EXERCISES IN STATISTICS

Series A, No. 6

1. Let x_1 and x_2 have the joint p.d.f $f(x_1, x_2) = x_1 + x_2$ with $0 \le x_1, x_2 \le 1$. Find the conditional mean of x_2 given x_1.

Answer. First let us show that

$$\int_{x_2}\int_{x_1} f(x_1, x_2)dx_2dx_1 = 1 :$$

We have

$$\begin{aligned}\int_{x_2}\int_{x_1}(x_1 + x_2)dx_1dx_2 &= \int_{x_2}\left[\frac{x_1^2}{2} + x_2x_1\right]_0^1 dx_2 \\ &= \int_{x_2}\left(\frac{1}{2} + x_2\right)dx_2 = \left[\frac{x_2}{2} + \frac{x_2^2}{2}\right]_0^1 = 1.\end{aligned}$$

From this we can see that $f(x_2) = [\frac{1}{2} + x_2]$. Likewise, by an argument of symmetry, we have $f(x_1) = [\frac{1}{2} + x_1]$.

Next consider

$$\begin{aligned}E(x_2|x_1) &= \int_{x_2} x_2 \frac{f(x_1, x_2)}{f(x_1)} dx_2 \\ &= \frac{1}{(\frac{1}{2} + x_1)}\int_{x_2} x_2(x_1 + x_2)dx_2 = \frac{1}{(\frac{1}{2} + x_1)}\left[\frac{x_1x_2^2}{2} + \frac{x_2^3}{3}\right]_{x_2=0}^1 \\ &= \frac{1}{(x_1 + \frac{1}{2})}\frac{(3x_1 + 2)}{6} = \frac{3x_1 + 2}{6x_1 + 3}.\end{aligned}$$

2. Find the $P(x < y|x < 2y)$ when $f(x, y) = e^{-(x+y)}$. Draw a diagram to represent the event.

Answer.

$$\begin{aligned}P(x < ay) &= \int_{y=0}^{\infty}\int_{x=0}^{ay} f(x, y)dxdy = \int_{y=0}^{\infty} e^{-y}\left\{\int_{x=0}^{ay} e^{-x}\right\}dy \\ &= \int_{y=0}^{\infty} e^{-y}\left[-e^{-x}\right]_0^{ay} dy = \int_{y=0}^{\infty} e^{-y}\left[1 - e^{-ay}\right]dy \\ &= \left[-e^{-y} + \frac{e^{-(1+a)y}}{1 + a}\right]_0^{\infty} = 1 - \frac{1}{1 + a}\end{aligned}$$

Hence we get

$$P(x < y) = 1 - \frac{1}{2} = \frac{1}{2}$$

$$P(x < 2y) = 1 - \frac{1}{3} = \frac{2}{3}$$

$$P(x < y | x < 2y) = \frac{P(x < y)}{P(x < 2y)} = \frac{1}{2}.\frac{3}{2} = \frac{3}{4}.$$

3. The variance of $x_1 + x_2$ is $V(x_1) + V(x_2) + 2C(x_1, x_2)$. Extend this result to find the variance of $x_1 + x_2 + x_3$.

Answer.

$$\begin{aligned} V(x_1 + x_2 + x_3) &= V\{(x_1 + x_2) + x_3\} \\ &= V(x_1 + x_2) + V(x_3) + 2C(x_1 + x_2, x_3) \\ &= V(x_1) + V(x_2) + V(x_3) + 2C(x_1 + x_2, x_3) + 2C(x_1, x_2). \end{aligned}$$

But

$$\begin{aligned} C(x_1 + x_2, x_3) &= E\Big[\big\{(x_1 + x_2) - E(x_1 + x_2)\big\}\big\{x_3 - E(x_3)\big\}\Big] \\ &= E\Big[\big\{x_1 - E(x_1)\big\} + \big\{x_2 - E(x_2)\big\}\Big]\Big[x_3 - E(x_3)\Big] \\ &= E\Big[\big\{x_1 - E(x_1)\big\}\big\{x_3 - E(x_3)\big\}\Big] \\ &\qquad\qquad + E\Big[\big\{x_2 - E(x_2)\big\}\big\{x_3 - E(x_3)\big\}\Big] \\ &= C(x_1, x_3) + C(x_2, x_3) \end{aligned}$$

Substituting this result into the expression for $V(x_1 + x_2 + x_3)$ gives

$$V(x_1+x_2+x_3) = V(x_1)+V(x_2)+V(x_3)+2C(x_1, x_2)+2C(x_1, x_3)+2C(x_2, x_3).$$

4. Find $V(x_1 - x_2)$. Prove that

$$C(x_1, x_2) \leq \sqrt{V(x_1)V(x_2)} \leq \frac{1}{2}\{V(x_1) + V(x_2)\}.$$

Answer. For the first part,

$$\begin{aligned} V(x_1 - x_2) &= E\big\{(x_1 - x_2) - E(x_1 - x_2)\big\}^2 \\ &= E\Big[\big\{x_1 - E(x_1)\big\} - \big\{x_2 - E(x_2)\big\}\Big]^2 \\ &= E\big\{x_1 - E(x_1)\big\}^2 - 2E\Big[\big\{x_1 - E(x_1)\big\}\big\{x_2 - E(x_2)\big\}\Big] \\ &\qquad + E\big\{x_2 - E(x_2)\big\}^2 \\ &= V(x_1) + V(x_2) - 2C(x_1, x_2) \end{aligned}$$

For the second part, we may assume that

$$-1 \leq \rho = \frac{C(x_1, x_2)}{\sqrt{V(x_1)V(x_2)}} \leq 1.$$

It follows that

$$C(x_1, x_2) \leq \sqrt{V(x_1) + V(x_2)}$$

Also we have

$$0 \leq \left(\sqrt{V(x_1)} - \sqrt{V(x_2)}\right)^2 = V(x_1) - 2\sqrt{V(x_1)}\sqrt{V(x_2)} + V(x_2)$$

whence

$$\frac{1}{2}\{V(x_1) + V(x_2)\} \geq \sqrt{V(x_1) + V(x_2)}$$

and the result follows from combining the inequalities.