additive effects determine all the phenotypic differences. If the slope is less than 1 but greater than zero, as in Figure 22.14b, both additive genes and nonadditive factors (genes with dominance, genes with epistasis, and environmental factors) affect the phenotypic variation. It is possible to show mathematically that when the mean phenotype of the offspring is regressed against the mean phenotype of the parents, the narrow-sense heritability \( (h^2) \) equals the slope of the regression line \((b)\):

\[
h^2 = b \quad \text{for the regression of mean offspring phenotype and mean parental phenotype}
\]

When the mean phenotype of the offspring is regressed against the phenotype of only one parent, the narrow-sense heritability is twice the slope:

\[
h^2 = 2b \quad \text{for the regression of mean offspring phenotype and one parental phenotype}
\]

Heritability values for a number of traits in different species are given in Table 22.6. These heritability values are based upon various populations and have been determined using a variety of methods, including the parent-offspring method. Estimates of heritability are rarely precise, and most measured heritability values have large errors. This lack of precision is reflected in the fact that heritabilities calculated for the same trait in the same organism often vary widely. Heritability values calculated for human traits must be viewed with special caution, given the difficulties of separating genetic and environmental influences in humans.

**Identification of Genes Influencing a Quantitative Trait**

One of the limitations of the traditional approach of quantitative genetics is that it tells us little about the genes involved in quantitative traits. While heritability may indicate how much of the phenotypic variance in a trait is due to genetic differences, it tells us nothing

---

Figure 22.14

Three hypothetical regressions of mean parental wing length on mean offspring wing length in *Drosophila*. In each case, the slope of the regression line \((b)\) equals the narrow-sense heritability \((h^2)\).