ECE 316-Solutions of Problem Set 5

Solution 1

Let X denote the r.v. of the number of 1's and Y denote the r.v of the number of 2's in n rolls of the die. The probability that a 1 occurs (or a 2 occurs) in a roll of the die is $\frac{1}{6}$.

Therefore:

$$P(X=k)=P(Y=k)=\binom{n}{k}\left(\frac{1}{6}\right)^{k}\left(\frac{5}{6}\right)^{n-k}, \quad k=0,1,2,\ldots,n$$

The joint distribution:

$$P(X=k,Y=j)=\binom{n}{j,k}\left(\frac{1}{6}\right)^{j+k}\left(\frac{4}{6}\right)^{n-j-k}, \quad j\geq 0, k\geq 0,\ j+k\leq n$$

where $\binom{n}{j,k}=\frac{n!}{j!k!(n-k-j)!}$ (multinomial coefficient (see page 11 of Chapter 1 in your text). This follows from the fact that if there are k 1's, j 2's then the remaining n-j-k must be some other number i.e. (3,4,5 or 6).

You can check that $\sum_{j=0}^{n-k} P(X=k,Y=j)=P(X=k)$

Now $E[X]=E[Y]=n\frac{1}{6}=\frac{n}{6}$ from the mean of the $B(n,p)$ distribution.

So:

$$cov(X,Y)=E[XY]-E[X]E[Y]=\sum_{i,j:i+j\leq n} ijP(X=i,Y=j)-\left(\frac{n}{36}\right)^2$$

Let us consider the first term on the rhs:

$$\begin{aligned}
E[XY] &= \sum_{i,j:i+j\leq n} ij\binom{n}{i,j,n-i-j}\left(\frac{1}{6}\right)^{i}\left(\frac{1}{6}\right)^{j}\left(\frac{4}{6}\right)^{n-j-i}\\
&= \sum_{i,j:i+j\leq n} ij\frac{n!}{i!j!(n-i-j)!}\left(\frac{1}{6}\right)^{i}\left(\frac{1}{6}\right)^{j}\left(\frac{4}{6}\right)^{n-j-i}\\
&= \sum_{i=1}^{n}\sum_{j=1}^{n-i}\frac{n!}{(i-1)!(j-1)!(n-i-j)!}\left(\frac{1}{6}\right)^{i}\left(\frac{1}{6}\right)^{j}\left(\frac{4}{6}\right)^{n-j-i}\\
&\quad \textit{substitute } i-1=p \textit{ and } j-1=q \textit{ in the summation}\\
&= n(n-1)\sum_{p=0}^{n-2}\sum_{q=0}^{n-2-i}\frac{(n-2)!}{p!q!(n-2-p-q)!}\left(\frac{1}{6}\right)^{p+1}\left(\frac{1}{6}\right)^{q+1}\left(\frac{4}{6}\right)^{n-2-p-q}\\
&= n(n-1)\left(\frac{1}{6}\right)^{2}\left(\frac{1}{6}+\frac{1}{6}+\frac{4}{6}\right)^{n-2}\\
&= n(n-1)\frac{1}{36}
\end{aligned}$$

Hence

$$cov(X,Y)=n(n-1)\frac{1}{36}-\frac{n^2}{36}=\frac{-n}{36}<0$$

It is important to note that the covariance can be negative but the variance is always non-negative. This negative correlation is easy to see: if a 1 occurs in one throw then a 2 cannot occur (they are negatively correlated!).

Problem 2

At first it would seem that $f_{XY}(x,y)$ does not depend on y. But if we see the domain of definition for y we see that we can write:

$$f_{XY}(x,y) = 4e^{-2x}\mathbf{1}_{[0\le y\le x]}, \quad x,y \in [0,\infty)$$

so indeed the joint density depends on both.

Note the constant should be 4 and not 2 as stated otherwise the integral $\int_0^\infty \int_0^\infty f_{XY}(x,y)$ will not be 1.

Throughout, for calculations we use the identity:

$$\int_0^\infty x^n e^{-ax}dx = \frac{n!}{a^{n+1}}, \quad n = 0,1,2,$$

a) To compute $f_{X|Y}(x|y)$ we need to compute $f_Y(y) = \int_0^\infty f_{XY}(x,y)dx$

Now $f_Y(y) = \int_0^\infty 4e^{-2x}\mathbf{1}_{[y\le x]}dx = \int_y^\infty 4e^{-2x}dx = 2e^{-2y}$ and then

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f_{XY}(x,y)}{f_Y(y)} \\ &= \frac{4e^{-2x}\mathbf{1}_{[y\le x]}}{2e^{-2y}} \\ &= 2e^{-2(x-y)}\mathbf{1}_{[y\le x]} \end{aligned}$$

b) To compute $cov(X,Y)$ we need to compute $[X]$, $E[Y]$ and $E[XY]$ since

$$cov(X,Y) = E[XY] - E[X]E[Y]$$

Now

$$E[Y] = 2\int_0^\infty ye^{-2y}dy = \frac{2}{4} = \frac{1}{2}$$

Similarly

$$f_X(x) = \int_0^\infty f_{XY}(x,y)dy = 4e^{-2x}\int_0^x 1.dy = 4xe^{-2x}$$

and hence:

$$E[X] = \int_0^\infty xf_X(x)dx = 4\int_0^\infty x^2e^{-2x}dx = \frac{8}{8} = 1$$

Now

$$E[XY] = \int_0^\infty \int_0^\infty xyf_{XY}(x,y)dxdy = 2\int_0^\infty x^3e^{-2x}dx = \frac{12}{2^4} = \frac{3}{4}$$

So $cov(X,Y) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$

Problem 3

$$
\begin{aligned}
f_X(z) &= \frac{1}{\sqrt{2\pi}}\exp(-\frac{z^2}{2}) \qquad -\infty < z < \infty \\
E(Z) = E[X\mathbf{1}_{[X>x]}] &= \int_x^\infty z\frac{1}{\sqrt{2\pi}}\exp(-\frac{z^2}{2})dz \\
&= -\frac{1}{\sqrt{2\pi}}\exp(-\frac{z^2}{2})|_x^\infty \\
&= \frac{1}{\sqrt{2\pi}}\exp(-\frac{x^2}{2})
\end{aligned}
$$

Problem 4

This is an important result.

First note by definition:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = h(x)\int_{-\infty}^{\infty} g(y)dy = C_1h(x)$$

where $C_1 = \int_{-\infty}^{\infty} g(y)dy$

Similarly, we obtain

$$f_Y(y) = C_2g(y)$$

where $C_2 = \int_{-\infty}^{\infty} h(x)dx$.

On the other hand we know

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = \int_{-\infty}^{\infty} h(x)dx\int_{-\infty}^{\infty} g(y)dy = C_2C_1$$

Therefore $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ or the random variables are independent.

Now consider the joint density given in the example, then it can be written as:

$$f_{X,Y}(x,y) = Cxy\mathbf{1}_{[0<x<1]}\mathbf{1}_{[0<y<1]}\mathbf{1}_{[0<x+y<1]}, \quad -\infty < x < \infty;\ -\infty < y < \infty$$

which clearly cannot be written as $h(x)g(y)$ for some functions $h(.)$ and $g(.)$ because the last indicator function introduces dependency because it involves both x and y.

Therefore $f_{X,Y}(x,y) \neq f_X(x)f_Y(y)$ and so the random variables cannot be independent.

Problem 5

Instead of solving this problem for three random variables we can solve it for the case of n independent random variables.

$X_1, \cdots X_n$ are i.i.d r.v having uniform distribution over $(0,1)$ Let $Y = \max(X_1, \cdots X_n)$, $Z =$

$\min(X_1, \cdots X_n)$

$$P(Y \leq y) = P(X_1 \leq y, X_2 \leq y, \ldots, X_n \leq y) = P(X_1 \leq y)P(X_2 \leq y) \cdots P(X_n \leq y) \quad \textit{by independence}$$
$$= \begin{cases} 0, & y < 0 \\ y^n, & 0 < y < 1 \\ 1, & y > 1 \end{cases}$$

$$f_Y(y) = ny^{n-1} 0 < y < 1$$

$$\begin{aligned} P(Z \leq z) &= 1 - P(Z > z) \\ &= 1 - P(X_1 > z)P(X_2 > z) \cdots P(X_n > z) \\ &= \begin{cases} 1, & y < 0 \\ 1 - (1-z)^n, & 0 < z < 1 \\ 0, & z > 1 \end{cases} \end{aligned}$$

$$f_Z(z) = n(1-z)^{n-1} \qquad 0 < z < 1$$

$$E(Y) = \frac{n}{n+1}$$

$$E(Z) = \frac{1}{n+1}$$

Of course since Y and Z are non-negative random variables we can directly integrate the complementary distributions $P(Y > x)$ and $P(Z > x)$ instead of calculating the p.d.f.'s. In this case this is an easier approach.

Remark:

We see that as $n \to \infty$ $E[Y] \to 1$ and $E[Z] \to 0$ i.e. as the number of random variables increases the largest value we obtain is closer and closer to 1 the maximum a given r.v can be and the minimum goes towards 0 which is the minimum value a given r.v. can have. This confirms the intuition that the more number random variables we observe the larger the maximum value we are going to observe and the smaller the value of the minimum we will observe.

Problem 6

$$\begin{aligned}
E(X) &= \int_{y=0}^{\infty}\int_{x=0}^{\infty} x\frac{1}{y}\exp(-(y+x/y))dxdy \\
&= \int_{y=0}^{\infty}\frac{1}{y}\exp(-y)\int_{x=0}^{\infty} x\exp(-x/y)dxdy \\
&= \int_{y=0}^{\infty}\frac{1}{y}\exp(-y)y^2dy \\
&= 1
\end{aligned}$$

$$\begin{aligned}
E(Y) &= \int_{y=0}^{\infty}\int_{x=0}^{\infty} y\frac{1}{y}\exp(-(y+x/y))dxdy \\
&= 1
\end{aligned}$$

$$\begin{aligned}
E(XY) &= \int_{y=0}^{\infty}\int_{x=0}^{\infty} xy\frac{1}{y}\exp(-(y+x/y))dxdy \\
&= \int_{y=0}^{\infty} y\frac{1}{y}\exp(-y)\int_{x=0}^{\infty} x\exp(-x/y)dxdy \\
&= \int_{y=0}^{\infty}\exp(-y)y^2dy \\
&= 2 \\
Cov(X,Y) &= E(XY)-E(X)E(Y) = 1
\end{aligned}$$

Problem 7

X_1, X_2, X_3, X_4 are pairwise uncorrelated random variables each having mean 0 and variance 1.

$$\begin{aligned}
cov(X_1+X_2, X_2+X_3) &= cov(X_1,X_2)+cov(X_1,X_3)+cov(X_2,X_2)+cov(X_2,X_3) \\
&= 0+0+1+0 = 1 \\
var(X_1+X_2) &= var(X_1)+var(X_2)+2cov(X_1,X_2) \\
&= 1+1+0 = 2 \\
var(X_2+X_3) &= var(X_2)+var(X_3)+2cov(X_2,X_3) \\
&= 1+1+0 = 2 \\
\rho(X_1+X_2, X_2+X_3) &= \frac{cov(X_1+X_2, X_2+X_3)}{\sqrt{var(X_1+X_2)var(X_2+X_3)}} \\
&= 0.5 \\
cov(X_1+X_2, X_3+X_4) &= cov(X_1,X_3)+cov(X_1,X_4)+cov(X_2,X_3)+cov(X_2,X_4) \\
&= 0 \\
\rho(X_1+X_2, X_3+X_4) &= 0
\end{aligned}$$

Problem 8

$$
\begin{aligned}
f(x,y) &= \frac{e^{-x/y}e^{-y}}{y} \qquad 0<x<\infty, 0<y<\infty \\
f_Y(y) &= \int_{x=0}^{\infty} \frac{e^{-x/y}e^{-y}}{y} dx \\
&= e^{-y} \qquad y>0 \\
f_{X|Y}(x|y) &= \frac{f(x,y)}{f(y)} = \frac{e^{-x/y}}{y} \\
E(X^2|Y=y) &= \int_{x=0}^{\infty} x^2 \frac{e^{-x/y}}{y} dx \\
&= 2y^2
\end{aligned}
$$

Problem 9

Let $Y = logX$ then Y is $N(m, \sigma^2)$ so

$$
\begin{aligned}
M_Y(t) &= E[e^{tY}] = e^{tm+\frac{1}{2}t^2\sigma^2} \\
&= E[e^{tlogX}] = E[e^{logX^t}] \\
&= E[X^t]
\end{aligned}
$$

Hence

$$E[X] = M_Y(1) = e^{m+\frac{1}{2}\sigma^2}$$

and

$$
\begin{aligned}
var(X) &= E[X^2] - (E[X])^2 = M_Y(2) - (M_Y(1))^2 \\
&= M_Y(2) - e^{2m+\sigma^2} \\
&= e^{2m+2\sigma^2} - e^{2m+\sigma^2} \\
&= e^{2m+\sigma^2}(e^{\sigma^2} - 1)
\end{aligned}
$$

Problem 10

Let $\psi(t) = \log(M(t))$.

Then

$$\psi\text{"}(t) = \frac{d^2}{dt^2}\psi(t) = \frac{M''(t)}{M(t)} - \left(\frac{M'(t)}{M(t)}\right)^2$$

Now $M(0) = 1$, $M'(0) = E[X]$ and $M''(0) = E[X^2]$ by the definition of $M(t)$ being the moment generating function of X.

Therefore

$$\psi''(0) = E[X^2] - (E[X])^2 = var(X)$$

Problem 11

$$Y = \sum_{i=1}^{N} X_i$$

Then:

$$\begin{aligned} M_Y(t) &= E[e^{t\sum_{i=1}^{N} X_i}] \\ &= E[E[e^{t\sum_{i=1}^{N} X_i}|N]] \\ &= E[(M(t))^N] \\ &= \phi(M(t) \end{aligned}$$

where we have used $E[e^{t\sum_{i=1}^{N} X_i}|N = n] = (M(t))^n$ from the independence of x_i and N and

$$E[E[(M(t))^N|N]] = E[(M(t))^N] = \phi(M(t))$$

by definition of the generating function of N.

Problem 12

Our estimate of X is $\widehat{X} = aY + b$ where a and b should be chosen to minimize the mean squared error which is $J = E(X - \widehat{X})^2$. To minimize J, we set $\frac{\partial J}{\partial a} = 0$ and $\frac{\partial J}{\partial b} = 0$

Now:

$$J = E(X - aY - b)^2 = E[X^2] - 2aE[XY] - 2bE[X] - +2abE[Y] + a^2E[Y^2] + b^2$$

Therefore,

$$\begin{aligned} \frac{\partial J}{\partial a} &= 0 \\ \Rightarrow E(2(X - (aY + b))(-Y)) &= 0 \\ \Rightarrow E(XY) &= aE(Y^2) + bE(Y) \end{aligned} \tag{1}$$

$$\begin{aligned} \frac{\partial J}{\partial b} &= 0 \\ \Rightarrow E((2(X - (aY + b))(-1)) &= 0 \\ \Rightarrow E(X) &= aE(Y) + b \end{aligned} \tag{2}$$

Solving 1 and 2 for a and b

$$\begin{aligned} a &= \frac{E(XY) - E(X)E(Y)}{E(Y^2) - E(Y)^2} \\ &= \frac{cov(X,Y)}{Var(Y)} \end{aligned}$$

and

$$b = E(X) - aE(Y)$$

Problem 13

Let $J = E((X-a)^2)$. To minimize J, we should have $\frac{dJ}{da} = 0$. Therefore

$$\begin{aligned} \frac{dJ}{da} &= 0 \\ \Rightarrow 2E(X-a) &= 0 \\ \Rightarrow a &= E(X) \end{aligned}$$

The second part needs some proof.

First note that $|X-a| = (X-a)\mathbf{1}_{[X>a]} - (X-a)\mathbf{1}_{[X\leq a]}$.

Therefore:

$$E|X-a| = \int_a^\infty (x-a)f_X(x)dx - \int_{-\infty}^a (x-a)f_X(x)dx$$

Differentiating (using Liebniz's rule) w.r.t a and setting the derivative to 0 gives:

$$-\int_a^\infty f_X(x)dx + \int_{-\infty}^a f_X(x)dx = 0$$

Noting that $\int_{-\infty}^a f_X(x)dx = F_X(a)$ and $\int_a^\infty f_X(x)dx = 1 - F_X(a)$ we obtain:

$$F_X(a) = 1 - F_X(a)$$

or $F_X(a) = 0.5$ so a corresponds to the median.

Problem 14

Without loss of generality assume $E[X] = E[Y] = 0$ then

$$Cov(X+Y, X-Y) = E[(X-Y)(X+Y)] = E[X^2] - E[Y^2] = (Var(X) - Var(Y) = 0$$

since X and Y are identically distributed.

Problem 15

Once again without loss of generality assume that $E[X] = E[Y] = 0$ (convince yourselves that if $E[X] = m_X$ and $E[Y] = m_Y$ the answer still holds. Now $Y = a + bX$,

Hence $cov(XY) = E(XY) = aE[X] + bE[X^2] = bvar(X)$

Now $var(Y) = b^2Var(X)$

Therefore

$$\rho(X,Y) = \frac{bvar(X)}{|b|var(X)} = sign(b)$$

from which the answer follows.

Additional Problem

Let us show the Cauchy-Schwarz inequality that states:

Let X and Y be jointly distributed with $E[X^2] < \infty$ and $E[Y^2] < \infty$

$$|cov(X,Y)| \leq \sqrt{var(X)var(Y)}$$

Without loss of generality assume $E[X] = E[Y] = 0$.

Assume $E[Y^2] > 0$ and let $a = \frac{E[XY]}{E[Y^2]}$.

Then

$$\begin{aligned} 0 &\leq E[Y^2]E[(X - aY)^2] \\ &= E[Y^2](E[X^2] + a^2E[Y^2] - 2aE[XY] \\ &= E[Y^2]E[X^2] - (E[XY])^2 \end{aligned}$$

or

$$(cov(X,Y))^2 = (E[XY])^2 \leq E[X^2]E[Y^2] = var(X)var(Y) \quad (since\ the\ r.v's\ are\ zero\ mean)$$

from which the result follows.