NATIONAL UNIVERSITY OF SINGAPORE
DEPARTMENT OF STATISTICS & APPLIED PROBABILITY

ST2334 PROBABILITY AND STATISTICS

SEMESTER I, AY 2022/2023

Tutorial 06: Solution

This set of questions will be discussed by your tutors during the tutorial in Week 9.

Please work on the questions before attending the tutorial.

1. Suppose that X and Y are random variables having the joint probability function below.

$f(x,y)$		$x=2$	$x=4$
y	1	0.10	0.15
	3	0.20	0.30
	5	0.10	0.15

(a) Determine whether X and Y are independent.

(b) Find $E(Y|X=2)$.

(c) Find $E(X|Y=3)$.

(d) Find $E(2X-3Y)$.

(e) Find $E(XY)$.

(f) Find $V(X)$ and $V(Y)$.

SOLUTION

(a) The marginal distributions are added to the table.

$f(x,y)$		$x=2$	$x=4$	$f_Y(y)$
y	1	0.10	0.15	0.25
	3	0.20	0.30	0.50
	5	0.10	0.15	0.25
$f_X(x)$		0.40	0.60	1

It can be verified that

$$f(x,y) = f_X(x)f_Y(y) \quad \text{for } x = 2,4 \text{ and } y = 1,3,5.$$

Hence, X and Y are independent.

(b) The conditional probability function $f_{Y|2}(y|2)$ is given by

y	1	3	5
$f_{Y\|X}(y\|2)$	$0.10/0.40 = 1/4$	$0.20/0.40 = 2/4$	$0.10/0.40 = 1/4$

Thus $E(Y|X=2) = 1(1/4)+3(2/4)+5(1/4) = 3$.
Alternatively, since X and Y are independent, $E(Y|X=2) = E(Y) = 1(0.25)+3(0.5)+5(0.25) = 3$.

(c) The conditional probability function $f_{X|Y}(x|3)$ is given by

x	2	4
$f_{X\mid Y}(x\mid 3)$	$0.20/0.50 = 2/5$	$0.30/0.50 = 3/5$

Thus $E(X|Y=3) = 2(2/5)+4(3/5) = 16/5 = 3.2$.
Alternatively, since X and Y are independent, $E(X|Y=3) = E(X) = 2(0.4)+4(0.6) = 3.2$.

(d) $E(X) = 3.2$ and $E(Y) = 3$; therefore $E(2X-3Y) = 2E(X)-3E(Y) = 2(3.2)-3(3) = -2.6$.

(e) $E(XY) = (2)(1)(0.10)+(2)(3)(0.10)+(2)(5)(0.10)+(4)(1)(0.10)+(4)(3)(0.10)+(4)(5)(0.10) = 9.6$.
Alternatively, since X and Y are independent, we have $E(XY) = E(X)E(Y) = (3.2)(3) = 9.6$.

(f) $E(X^2) = 2^2(0.4)+4^2(0.6) = 11.2$; thus $V(X) = E(X^2)-[E(X)]^2 = 11.2-(3.2)^2 = 0.96$.
$E(Y^2) = 1^2(0.25)+3^2(0.5)+5^2(0.25) = 11$; thus $V(Y) = E(Y^2)-[E(Y)]^2 = 11-3^2 = 2$.

2. A service facility operates with two service lines. On a randomly selected day, let X be the proportion of time that the first line is in use whereas Y is the proportion of time that the second line is in use. Suppose that the joint probability density function for (X,Y) is given below.

$$f(x,y) = \begin{cases} \frac{3}{2}(x^2+y^2), & 0 \le x \le 1, 0 \le y \le 1; \\ 0, & \text{elsewhere} \end{cases}$$

(a) Determine whether X and Y are independent.

(b) Find the mean and variance of X and Y.

(c) Find the covariance of X and Y.

(d) Find the mean and variance of $X+Y$.

SOLUTION

The marginal distributions for X and Y are

$$\begin{aligned} f_X(x) &= \int_0^1 \frac{3}{2}(x^2+y^2)dy = \frac{3}{2}\left(x^2y+\frac{y^3}{3}\right)\Big|_{y=0}^1 = \frac{3}{2}\left(x^2+\frac{1}{3}\right), \quad \text{for } 0 \le x \le 1; \\ f_Y(y) &= \int_0^1 \frac{3}{2}(x^2+y^2)dx = \frac{3}{2}\left(\frac{x^3}{3}+xy^2\right)\Big|_{x=0}^1 = \frac{3}{2}\left(\frac{1}{3}+y^2\right), \quad \text{for } 0 \le y \le 1. \end{aligned}$$

(a) X and Y are dependent, since $f(x,y) \ne f_X(x)f_Y(y)$, or $f(x,y)$ can not be factorized as $Cg_1(x)g_2(y)$.

(b) For X,

$$\begin{aligned} E(X) &= \frac{3}{2}\int_0^1 x\left(x^2+\frac{1}{3}\right)dx = \frac{3}{2}\left(\frac{x^4}{4}+\frac{x^2}{6}\right)\Big|_0^1 = \frac{5}{8}; \\ E(X^2) &= \frac{3}{2}\int_0^1 x^2\left(x^2+\frac{1}{3}\right)dx = \frac{3}{2}\left(\frac{x^5}{5}+\frac{x^3}{9}\right)\Big|_0^1 = \frac{7}{15}; \\ V(X) &= E(X^2)-[E(X)]^2 = \frac{73}{960}; \end{aligned}$$

For Y, method 1: we can repeat the work in Part (b),

$$\begin{aligned} E(Y) &= \frac{3}{2}\int_0^1 y\left(y^2+\frac{1}{3}\right)dy = \frac{3}{2}\left(\frac{y^4}{4}+\frac{y^2}{6}\right)\Big|_0^1 = \frac{5}{8}; \\ E(Y^2) &= \frac{3}{2}\int_0^1 y^2\left(y^2+\frac{1}{3}\right)dy = \frac{3}{2}\left(\frac{y^5}{5}+\frac{y^3}{9}\right)\Big|_0^1 = \frac{7}{15}; \\ V(Y) &= E(Y^2)-[E(Y)]^2 = 73/960. \end{aligned}$$

Method 2: observe that the joint p.d.f. is symmetric in x and y, i.e., $f(x,y) = f(y,x)$; so $E(Y) = E(X) = 5/8$; $V(Y) = V(X) = 73/960$.

(c)

$$\begin{aligned} E(XY) &= \frac{3}{2}\int_0^1\int_0^1 xy(x^2+y^2)dxdy = \frac{3}{2}\int_0^1\int_0^1 (x^3y+xy^3)dxdy \\ &= \frac{3}{2}\int_0^1 \left(\frac{x^4y}{4}+\frac{x^2y^3}{2}\right)\Big|_{x=0}^1 dy = \frac{3}{2}\int_0^1\left(\frac{y}{4}+\frac{y^3}{2}\right)dy \\ &= \frac{3}{2}\left(\frac{y^2}{8}+\frac{y^4}{8}\right)\Big|_0^1 = \frac{3}{8}. \end{aligned}$$

Thus

$$Cov(X,Y) = E(XY)-E(X)E(Y) = \frac{3}{8}-\frac{5}{8}\cdot\frac{5}{8} = -1/64.$$

(d) $E(X+Y) = E(X)+E(Y) = \dfrac{5}{8}+\dfrac{5}{8} = \dfrac{5}{4}$.

$V(X+Y) = V(X)+V(Y)+2Cov(X,Y) = \dfrac{73}{960}+\dfrac{73}{960}+2\cdot\dfrac{-1}{64} = \dfrac{29}{240}$.

3. The random variables X and Y have the joint probability density function given by

$$f(x,y) = \begin{cases} x+y, & 0\le x\le 1, 0\le y\le 1; \\ 0, & \text{elsewhere} \end{cases}$$

Find

(a) $Cov(X,Y)$;

(b) $E(Y|X=0.2)$;

(c) $E(X|Y=0.5)$.

SOLUTION

(a) We need to compute $E(X)$, $E(Y)$, and $E(XY)$.

$$\begin{aligned} f_X(x) &= \int_0^1 (x+y)dy = \left(xy+\frac{y^2}{2}\right)\Big|_{y=0}^1 = x+\frac{1}{2}, \quad \text{for } 0\le x\le 1. \\ E(X) &= \int_0^1 x\left(x+\frac{1}{2}\right)dx = \left(\frac{x^3}{3}+\frac{x^2}{4}\right)\Big|_0^1 = \frac{7}{12}. \\ f_Y(y) &= \int_0^1 (x+y)dx = \left(\frac{x^2}{2}+xy\right)\Big|_{x=0}^1 = y+\frac{1}{2}, \quad \text{for } 0\le y\le 1; \\ E(Y) &= \int_0^1 y\left(y+\frac{1}{2}\right)dy = \left(\frac{y^3}{3}+\frac{y^2}{4}\right)\Big|_0^1 = \frac{7}{12}. \end{aligned}$$

$$\begin{aligned} E(XY) &= \int_0^1 \int_0^1 xy(x+y)dxdy = \int_0^1 \int_0^1 (x^2y+xy^2)dxdy = \int_0^1 \left(\frac{x^3y}{3}+\frac{x^2y^2}{2}\right)\Big|_{x=0}^1 dy \\ &= \int_0^1 \left(\frac{y}{3}+\frac{y^2}{2}\right)dy = \frac{1}{3}. \end{aligned}$$

Therefore

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 1/3 - (7/12)(7/12) = -1/144.$$

(b) To compute $E(Y|X=0.2)$, we need to obtain $f_{Y|X}(y|0.2)$ first. For $0 \le y \le 1$,

$$f_{Y|X}(y|x=0.2) = \frac{f(0.2,y)}{f_X(0.2)} = \frac{0.2+y}{0.2+0.5} = \frac{2+10y}{7},$$

and $f_{Y|X}(y|0.2) = 0$ for $y \notin [0,1]$. Therefore

$$E(Y|X=0.2) = \int_0^1 y\left(\frac{2+10y}{7}\right)dy = \frac{1}{7}\left(y^2+\frac{10y^3}{3}\right)\Big|_0^1 = \frac{13}{21}.$$

(c) Similar in strategy to Part (b), for $0 \le x \le 1$;

$$f_{X|Y}(x|y=0.5) = \frac{f(x,0.5)}{f_Y(0.5)} = \frac{x+0.5}{0.5+0.5} = x+0.5;$$

and $f_{X|Y}(x|0.5) = 0$ elsewhere.

$$E(X|Y=0.5) = \int_0^1 x(x+0.5)dx = \left(\frac{x^3}{3}+\frac{x^2}{4}\right)\Big|_0^1 = \frac{7}{12}.$$

4. Given that $V(X) = 5$ and $V(Y) = 3$, and define $Z = -2X + 4Y - 3$.

(a) Find $V(Z)$ if X and Y are independent.

(b) Find $V(Z)$ if $Cov(X,Y) = 1$.

SOLUTION

In general,

$$\begin{aligned} V(Z) &= V(-2X+4Y-3) = V(-2X+4Y) = (-2)^2V(X) + 4^2V(Y) + 2(-2)(4)Cov(X,Y) \\ &= 4V(X) + 16V(Y) - 16Cov(X,Y) = 20 + 48 - 16Cov(X,Y) = 68 - 16Cov(X,Y). \end{aligned}$$

(a) If X and Y are independent, so that $Cov(X,Y) = 0$, we have $V(Z) = 68$.

(b) If $Cov(X,Y) = 1$, we have $V(Z) = 68 - 16 = 52$.

5. An employee is selected from a staff of 10 to supervise a certain project by selecting a tag at random from a box containing 10 tags numbered from 1 to 10.

(a) Find the formula for the probability distribution of X representing the number on the tag that is drawn.

(b) What is the probability that the number drawn is less than 4?

(c) Find the mean and variance of X.

SOLUTION

(a) X follows a discrete uniform distribution; p.m.f. is given by

$$f_X(x) = \begin{cases} \dfrac{1}{10}, & x = 1, 2, \ldots, 10 \\ 0, & \text{elsewhere} \end{cases}$$

(b) $P(X < 4) = f_X(1) + f_X(2) + f_X(3) = 3/10 = 0.3$.

(c) $\mu_X = \displaystyle\sum_{x=1}^{10} x\left(\frac{1}{10}\right) = 5.5$.

$\sigma_X^2 = \displaystyle\sum_{x=1}^{10} (x - 5.5)^2 \frac{1}{10} = 8.25$.

6. According to Chemical Engineering Progress (Nov, 1990), approximately 30% of all pipework failures in chemical plants are caused by operator error.

   (a) What is the probability that out of the next 20 pipework failures at least 10 are due to operator error?

   (b) What is the probability that no more than 4 out of 20 such failures are due to operator error?

   (c) What is the probability that for out of 20 such failures, exactly 5 are operational errors.

   SOLUTION
   Let X =number of pipework failures caused by operator error out of 20 pipework. Then $X \sim \text{Binomial}(20, 0.3)$.

   (a) $P(X \geq 10) = 1 - P(X < 9) = 1 - 0.9520 = 0.0480$.

   (b) $P(X \leq 4) = 0.2375$.

   (c) $P(X = 5) = 0.1789$.

7. In testing a certain kind of truck tire over a rugged terrain, it is found that 25% of the trucks fail to complete the test without a blowout. Of the next 15 trucks tested, find

   (a) The probability of zero blowouts.

   (b) The probability of at least 8 blowouts.

   (c) Expected number of blowouts; variance of number of blowouts.

   SOLUTION
   X =number of trucks out of 15 trucks with blowout. $X \sim \text{Binomial}(15, 0.25)$.

   (a) $P(X = 0) = 0.0134$;

   (b) $P(X \geq 8) = 1 - 0.9824 = 0.0173$;

   (c) $E(X) = np = (15)(0.25) = 3.75$;
   $V(X) = np(1 - p) = (15)(0.25)(0.75) = 2.8125$.