**EXERCISES IN STATISTICS**

**Series A, No. 7**

1. Let $x$ and $y$ be jointly distributed random variables with conditional expectations which can be written as $E(y|x) = \alpha + \beta x$ and $E(x|y) = \gamma + \delta y$. Express $\beta$ and $\delta$ in terms of the moments of the joint distributions and show that $\beta \leq 1/\delta$.

**Answer.** We have

$$E(y|x) = \alpha + \beta x \quad \text{with} \quad \beta = \frac{C(x,y)}{V(x)} \quad \text{and} \quad \alpha = E(y) - \beta E(x).$$

Likewise

$$E(x|y) = \gamma + \delta y \quad \text{with} \quad \delta = \frac{C(x,y)}{V(y)} \quad \text{and} \quad \gamma = E(x) - \beta E(y).$$

On forming the product of the two slope coefficients we can invoke the Cauchy–Schwarz Inequality:

$$\beta\delta = \frac{C(x,y)^2}{V(x)V(y)} = \big\{Corr(x,y)\big\}^2 \leq 1.$$

Hence $\beta \leq 1/\delta$.

2. A marksman's scores are a sequence of random variables $y_i$ with $E(y_i) = 90$ and $V(y_i) = 16$ for all $i$. The correlation between successive scores is 0.9, and the expectation of a score conditional upon the previous score is given by $E(y_i|y_{i-1}) = \alpha + \beta y_{i-1}$ where $\alpha = (1-\beta)E(y_i)$. Find the expected score given that the previous score was 80.

**Answer.** Substituting for $\alpha = (1-\beta)E(y_i)$. gives

$$\begin{aligned} E(y_i|y_{i-1}) &= \alpha + \beta y_{i-1} = (1-\beta)E(y_i) + \beta y_{i-1} \\ &= E(y_i) + \beta\big\{y_{i-1} - E(y_i)\big\} \end{aligned}$$

Also, since $V(y_i) = V(y_{i-1})$, we have

$$Corr(y, y_{i-1}) = \frac{C(y_i, y_{i-1})}{\sqrt{V(y_i)V(y_{i-1})}} = \frac{C(y_i, y_{i-1})}{V(y_{i-1})} = \beta.$$

With $E(y_i) = E(y_{i-1}) = 90$, $Corr(y_i, y_{i-1}) = 9/10$, and $y_{i-1} = 80$, this gives

$$E(y_i|y_{i-1}) = 90 + \frac{9}{10}(80 - 90) = 81.$$

3. The expected rainfalls in September, October and November are 10 ins, 8 ins 6 ins respectively, with a variance–covariance matrix of

$$\begin{bmatrix} 6 & -3 & 1.5 \\ -3 & 6 & -3 \\ 1.5 & -3 & 6 \end{bmatrix}.$$

Calculate the expected rainfall throughout these three months and find its variance.

If the September rain was unusually high, in what direction would ones estimates of rainfall be revised

(a) in the two months following, and

(b) for the three month period?

**Answer.** The Expected rainfall is

$$\begin{aligned} E(y) &= E\left(\sum x_i\right) = \sum E(x_i) \\ &= 10 + 8 + 6 = 24. \end{aligned}$$

Its variance is

$$\begin{aligned} V(x_1 + x_2 + x_3) &= V(x_1) + V(x_2) + V(x_3) \\ &\qquad + 2\big\{C(x_1, x_2) + C(x_1, x_3) + C(x_2, x_3)\big\} \\ &= (6 + 6 + 6) + 2\{1.5 - 3 - 3\} = 9. \end{aligned}$$

Let $x$, $y$, $z$ be the rainfall in September, October and November, and let us write $x = E(x) + \Delta x$. Then we have

$$E(x + y + z|z) = x + E(y|x) + E(z|x),$$

with

$$\begin{aligned} E(y|x) &= E(y) + \beta\{x - E(x)\}; \quad \beta = C(x, y)/V(x); \\ E(z|x) &= E(z) + \delta\{x - E(x)\}; \quad \delta = C(x, z)/V(x), \end{aligned}$$

which may be combined to give

$$\begin{aligned} E(x + y + z|z) &= x + E(y) + E(z) + \{\beta + \delta\}\{x - E(x)\} \\ &= E(x) + E(y) + E(z) + \{1 + \beta + \delta\}\Delta x. \end{aligned}$$

Given that

$$\beta = \frac{C(x,y)}{V(x)} = -\frac{3}{6} = -\frac{1}{2} \quad \text{and} \quad \delta = \frac{C(x,z)}{V(z)} = \frac{1.5}{6} = \frac{1}{4},$$

It follows that

$$E(x+y+z|z) = E(x) + E(y) + E(z) + \left\{1 - \frac{1}{2} + \frac{1}{4}\right\}\Delta x.$$

4. A man runs on Hampstead Heath twice a week. The average duration of one of his outings is 40 minutes with a standard deviation of 5 minutes. His average running time per week is 1 hour 20 minutes with a standard deviation of 4 minutes. Given that he ran for only 15 minutes on Monday, what is the expected duration of his Friday outing?

**Answer.** The duration of the Monday outing is $x$, the duration of the Friday outing is is $y$. We have

$$E(x) = E(y) = 40, \quad V(x) = V(y) = 25, \quad \text{and}$$
$$V(x+y) = V(x) + V(y) + 2C(x,y) = 16.$$

Hence

$$C(x,y) = \frac{16 - 25 - 25}{2} = -17, \quad \text{and} \quad \beta = \frac{C(x,y)}{V(x)} = -\frac{17}{25}.$$

If $x = 15$, then

$$\begin{aligned} E(y|x) &= E(y) + \beta\{x - E(x)\} \\ &= 40 - \frac{17}{25}\{15 - 40\} = 57. \end{aligned}$$

5. An investor has a choice of three financial assets. The expected yields of these assets are given in the vector $[\,0.04 \quad 0.03 \quad 0.05\,]$ and the variances and covariances of the yields are given in the matrix

$$10^{-4}\begin{bmatrix} 3.29 & -0.83 & 0 \\ -0.83 & 3.41 & 0 \\ 0 & 0 & 2.01 \end{bmatrix}.$$

Derive expressions for the expected yield and variance of a portfolio containing $\lambda Q$ of the first asset and $(1-\lambda)Q$ of the second asset, and ascertain whether there is any value of $\lambda$ such that the variance of this portfolio is less than the variance of an investment $Q$ in the third asset.

**Answer.** The investor has two objectives (a) to obtain high returns and (b) to minimise risk. A trade-off must be established between the two objectives. However, some portfolios dominate others in the sense that they fulfil one of the objectives better that the alternative portfolio whilst fulfilling the other objective at least as well. The question we must answer is whether or not there exists a mixed portfolio containing the first and the second assets which dominates a portfolio containing only the third asset.

First we find the expected yield and variance of a generic portfolio containing $\lambda Q$ of the first asset and $(1-\lambda)Q$ of the second asset. Let $x_1$, $x_2$ and $x_3$ be the yields of the three assets. We find that

$$\begin{aligned} E\{\lambda Q x_1 + (1-\lambda)Q x_2\} &= Q\{\lambda E(x_1) + (1-\lambda)E(x_2)\} \\ &= Q\Big\{E(x_2) + \lambda\big[E(x_1) - E(x_2)\big]\Big\} \\ &= Q\{0.03 + \lambda 0.01\}. \end{aligned}$$

Also

$$\begin{aligned} &V\{\lambda Q x_1 + (1-\lambda)Q x_2\} \\ &\quad = Q^2\Big\{\lambda^2 V(x_1) + (1-\lambda)^2 V(x_2) + 2\lambda(1-\lambda)C(x_1,x_2)\Big\} \\ &\quad = Q^2\Big[\lambda^2\big\{V(x_1) + V(x_2) - 2C(x_1,x_2)\big\} \\ &\qquad\qquad + \lambda\big\{2C(x_1,x_2) - 2V(x_2)\big\} + V(x_2)\Big] \\ &\quad = Q^2\left[\lambda^2\left\{\begin{matrix} 3.29 \\ +3.41 \\ +1.66 \end{matrix}\right\} - \lambda\left\{\begin{matrix} 1.66 \\ +6.82 \end{matrix}\right\} + 3.41\right] \\ &\quad = Q^2(8.36\lambda^2 - 8.48\lambda + 3.41). \end{aligned}$$

Now we ask whether there is a value $\lambda \in (0,1)$ such that this variance is less that $V(Qx_3) = Q^2 2.01$. The answer can be found by investigating the roots of

$$V\big\{\lambda x_1 + (1-\lambda)x_2\big\} - V(x_3) = 8.36\lambda^2 - 8.48\lambda + 1.40 = 0.$$

The roots $\lambda_1 = 0.2075$ and $\lambda_2 = 0.8009$ are the points at which the alternative portfolios have the same variance. The optimum value of $\lambda$ which minimises the variance of the mixed portfolio, and which can be obtained by ordinary calculus, lies somewhere in the interval $[\lambda_1, \lambda_2]$