EEB 600A, Problem Set Five

Solutions

1 : Use the properties of variances to show

(a): $Var(ax) = a^2 Var(x)$

$$Var(ax) = Cov(ax, ax) = aCox(x, ax) = a^2 Cov(x, x) = a^2 Var(x)$$

(b): $Var(x + y) = Var(x) + Var(y) + 2Cov(x, y)$

$$\begin{aligned} Var(x + y) &= Cov(x + y, x + y) = Cov(x, x + y) + Cov(y, x + y) \\ &= Cov(x, x) + Cov(x, y) + Cov(y, x) + Cov(y, y) \\ &= Var(x) + 2Cov(x, y) + Var(y) \end{aligned}$$

2 : Data was measured on 50 individuals for arm size (x) and brain size (y), with the following results:

$$\overline{x} = 10, \quad \overline{y} = 50, \quad \sum_{i=1}^{50} (\, x_i - \overline{x}\,)^2 = 100, \quad \sum_{i=1}^{50} (\, y_i - \overline{y}\,)^2 = 400, \quad \sum_{i=1}^{50} (\, x_i - \overline{x}\,)(\, y_i - \overline{y}\,) = 175$$

(a) Compute the variances of x and y, their covariance, and their correlation.

$$\textbf{\textit{Var}}(x) = \frac{100}{49} = 2.04, \quad \textbf{\textit{Var}}(y) = \frac{400}{49} = 8.16, \quad \textbf{\textit{Cov}}(x, y) = \frac{175}{49} = 3.57$$

$$\textbf{\textit{Corr}}(x, y) = \frac{3.57}{\sqrt{2.04} \cdot \sqrt{8.16}} = 0.88$$

(b) What is the best linear regression of arm size on brain size?

$$b_{x\,|\,y} = \frac{3.57}{8.16} = 0.44, \qquad a = \overline{x} - b_{x\,|\,y}\,\overline{y} = 10 - 0.44 \cdot 50 = -11.88$$

Hence, the regression is (Arm size) = -11.88+ 0.44(Brain size)

(c) What is the best linear regression of brain size on arm size?

$$b_{y\,|\,x} = \frac{3.57}{2.04} = 1.75, \qquad a = \overline{y} - b_{y\,|\,x}\,\overline{x} = 50 - 1.75 \cdot 10 = -11.88$$

Hence, the regression is (Brain size) = 32.50 + 1.75(Arm size)

(d) What fraction of the total variance in brain size does the regression account for?

Fraction of the total variance explained by the regression is the squared correlation, or $0.88^2 = 0.766$

3 : Suppose the slope of a midparent-offspring regression is 0.5

(a): What is the expected slope of a single-parent - offspring regression?

Expected slope is 0.25

(b): If the trait variance is 200, what is $Var(A)$?

slope = h^2 ***= Var(A)/Var(z) = Var(A)/200 = 0.5, or Var(A) = 100.***

(c): Suppose that for 10 generations we select midparents whose average value is 5 units larger than the current population mean. What is the expected mean after this selection?

For a single generation, $R = h^2 S = 0.5 * 5 = 2.5$ ***hence, afte 10 generations expect a change in mean of*** $10 \cdot 2.5 = 25$

4 : Suppose the genotypes bb: Bb: BB have genotypic values of 0: (1+h)a : 2a. Then if p = freq(B), then it can be shown (e.g. Lynch and Walsh, Chapter 4) that the genetic variances are

$$\sigma_G^2 = 2p(1-p)\cdot(1+h)\cdot a + p^2 \cdot 2a$$

$$\sigma_A^2 = 2p(1-p)a^2[1+h(2p-1)]^2$$

$$\sigma_D^2 = [\,2p(1-p)ah\,]^2$$

(a): Suppose $h = 0$, so that the genotypes are 0 : a : 2a, a completely additive locus. What are σ_A^2, σ_D^2, and σ_G^2?

$$\sigma_G^2 = \sigma_A^2 = 2p(1-p)a^2, \quad \sigma_D^2 = 0$$

As a function of p, what is the maximal value of σ_A^2?

p = 1/2 (as** $dp(1-p)/dp = 1 - 2p = 0$ **gives p = 1/2.

Are there any values of p for which $\sigma_A^2 = 0$?

p = 0 or 1.

(b): Now suppose $h = 1$, so that the genotypes are 0: 2a: 2a, a completely dominant locus. What are σ_A^2, σ_D^2, and σ_G^2?

$$\sigma_A^2 = 2p(1-p)a^2[4p^2], \quad \sigma_D^2 = [\,2p(1-p)a]^2, \quad \sigma_G^2 = 4p(1-p)a + p^2 \cdot 2a$$

Are there any values of p for which $\sigma_A^2 = 0$?

$p = 0, 1$.

(c): Finally, suppose that the genotypes are 0: 3a: 2a (overdominance). What are σ_A^2, σ_D^2, and σ_G^2?

Here** $h = 2$**, giving

$$\sigma_G^2 = 2p(1-p)\cdot(3)\cdot a + p^2 \cdot 2a, \quad \sigma_A^2 = 2p(1-p)a^2[1+2(2p-1)]^2, \quad \sigma_D^2 = [\,2p(1-p)a2\,]^2$$

Are there any values of p for which $\sigma_A^2 = 0$?

$p = 0, 1$ ***and*** $1 + 2(2p-1) = 0$ ***or*** $4p - 1 = 0$ ***giving p = 1/4***