Section 5.9

Tuesday, November 15, 2016 12:34 PM

Section 5.9

$Y_1 \sim \text{beta}(3, 1)$ and

$$f_{Y_2}(y_2) = \begin{cases} \frac{3}{2}(1 - y_2^2), & 0 < y_2 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

The variance of Y_1 is

$$V(Y_1) = \frac{3(1)}{(3+1+1)(3+1)^2} = \frac{3}{80} \implies \sigma_{Y_1} = \sqrt{\frac{3}{80}} \approx 0.194.$$

Simple calculations using $f_{Y_2}(y_2)$ show that $E(Y_2^2) = 1/5$ and $E(Y_2) = 3/8$ so that

$$V(Y_2) = \frac{1}{5} - \left(\frac{3}{8}\right)^2 = 0.059 \implies \sigma_{Y_2} = \sqrt{0.059} \approx 0.244.$$

Finally, the correlation is

$$\rho = \frac{\text{Cov}(Y_1, Y_2)}{\sigma_{Y_1}\sigma_{Y_2}} \approx \frac{0.01875}{(0.194)(0.244)} \approx 0.40. \quad \square$$

5.9 Expectations and variances of linear functions of random variables

TERMINOLOGY: Suppose that $Y_1, Y_2, ..., Y_n$ are random variables and that $a_1, a_2, ..., a_n$ are constants. The function

$$U = \sum_{i=1}^{n} a_i Y_i = a_1 Y_1 + a_2 Y_2 + \cdots + a_n Y_n$$

is called a **linear combination** of the random variables $Y_1, Y_2, ..., Y_n$.

EXPECTED VALUE OF A LINEAR COMBINATION:

$$E(U) = E\left(\sum_{i=1}^{n} a_i Y_i\right) = \sum_{i=1}^{n} a_i E(Y_i)$$

VARIANCE OF A LINEAR COMBINATION:

$$\begin{aligned} V(U) = V\left(\sum_{i=1}^{n} a_i Y_i\right) &= \sum_{i=1}^{n} a_i^2 V(Y_i) + 2\sum_{i<j} a_i a_j \text{Cov}(Y_i, Y_j) \\ &= \sum_{i=1}^{n} a_i^2 V(Y_i) + \sum_{i \neq j} a_i a_j \text{Cov}(Y_i, Y_j) \end{aligned}$$

$= a^2 + 2ab + b^2$

$(a+b)(a+b) = a^2 + ab + ba + b^2$

$(a+b+c)(a+b+c)$

$= a^2 + b^2 + c^2 + 2ab + 2ac + 2bc$

$(\Sigma a_i Y_i)(\Sigma a_i Y_i)$

$= \Sigma a_i^2 Y_i^2 + \sum_{i \neq j} a_i a_j Y_i Y_j$

$\text{Cov}(Y_i, Y_j) = \text{Cov}(Y_j, Y_i)$

Example 5.17. Achievement tests are commonly seen in educational or employment settings. For a large population of test-takers, let Y_1, Y_2, and Y_3 represent scores for different parts of an exam. Suppose that $Y_1 \sim \mathcal{N}(12, 4)$, $Y_2 \sim \mathcal{N}(16, 9)$, and $Y_3 \sim \mathcal{N}(20, 16)$. Suppose additionally that Y_1 and Y_2 are independent, $\text{Cov}(Y_1, Y_3) = 0.8$, and $\text{Cov}(Y_2, Y_3) = -6.7$. Two different summary measures are computed to assess a subject's performance:

$$U_1 = 0.5Y_1 - 2Y_2 + Y_3 \quad \text{and} \quad U_2 = 3Y_1 - 2Y_2 - Y_3.$$

Find $E(U_1)$ and $V(U_1)$.

SOLUTION: The expected value of U_1 is

$$\begin{aligned} E(U_1) = E(0.5Y_1 - 2Y_2 + Y_3) &= 0.5E(Y_1) - 2E(Y_2) + E(Y_3) \\ &= 0.5(12) - 2(16) + 20 = -6. \end{aligned}$$

The variance of U_1 is

$$\begin{aligned} V(U_1) &= V(0.5Y_1 - 2Y_2 + Y_3) \\ &= (0.5)^2V(Y_1) + (-2)^2V(Y_2) + (1)^2V(Y_3) \\ &\quad + 2(0.5)(-2)\text{Cov}(Y_1, Y_2) + 2(0.5)(1)\text{Cov}(Y_1, Y_3) + 2(-2)(1)\text{Cov}(Y_2, Y_3) \\ &= (0.25)(4) + 4(9) + 16 + 2(0.5)(-2)(0) + 2(0.5)(0.8) + 2(-2)(-6.7) = 80.6. \end{aligned}$$

EXERCISE: Find $E(U_2)$ and $V(U_2)$. □

COVARIANCE BETWEEN TWO LINEAR COMBINATIONS: Suppose that

$$\begin{aligned} U_1 &= \sum_{i=1}^{n} a_iY_i = a_1Y_1 + a_2Y_2 + \cdots + a_nY_n \\ U_2 &= \sum_{j=1}^{m} b_jX_j = b_1X_1 + b_2X_2 + \cdots + b_mX_m. \end{aligned}$$

Then,

$$\text{Cov}(U_1, U_2) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_ib_j\text{Cov}(Y_i, X_j).$$

EXERCISE: In Example 5.17, compute $\text{Cov}(U_1, U_2)$.