Probability and Statistics: MA6.101

Homework 6 Solutions

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

Q1: Let X and Y be two jointly continuous random variables with joint probability density function (PDF) given by:

$$f_{XY}(x,y) = \begin{cases} \frac{x^2}{6} + \frac{y^2}{6} + \frac{xy}{8}, & 0 \le x \le 2, 0 \le y \le 3 \\ 0, & \text{otherwise} \end{cases}$$

For $0 \le y \le 3$, find:

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

A:
Given the joint probability density function (PDF):

$$f_{XY}(x,y) = \begin{cases} \frac{x^2}{6} + \frac{y^2}{6} + \frac{xy}{8}, & 0 \le x \le 2, 0 \le y \le 3 \\ 0, & \text{otherwise} \end{cases}$$

we need to find $\mathbb{E}[X \mid Y = 2]$ and $\mathrm{Var}(X \mid Y = 2)$.
Step 1: Find $\mathbb{E}[X \mid Y = 2]$

$$\mathbb{E}[X \mid Y = 2] = \int_{-\infty}^{\infty} x f_{X|Y}(x|2)\, dx$$

Since $0 \le x \le 2$ and $0 \le y \le 3$, the conditional distribution is:

$$f_{X|Y}(x|2) = \frac{f_{XY}(x,2)}{\int_0^2 f_{XY}(x,2)\, dx}$$

The joint distribution for $Y = 2$ is:

$$f_{XY}(x,2) = \frac{x^2}{6} + \frac{4}{6} + \frac{2x}{8} = \frac{x^2}{6} + \frac{2}{3} + \frac{x}{4}$$

Thus,

$$\mathbb{E}[X \mid Y = 2] = \frac{\int_0^2 x\left(\frac{x^2}{6} + \frac{2}{3} + \frac{x}{4}\right) dx}{\int_0^2 \left(\frac{x^2}{6} + \frac{2}{3} + \frac{x}{4}\right) dx}$$

Step 2: Solve the integral in the denominator

$$\int_0^2 \left(\frac{x^2}{6} + \frac{2}{3} + \frac{x}{4}\right) dx = \frac{1}{6}\int_0^2 x^2\, dx + \frac{2}{3}\int_0^2 1\, dx + \frac{1}{4}\int_0^2 x\, dx$$

$$= \frac{1}{6}\left[\frac{x^3}{3}\right]_0^2 + \frac{2}{3}[x]_0^2 + \frac{1}{4}\left[\frac{x^2}{2}\right]_0^2 = \frac{1}{6} \times \frac{8}{3} + \frac{2}{3} \times 2 + \frac{1}{4} \times 2$$

$$= \frac{4}{9} + \frac{4}{3} + \frac{1}{2} = \frac{20.5}{9}$$

Step 3: Solve the integral in the numerator

$$\int_0^2 x\left(\frac{x^2}{6} + \frac{2}{3} + \frac{x}{4}\right) dx = \frac{1}{6}\int_0^2 x^3\,dx + \frac{2}{3}\int_0^2 x\,dx + \frac{1}{4}\int_0^2 x^2\,dx$$

$$= \frac{1}{6}\left[\frac{x^4}{4}\right]_0^2 + \frac{2}{3}\left[\frac{x^2}{2}\right]_0^2 + \frac{1}{4}\left[\frac{x^3}{3}\right]_0^2 = \frac{1}{6}\times\frac{16}{4} + \frac{2}{3}\times 2 + \frac{1}{4}\times\frac{8}{3}$$

$$= \frac{16}{24} + \frac{4}{3} + \frac{2}{3} = \frac{8}{3}$$

Step 4: Calculate $\mathbb{E}[X \mid Y = 2]$

$$\mathbb{E}[X \mid Y = 2] = \frac{\frac{8}{3}}{\frac{20.5}{9}} = \frac{8}{3}\times\frac{9}{20.5} = \frac{24}{20.5} \approx 1.17$$

Step 5: Find $\mathbb{E}[X^2 \mid Y = 2]$

$$\mathbb{E}[X^2 \mid Y = 2] = \frac{\int_0^2 x^2\left(\frac{x^2}{6} + \frac{2}{3} + \frac{x}{4}\right) dx}{\int_0^2 \left(\frac{x^2}{6} + \frac{2}{3} + \frac{x}{4}\right) dx}$$

Using the same denominator as in step 2, calculate the numerator:

$$\int_0^2 x^2\left(\frac{x^2}{6} + \frac{2}{3} + \frac{x}{4}\right) dx = \frac{1}{6}\int_0^2 x^4\,dx + \frac{2}{3}\int_0^2 x^2\,dx + \frac{1}{4}\int_0^2 x^3\,dx$$

$$= \frac{1}{6}\left[\frac{x^5}{5}\right]_0^2 + \frac{2}{3}\left[\frac{x^3}{3}\right]_0^2 + \frac{1}{4}\left[\frac{x^4}{4}\right]_0^2 = \frac{1}{6}\times\frac{32}{5} + \frac{2}{3}\times\frac{8}{3} + \frac{1}{4}\times\frac{16}{4}$$

$$= \frac{32}{30} + \frac{16}{9} + 4 = 6.85$$

Thus,

$$\mathbb{E}[X^2 \mid Y = 2] = \frac{6.85}{\frac{20.5}{9}} \approx 3.0$$

Step 6: Find Var$(X \mid Y = 2)$

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

$$\mathrm{Var}(X \mid Y = 2) = 3.0 - (1.17)^2 = 3.0 - 1.37 = 1.63$$

Thus, the conditional expectation and variance are:

$$\mathbb{E}[X \mid Y = 2] \approx 1.17, \quad \mathrm{Var}(X \mid Y = 2) \approx 1.63$$

Q2: Let X and Y be two independent $N(0, 1)$ random variables, and $U = X + Y$.

(a) Find the conditional PDF of U given $X = x$, $f_{U|X}(u|x)$.

(b) Find the PDF of U, $f_U(u)$.

(c) Find the conditional PDF of X given $U = u$, $f_{X|U}(x|u)$.

(d) Find $E[X|U = u]$, and $\mathrm{Var}(X|U = u)$.

A:

(a)

$$\boxed{f_{U|X}(u|x) \sim N(x, 1)}$$

To solve this, write the CDF of random variable $U|X$, which is given by

$$\begin{aligned} F_{U|X}(u|x) &= \mathbb{P}(U \leq u|X = x) \\ &= \mathbb{P}(X + Y \leq u|X = x) \\ &= \mathbb{P}(x + Y \leq u|X = x) \\ &= \mathbb{P}(Y \leq u - x|X = x) \\ &= F_{Y|X}(u - x|x) = F_Y(u - x) \end{aligned}$$

since X and Y are independent. Now converting this into PDF, we get

$$f_{U|X}(u|x) = f_Y(u - x)$$

When you substitute the value in $N(0, 1)$ then you will see that this takes the form of Gaussian RV with mean x and variance 1.

(b)

$$\boxed{f_U(u) \sim N(0, 2)}$$

Solve this by getting marginal PDF from the PDF calculated above. Try to make the expression in numerator of exponential while calculating a whole square (this method is also known as completing the square), and then convert the exponential obtained into a Gaussian RV with some mean and variance (you need to find corresponding values of mean and variance, which when simplified gives the expression in exponent).
Then use the fact that integration of PDF for a Gaussian RV will sum to 1, giving at the end PDF of Gaussian with mean 0 and variance 2.
A generalised result: Sum of N Gaussian RV is a Gaussian RV with mean and variance as sum of all the N Gaussian means and variances respectively.

(c)

$$\boxed{f_{X|U}(x|u) \sim N(\frac{u}{2}, \frac{1}{2})}$$

Apply Bayes Theorem using the results already calculated.

(d)

$$\mathbb{E}_X[X|U = u] = \mathbb{E}_X[N(\frac{u}{2}, \frac{1}{2})] = \frac{u}{2}$$

(e)

$$\mathrm{Var}_X[X|U = u] = \mathrm{Var}_X[N(\frac{u}{2}, \frac{1}{2})] = \frac{1}{2}$$

Q3: Fraser runs a dolphin-watch business. Every day, he is unable to run the trip due to bad weather with probability p, independently of all other days. Fraser works every day except the bad-weather days, which he takes as holiday.

Let Y be the number of consecutive days Fraser has to work between bad-weather days. Let X be the total number of customers who go on Fraser's trip in this period of Y days. Conditional on Y, the distribution of X is

$$(X|Y) \sim \text{Poisson}(\mu Y).$$

(a) Name the distribution of Y, and state $E(Y)$ and $\text{Var}(Y)$.

(b) Find the expectation and the variance of the number of customers Fraser sees between bad-weather days, $E(X)$ and $\text{Var}(X)$.

[Poisson(λ) Random Variable is a discrete random variable with mean λ and variance λ.]

A:

(a) Let 'success' be 'bad-weather day' and 'failure' be 'work-day'.
Then $\mathbb{P}(\text{success}) = \mathbb{P}(\text{bad-weather}) = p$.
Y is the number of failures before the first success.
So

$$Y \sim \text{Geometric}(p).$$

Thus

$$\mathbb{E}(Y) = \frac{1-p}{p},$$
$$\text{Var}(Y) = \frac{1-p}{p^2}.$$

(b) We know $(X|Y) \sim \text{Poisson}(\mu Y)$: so
$$\mathbb{E}(X|Y) = \text{Var}(X|Y) = \mu Y.$$

By the Law of Total Expectation:

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}_Y\{\mathbb{E}(X|Y)\} \\ &= \mathbb{E}_Y(\mu Y) \\ &= \mu \mathbb{E}_Y(Y) \\ \therefore \mathbb{E}(X) &= \frac{\mu(1-p)}{p}.\end{aligned}$$

By the Law of Total Variance:

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}_Y(\mathrm{Var}(X|Y)) + \mathrm{Var}_Y(\mathbb{E}(X|Y)) \\ &= \mathbb{E}_Y(\mu Y) + \mathrm{Var}_Y(\mu Y) \\ &= \mu \mathbb{E}_Y(Y) + \mu^2 \mathrm{Var}_Y(Y) \\ &= \mu\left(\frac{1-p}{p}\right) + \mu^2\left(\frac{1-p}{p^2}\right) \\ &= \frac{\mu(1-p)(p+\mu)}{p^2}.\end{aligned}$$

Q4: The following is one formulation of a famous "two envelope" paradox.
Jill is a money-loving individual who, given two options, invariably chooses the one that gives her the most money in expectation. One day Harry, a trusted (and capable of delivering) individual, offers her the following deal as a gift. He will secretly toss a fair coin until the first time that it comes up tails. If there are n heads before the first tails, he will place 10^n dollars in one envelope and 10^{n+1} dollars in the second envelope.(Thus, the probability that one envelope has 10^n dollars and the other has 10^{n+1} dollars is 2^{-n-1} for $n \geq 0$.) Harry will then hand Jill the pair of envelopes (randomly ordered, indistinguishable from the outside) and invite her to choose one. After Jill chooses an envelope she will be allowed to open it. Once she does, she will be allowed to either keep the money in the first envelope or switch to the second envelope and keep whatever amount of money is in the second envelope. However, if she decides to switch envelopes, she has to pay a one dollar "switching fee."

(a) If Jill finds 100 dollars in the first envelope she opens, what is the conditional probability that the other envelope contains 1000 dollars? What is the conditional probability that the other envelope contains 10 dollars?

(b) If Jill finds 100 dollars in the first envelope she opens, how much money does Jill expect to win from the game if she does not switch envelopes? (Answer: 100 dollars.) How much does she expect to win (net, after the switching fee) if she does switch envelopes?

(c) Generalize the answers above to the case that the first envelope contains 10^n dollars (for $n \geq 0$) instead of 100.

A: We solve (c) and then apply its results to (a) and (b)
Let Y be the random variable for number of heads before the first tails, X_1 be the amount of money in the envelope chosen by Jill, and X_2 be the amount of money in the other envelope.
The pmf of Y is

$$p_Y(y) = \frac{1}{2^{y+1}}$$

PMF of X_1

First let's find the pmf of X_1

$$p_{X_1}(x) = \sum_y p_{X_1|Y}(x|y) p_Y(y)$$

If $x = 10^n$, then $p_{X_1|Y}(x|y) = 0$ for $y \notin \{n-1, n\}$

$$\implies p_{X_1}(10^n) = p_{X_1|Y}(10^n|n-1)p_Y(n-1) + p_{X_1|Y}(10^n|n)p_Y(n)$$

Now, $p_{X_1|Y}(10^n|n-1) = p_{X_1|Y}(10^n|n) = \frac{1}{2}$ since the envelopes are chosen randomly

$$\begin{aligned}
\implies p_{X_1}(10^n) &= \frac{1}{2}\left(\frac{1}{2^n} + \frac{1}{2^{n+1}}\right) \\
&= \frac{3}{2^{n+2}}
\end{aligned}$$

Conditional PMF of X_2 given X_1

The conditional pmf of X_2 conditioned on X_1, $p_{X_2|X_1}(x_2|x_1)$ is given by (the identity can be verified easily)

$$\begin{aligned}
p_{X_2|X_1}(x_2|x_1) &= \sum_y p_{X_2|X_1,Y}(x_2|x_1,y)p_{Y|X_1}(y|x_1) \\
&= \sum_y p_{X_2|X_1,Y}(x_2|x_1,y)p_{X_1|Y}(x_1|y)\frac{p_Y(y)}{p_{X_1}(x_1)}
\end{aligned}$$

Let $x_1 = 10^n$
If $x_2 = 10^{n+1}$, then $(X_1, X_2) = (10^n, 10^{n+1})$ is possible only for $Y = n$

$$\begin{aligned}
p_{X_2|X_1}(10^{n+1}|10^n) &= p_{X_2|X_1,Y}(10^{n+1}|10^n, n)p_{X_1|Y}(10^n|n)\frac{p_Y(n)}{p_{X_1}(10^n)} \\
&= 1 \times \frac{1}{2} \times \frac{\frac{1}{2^{n+1}}}{\frac{3}{2^{n+2}}} \\
&= \frac{1}{3}
\end{aligned}$$

$p_{X_2|X_1}(10^{n-1}|10^n) = 1 - \frac{1}{3} = \frac{2}{3}$

Net reward for switching

Let $R(X_1, A)$ represent the reward for action A when the first envelope contains X_1 dollars.

$$R(10^n, \text{keep}) = 10^n$$

$$\begin{aligned}
R(10^n, \text{switch}) &= -1 + \mathbb{E}(X_2|X_1 = 10^n) \\
&= -1 + 10^{n+1} \times \frac{1}{3} + 10^{n-1} \times \frac{2}{3} \\
&= 34 \times 10^{n-1} - 1
\end{aligned}$$

(a) Probability of 1000 dollars = $\frac{1}{3}$
Probability of 100 dollars = $\frac{2}{3}$
(b) Net gain after switching = 340-1 = 339 dollars

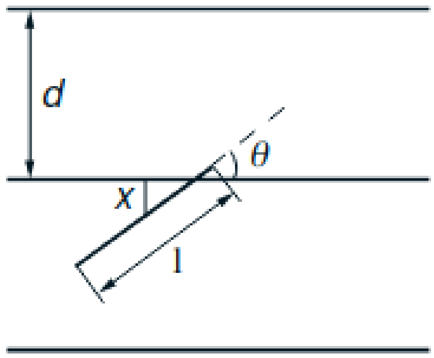

Figure 1: Figure for question 10

Q5: A surface is ruled with parallel lines, which are at distance d from each other See Figure 1. Suppose that we throw a needle of length l on the surface at random. What is the probability that the needle will intersect one of the lines?

A: We assume here that $l < d$ so that the needle cannot intersect two lines simultaneously. Let X be the distance from the midpoint of the needle to the nearest of the parallel lines, and let Θ be the acute angle formed by the axis of the needle and the parallel lines (see Fig. 1). We model the pair of random variables (X, Θ) with a uniform joint PDF over the rectangle $[0, d/2] \times [0, \pi/2]$, so that:

$$f_{X,\Theta}(x, \theta) = \begin{cases} \frac{4}{\pi d} & \text{if } x \in [0, d/2] \text{ and } \theta \in [0, \pi/2], \\ 0 & \text{otherwise.} \end{cases}$$

As can be seen from Fig. 3.17, the needle will intersect one of the lines if and only if:

$$X \leq \frac{l}{2} \sin \Theta,$$

so the probability of intersection is:

$$P\left(X \leq \frac{l}{2} \sin \Theta\right) = \int \int_{x \leq \frac{l}{2} \sin \theta} f_{X,\Theta}(x, \theta)\, dx\, d\theta.$$

Substituting $f_{X,\Theta}(x, \theta)$ into the expression, we get:

$$P\left(X \leq \frac{l}{2} \sin \Theta\right) = \frac{4}{\pi d} \int_0^{\pi/2} \int_0^{\frac{l}{2} \sin \theta} dx\, d\theta.$$

Evaluating the inner integral with respect to x:

$$= \frac{4}{\pi d} \int_0^{\pi/2} \frac{l}{2} \sin \theta\, d\theta.$$

Now, integrate with respect to θ:

$$= \frac{2l}{\pi d} \left(-\cos\theta\right)\Big|_0^{\pi/2}.$$

Substituting the limits of integration:

$$= \frac{2l}{\pi d}\left(-\cos\frac{\pi}{2} + \cos 0\right) = \frac{2l}{\pi d}.$$

Thus, the probability of intersection is $\frac{2l}{\pi d}$.
The probability of intersection can be empirically estimated by repeating the experiment a large number of times. Since it is equal to $\frac{2l}{\pi d}$, this provides us with a method for the experimental evaluation of π.