Probability and Statistics: MA6.101

Tutorial 6

Topics Covered: Conditional Probability, Conditional Expectation, Law of Iterated Expectation, Sums of Random Variables

Q1: Let X and Y be two independent $N(0,1)$ random variables, and define:

$$Z = 1 + X + XY^2$$
$$W = 1 + X$$

To find $\text{Cov}(Z,W)$.

A: We are asked to compute the covariance $\text{Cov}(Z,W)$, where:

$$Z = 1 + X + XY^2, \quad W = 1 + X$$

Using the properties of covariance:

$$\text{Cov}(Z,W) = \text{Cov}(1 + X + XY^2, 1 + X)$$

$$= \text{Cov}(X + XY^2, X)$$

$$= \text{Cov}(X,X) + \text{Cov}(XY^2, X)$$

$$= \text{Var}(X) + \mathbb{E}[X^2Y^2] - \mathbb{E}[XY^2]\mathbb{E}[X]$$

Since X and Y are independent, we have:

$$\text{Var}(X) = 1, \quad \mathbb{E}[X^2] = 1, \quad \mathbb{E}[Y^2] = 1, \quad \mathbb{E}[X] = 0$$

Thus:

$$\text{Cov}(Z,W) = 1 + \mathbb{E}[X^2]\mathbb{E}[Y^2] - 0 = 1 + 1 - 0 = 2$$

Therefore, the covariance is:

$$\text{Cov}(Z,W) = 2$$

Q2: The joint density function is given as $f_{X,Y}(x,y) = cx(y-x)e^{-y}$ for $0 \le x \le y < \infty$.

(a) Find c.

(b) Show that:

$$f_{X|Y}(x|y) = \frac{6x(y-x)}{y^3}, \quad 0 \le x \le y$$
$$f_{Y|X}(y|x) = (y-x)e^{x-y}, \quad 0 \le x \le y < \infty$$

(c) Deduce that:

$$\mathbb{E}(X|Y) = \frac{Y}{2}$$

A:

(a)

$$\int\int_{x,y} f_{X,Y}(x,y)dxdy = 1$$
$$\implies \int_y\int_x f_{X,Y}(x,y)dxdy = 1$$

Since x is upper bounded by y, we take the limit of x from 0 to y. And since we are initially calculating marginal pdf of $f_Y(y)$, the outside integral will be from 0 to ∞ as y can take all these values.

$$\int_0^\infty \int_0^y c \cdot x(y-x)e^{-y}\, dx\, dy = 1$$

$$\implies \int_0^\infty c \cdot e^{-y}\, dy \int_0^y x(y-x)\, dx = 1$$

$$\implies \int_0^\infty c \cdot e^{-y} \left[\frac{yx^2}{2} - \frac{x^3}{3}\right]_0^y dy = 1$$

$$\implies \int_0^\infty c \cdot e^{-y} \cdot \frac{y^3}{6}\, dy = 1$$

$$\implies \frac{c}{6}\int_0^\infty y^3 e^{-y}\, dy = 1$$

$$\implies \frac{c}{6}[-(y^3+3y^2+6y+6)e^{-y}]_0^\infty = 1$$

$$\implies \frac{c}{6} \cdot 6 = 1 \implies \boxed{c = 1}$$

(b) For conditional PDF

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{X,Y}(x,y)}{\int_x f(x,y) \cdot dx}$$

Similarly

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{f(x,y)}{\int_y f(x,y) \cdot dy}$$

Let us first calculate all the marginal pdfs.

$$f_X(x) = \int_0^\infty f_{X,Y}(x,y) \cdot dy$$

$$\implies f_X(x) = \int_0^x f_{X,Y}(x,y) \cdot dy + \int_x^\infty f_{X,Y}(x,y) \cdot dy$$

Since for the interval $0 \leq y \leq x$ does not have any density

$$\implies f_X(x) = 0 + \int_x^{\infty} f_{X,Y}(x,y) \cdot dy$$

$$\implies f_X(x) = \int_x^{\infty} x \cdot (y-x) \cdot e^{-y} \cdot dy$$

$$= x \cdot \int_x^{\infty} (y-x) \cdot e^{-y} \cdot dy$$

$$= x^2 e^{-x} + xe^{-x} - x^2 e^{-x}$$

$$\therefore f_X(x) = x \cdot e^{-x}$$

Now for $f_Y(y)$

$$f_Y(y) = \int_0^y f_{X,Y}(x,y) \cdot dx$$

$$\implies f_Y(y) = \int_0^y x \cdot (y-x) \cdot e^{-y} \cdot dx$$

$$= e^{-y} \cdot \int_x^{\infty} x \cdot (y-x) \cdot dy$$

$$= e^{-y} \cdot (\frac{y \cdot y^2}{2} - \frac{y^3}{3})$$

$$= e^{-y} \cdot (\frac{y^3}{6})$$

$$\therefore f_Y(y) = e^{-y} \cdot \frac{y^3}{6}$$

Substituting the values in the formulas specified above, we will get these expressions.

(c)

$$\mathbb{E}_X[X|Y=y] = \int_x x \cdot f_{X|Y}(x|y) \cdot dx$$

$$\mathbb{E}_X[X|Y=y] = \int_0^y x \cdot \frac{6x \cdot (y-x)}{y^3} \cdot dx$$

$$= \frac{6}{y^3} \cdot \int_0^y x^2 \cdot (y-x) \cdot dx$$

$$= \frac{6}{y^3}[\frac{yx^3}{3} - \frac{x^4}{4}]_0^y$$

$$= \frac{6}{y^3}(\frac{y^4}{12}) = \frac{y}{2}$$

$$\therefore \mathbb{E}_X[X|Y=y] = \frac{y}{2}$$

$\implies$ $\mathbb{E}_X[X|Y]$ is a random variable function $g(Y)$ and takes the value $\mathbb{E}_X[X|Y] = \frac{Y}{2}$

Q3: You throw a fair six-sided die until you get 6. What is the expected number of throws (including the throw giving 6) conditioned on the event that all throws gave even numbers?

A: Let N be the random variable representing the number of throws till the first 6, and let E be the event that all throws are even. The conditional expectation is given by

$$\mathbb{E}[N|E] = \sum_{n=1}^{\infty} n p_{N|E}(n|E)$$

The conditional pmf is given by

$$p_{N|E}(n|E) = \frac{P(\{N = n\} \cap E)}{P(E)}$$

$P(\{N = n\} \cap E)$ is equivalent to the probability of rolling $n-1$ 2s or 4s and then rolling a 6

$$P(\{N = n\} \cap E) = \left(\frac{2}{6}\right)^{n-1} \left(\frac{1}{6}\right)$$

$$\begin{aligned} P(E) &= \sum_{N} P(\{N = n\} \cap E) \\ &= \sum_{n=1}^{\infty} \left(\frac{2}{6}\right)^{n-1} \left(\frac{1}{6}\right) \\ &= \frac{1}{4} \end{aligned}$$

Substituting, we get

$$p_{N|E}(n|E) = \left(\frac{1}{3}\right)^{n-1} \left(\frac{2}{3}\right)$$

$$\begin{aligned} \mathbb{E}[N|E] &= \sum_{n=1}^{\infty} n p_{N|E}(n|E) \\ &= \sum_{n=1}^{\infty} n \left(\frac{1}{3}\right)^{n-1} \left(\frac{2}{3}\right) \\ &= \frac{2}{3} \times \frac{9}{4} \\ &= \frac{3}{2} \end{aligned}$$

Q4: Let X and Y be two independent Uniform$(0, 1)$ random variables, and define:

$$Z = \frac{X}{Y}$$

(a) Find CDF of Z .

(b) Find PDF of Z

A:
Let X and Y be two independent Uniform$(0, 1)$ random variables. We aim to find the cumulative distribution function (CDF) and probability density function (PDF) of the random variable:

$$Z = \frac{X}{Y}.$$

CDF of Z:

The CDF of Z is defined as:

$$F_Z(z) = P\left(Z \le z\right) = P\left(\frac{X}{Y} \le z\right).$$

Since X and Y are independent and uniformly distributed over $(0, 1)$, we can write this as:

$$F_Z(z) = P\left(X \le zY\right).$$

We now express the probability as an integral over the possible values of Y:

$$F_Z(z) = \int_0^1 P\left(X \le zy \mid Y = y\right) f_Y(y)\, dy.$$

For $X \sim$ Uniform$(0, 1)$, we have $P(X \le zy) = \min(1, zy)$, so the CDF becomes:

$$F_Z(z) = \int_0^1 \min(1, zy)\, dy.$$

Now, let's evaluate the integral in two parts, based on the value of z.

- If $z \le 1$, the integration of $\min(1, zy)$ is over $zy \le 1$, which simplifies to zy for $y \in [0, 1]$.
- If $z > 1$, the minimum value becomes 1 for $y \in [0, 1]$, so the integral is over the entire interval.

Therefore, for $z \le 1$:

$$F_Z(z) = \int_0^1 zy\, dy = \frac{z}{2}.$$

For $z > 1$:

$$F_Z(z) = \int_0^{1/z} zy\, dy + \int_{1/z}^1 1\, dy = \frac{1}{2z} + \left(1 - \frac{1}{z}\right).$$

Thus, the CDF of Z is:

$$F_Z(z) = \begin{cases} \frac{z}{2}, & \text{if } z \leq 1, \\ 1 - \frac{1}{2z}, & \text{if } z > 1. \end{cases}$$

PDF of Z:

To find the PDF of Z, we differentiate the CDF:

$$f_Z(z) = \frac{d}{dz} F_Z(z).$$

For $z \leq 1$:

$$f_Z(z) = \frac{d}{dz}\left(\frac{z}{2}\right) = \frac{1}{2}.$$

For $z > 1$:

$$f_Z(z) = \frac{d}{dz}\left(1 - \frac{1}{2z}\right) = \frac{1}{2z^2}.$$

Thus, the PDF of Z is:

$$f_Z(z) = \begin{cases} \frac{1}{2}, & \text{if } z \leq 1, \\ \frac{1}{2z^2}, & \text{if } z > 1. \end{cases}$$

Q5: Let X, Y, and Z be discrete random variables. Show the following generalizations of the law of iterated expectations.

(a) $\mathbb{E}[Z] = \mathbb{E}\left[\mathbb{E}[Z \mid X, Y]\right]$.

(b) $\mathbb{E}[Z \mid X] = \mathbb{E}\left[\mathbb{E}[Z \mid X, Y] \mid X\right]$.

A:

(a) To prove:

$$\mathbb{E}[Z] = \mathbb{E}\left[\mathbb{E}[Z \mid X, Y]\right]$$

By the law of iterated expectation, the expectation of Z can be computed by first conditioning on both X and Y, and then taking the expectation:

$$\mathbb{E}[Z] = \sum_x \sum_y \mathbb{P}(X = x, Y = y)\mathbb{E}[Z \mid X = x, Y = y]$$

Since $\mathbb{E}[Z \mid X = x, Y = y]$ is the conditional expectation, it is weighted by the joint probability $\mathbb{P}(X = x, Y = y)$, and taking the overall expectation gives us the desired result:

$$\mathbb{E}[Z] = \mathbb{E}\left[\mathbb{E}[Z \mid X, Y]\right]$$

(b) To prove:

$$\mathbb{E}[Z \mid X] = \mathbb{E}\left[\mathbb{E}[Z \mid X, Y] \mid X\right]$$

By the law of iterated expectation applied conditionally on X, we condition on both X and Y, and then take the expectation over Y, given X:

$$\mathbb{E}[Z \mid X = x] = \sum_y \mathbb{P}(Y = y \mid X = x)\mathbb{E}[Z \mid X = x, Y = y]$$

This shows that the conditional expectation of Z, given X, can be written as the expectation of the conditional expectation of Z given X and Y, with respect to Y conditioned on X. Hence, we conclude:

$$\mathbb{E}[Z \mid X] = \mathbb{E}\left[\mathbb{E}[Z \mid X, Y] \mid X\right]$$

Q6: If X and Y are arbitrary random variables for which the necessary expectations and variances exist, then prove that $\mathbf{Var}(Y) = \mathbb{E}[\mathrm{Var}_X(Y|X)] + \mathbf{Var}[\mathbb{E}_X(Y|X)]$.

A: We know that

$$\mathrm{Var}[Y] = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2$$

Applying law of iterated expectation with rv X on RHS above, we get

$$\mathrm{Var}[Y] = \mathbb{E}_X[\mathbb{E}_Y[Y^2|X]] - (\mathbb{E}_X[\mathbb{E}_Y[Y|X]])^2$$

We define conditional variance $\mathrm{Var}[Y|X]$ as

$$\mathrm{Var}_Y[Y|X = x] = \mathbb{E}_Y[Y^2|X = x] - (\mathbb{E}_Y[Y|X = x])^2$$

Note that $\mathrm{Var}_Y[Y|X]$ is also a random variable function in X. So we can substitute the value of $\mathbb{E}_Y[Y^2|X]$ from the above equation as:

$$\mathrm{Var}[Y] = \mathbb{E}_X[\mathrm{Var}_Y[Y|X] + (\mathbb{E}_Y[Y|X])^2] - (\mathbb{E}_X[\mathbb{E}_Y[Y|X]])^2$$

$$\implies \mathrm{Var}[Y] = \mathbb{E}_X[\mathrm{Var}_Y[Y|X]] + \mathbb{E}_X[(\mathbb{E}_Y[Y|X])^2] - (\mathbb{E}_X[\mathbb{E}_Y[Y|X]])^2$$

Let's define a random variable Z $= \mathbb{E}_Y[Y|X]$. Then we can write the above expression as:

$$\implies \mathrm{Var}[Y] = \mathbb{E}_X[\mathrm{Var}_Y[Y|X]] + \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2$$

$$\implies \mathrm{Var}[Y] = \mathbb{E}_X[\mathrm{Var}_Y[Y|X]] + \mathrm{Var}[Z]$$

by definition of Variance. Substituting the value of Z we get

$$\mathrm{Var}[Y] = \mathbb{E}_X[\mathrm{Var}_Y[Y|X]] + \mathrm{Var}_X[\mathbb{E}_Y[Y|X]]$$

Q7: Consider a gambler who at each gamble either wins or loses his bet with probabilities p and $1 - p$, independent of earlier gambles. When $p > \frac{1}{2}$, a popular gambling system, known as the Kelly strategy, is to always bet the fraction $2p - 1$ of the current fortune. Compute the expected fortune after n gambles, starting with x units and employing the Kelly strategy.

A: If the gambler's fortune at the beginning of a round is a, the gambler bets $a(2p-1)$. He therefore gains $a(2p-1)$ with probability p, and loses $a(2p-1)$ with probability $1-p$. Thus, his expected fortune at the end of a round is:

$$a\left(1+p(2p-1)-(1-p)(2p-1)\right)=a\left(1+(2p-1)^2\right)$$

Let X_k be the fortune after the kth round. Using the preceding calculation, we have :

$$E[X_{k+1}|X_k]=(1+(2p-1)^2)X_k$$

Taking expectation and using law of iterated expectations, we obtain :

$$E[X_{k+1}]=(1+(2p-1)^2)E[X_k]$$

and

$$E[X_1]=(1+(2p-1)^2)x$$

So, we conclude that :

$$E[X_n]=(1+(2p-1)^2)^n x$$

Q8: There are n letters and n envelopes. You put the letters randomly in the envelopes so that each letter is in one envelope. (Effectively a random permutation of n numbers chosen uniformly). Calculate the expected number of envelopes with the correct letter inside them.

A: Let X_i be the indicator random variable such that:

$$X_i=\begin{cases}1 & \text{if the } i\text{th letter ends up in the } i\text{th envelope,}\\ 0 & \text{otherwise.}\end{cases}$$

The expected value of X_i is:

$$E[X_i]=P(X_i=1)=\frac{1}{n}\quad\text{for any } i.$$

Let X be the number of letters that end up in their respective envelopes. Then,

$$X=X_1+X_2+\cdots+X_n.$$

The expected value of X is:

$$E[X]=E\left[\sum_{i=1}^{n}X_i\right]$$

Using the linearity of expectation, we have:

$$E[X]=\sum_{i=1}^{n}E[X_i]$$

Since $E[X_i] = \frac{1}{n}$ for each i, we get:

$$E[X] = \sum_{i=1}^{n} \frac{1}{n} = \frac{n}{n} = 1.$$

Therefore, we expect on average one letter to be in the correct envelope.