Quiz Section 8 – Solutions

Review

1) **Markov's Inequality**: Let X be a non-negative random variable, and $\alpha > 0$. Then, $\mathbb{P}(X \geqslant \alpha) \leqslant \frac{\mathbb{E}[X]}{\alpha}$.

2) **Chebyshev's Inequality**: Suppose Y is a random variable with $\mathbb{E}[Y] = \mu$ and $\text{Var}(Y) = \sigma^2$. Then, for any $\alpha > 0$, $\mathbb{P}(|Y - \mu| \geqslant \alpha) \leqslant \frac{\sigma^2}{\alpha^2}$.

3) **Chernoff Bound**: Suppose $X = X_1 + \cdots + X_n$ where the X_i are independent and in $[0, 1]$. Let $\mu = \mathbb{E}[X]$. Then, for any $0 < \delta \leqslant 1$, $\mathbb{P}(|X - \mu| \geqslant \delta\mu) \leqslant e^{-\delta^2\mu/4}$ and for any $\delta > 0$, $\mathbb{P}(X - \mu \geqslant \delta\mu) \leqslant e^{-\delta^2\mu/4}$.

4) **Multivariate: Discrete to Continuous:**

	Discrete	**Continuous**
Joint PMF/PDF	$p_{X,Y}(x,y) = \mathbb{P}(X = x, Y = y)$	$f_{X,Y}(x,y) \neq \mathbb{P}(X = x, Y = y)$
Joint range/support $\Omega_{X,Y}$	$\{(x,y) \in \Omega_X \times \Omega_Y : p_{X,Y}(x,y) > 0\}$	$\{(x,y) \in \Omega_X \times \Omega_Y : f_{X,Y}(x,y) > 0\}$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \leqslant x, s \leqslant y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(t,s)\,dsdt$
Normalization	$\sum_{x,y} p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_y p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$
Expectation	$\mathbb{E}[g(X,Y)] = \sum_{x,y} g(x,y)p_{X,Y}(x,y)$	$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{X,Y}(x,y)dxdy$
Independence must have	$\forall x,y, p_{X,Y}(x,y) = p_X(x)p_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$	$\forall x,y, f_{X,Y}(x,y) = f_X(x)f_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$
Conditional PMF/PDF	$p_{X\|Y}(x\|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$	$f_{X\|Y}(x\|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$
Conditional Expectation	$\mathbb{E}[X\|Y = y] = \sum_x x \cdot p_{X\|Y}(x\|y)$	$\mathbb{E}[X\|Y = y] = \int_{-\infty}^{\infty} x f_{X\|Y}(x\|y)dx$

Task 1 – Continuous joint density

The joint density of X and Y is given by

$$f_{X,Y}(x,y) = \begin{cases} xe^{-(x+y)} & x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

and the joint density of W and V is given by

$$f_{W,V}(w,v) = \begin{cases} 2 & 0 < w < v, 0 < v < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Are X and Y independent? Are W and V independent?

For two random variables X, Y to be independent, we must have $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ for all $x \in \Omega_X,\ y \in \Omega_Y$. Let's start with X and Y by finding their marginal PDFs. By definition, and using the fact that the joint PDF is 0 outside of $y > 0$, we get:

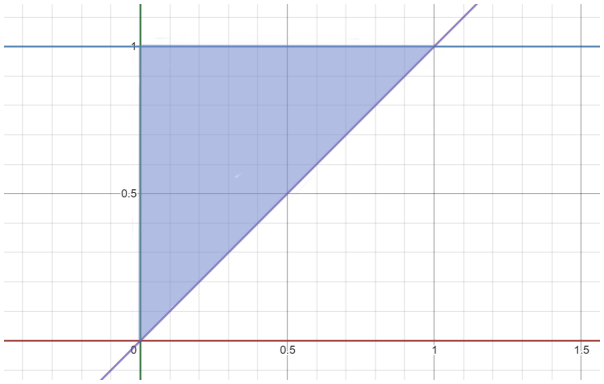
$$f_X(x) = \int_0^{\infty} xe^{-(x+y)}dy = e^{-x}x$$

We do the same to get the PDF of Y, again over the range $x > 0$:

$$f_Y(y) = \int_0^{\infty} xe^{-(x+y)}dx = e^{-y}$$

Since $e^{-x}x \cdot e^{-y} = xe^{-x-y} = xe^{-(x+y)}$ for all $x, y > 0$, X and Y are independent.

We can see that W and V are not independent simply by observing that $\Omega_W = (0,1)$ and $\Omega_V = (0,1)$, but $\Omega_{W,V}$ is not equal to their Cartesian product. Specifically, looking at their range of $f_{W,V}(w,v)$. Graphing it with w as the "x-axis" and v as the "y-axis", we see that :



The shaded area is where the joint pdf is strictly positive. Looking at it, we can see that it is not rectangular, and therefore it is not the case that $\Omega_{W,V} = \Omega_W \times \Omega_V$. Remember, the joint range being the Cartesian product of the marginal ranges is not sufficient for independence, but it is *necessary*. Therefore, this is enough to show that they are not independent.

Task 2 – Trapped Miner

A miner is trapped in a mine containing 3 doors.

- D_1: The 1^{st} door leads to a tunnel that will take him to safety after 3 hours.
- D_2: The 2^{nd} door leads to a tunnel that returns him to the mine after 5 hours.
- D_3: The 3^{rd} door leads to a tunnel that returns him to the mine after a number of hours that is Binomial with parameters $(12, \frac{1}{3})$.

At all times, he is equally likely to choose any one of the doors. What is the expected number of hours for this miner to reach safety?

Let T = number of hours for the miner to reach safety. (T is a random variable)
Let D_i be the event the i^{th} door is chosen. $i \in \{1,2,3\}$. Finally, let T_3 be the time it takes to return to the mine in the third case only (a random variable). Note that the expectation of T_3 is $12 \cdot \frac{1}{3}$ because it is binomially distributed with parameters $n = 12, p = \frac{1}{3}$. By Law of Total Expectation, linearity of expectation, and by applying the conditional expectations given by the problem statement:

$$\begin{aligned}
\mathbb{E}[T] &= \mathbb{E}[T \mid D_1] \cdot \mathbb{P}(D_1) + \mathbb{E}[T \mid D_2] \cdot \mathbb{P}(D_2) + \mathbb{E}[T \mid D_3] \cdot \mathbb{P}(D_3) \\
&= 3 \cdot \frac{1}{3} + (5 + \mathbb{E}[T]) \cdot \frac{1}{3} + (\mathbb{E}[T_3 + T]) \cdot \frac{1}{3} \\
&= 3 \cdot \frac{1}{3} + (5 + \mathbb{E}[T]) \cdot \frac{1}{3} + (\mathbb{E}[T_3] + \mathbb{E}[T]) \cdot \frac{1}{3} \\
&= 3 \cdot \frac{1}{3} + (5 + \mathbb{E}[T]) \cdot \frac{1}{3} + (4 + \mathbb{E}[T]) \cdot \frac{1}{3}
\end{aligned}$$

Solving this equation for $\mathbb{E}[T]$, we get

$$\mathbb{E}[T] = 12$$

Therefore, the expected number of hours for this miner to reach safety is 12.

Task 3 – Lemonade Stand

Suppose I run a lemonade stand, which costs me \$100 a day to operate. I sell a drink of lemonade for \$20. Every person who walks by my stand either buys a drink or doesn't (no one buys more than one). If it is raining, n_1 people walk by my stand, and each buys a drink independently with probability p_1. If it isn't raining, n_2 people walk by my stand, and each buys a drink independently with probability p_2. It rains each day with probability p_3, independently of every other day. Let X be my profit over the next week. In terms of n_1, n_2, p_1, p_2 and p_3, what is $\mathbb{E}[X]$?

Let R be the event it rains. Let X_i be how many drinks I sell on day i for $i = 1, ..., 7$. We are interested in $X = \sum_{i=1}^{7} (20X_i - 100)$. We have $X_i|R \sim$ Binomial(n_1, p_1), so $\mathbb{E}[X_i|R] = n_1 p_1$. Similarly, $X_i|R^C \sim$ Binomial(n_2, p_2), so $\mathbb{E}[X_i|R^C] = n_2 p_2$. By the law of total expectation,

$$\mu = \mathbb{E}[X_i] = \mathbb{E}[X_i|R]\mathbb{P}(R) + \mathbb{E}[X_i|R^C]\mathbb{P}\left(R^C\right) = n_1 p_1 p_3 + n_2 p_2 (1 - p_3)$$

Hence, by linearity of expectation,

$$\mathbb{E}[X] = \mathbb{E}[\sum_{i=1}^{7} (20X_i - 100)] = 20 \sum_{i=1}^{7} \mathbb{E}[X_i] - 700 = 140\mu - 700$$

$$= 140 \cdot (n_1 p_1 p_3 + n_2 p_2 (1 - p_3)) - 700.$$

Task 4 – A Dysfunctional Family

Rick and his grandson Morty are set to meet at a certain time. Since their relationship is a little strained, neither of them wants to be there on time. Let $X \sim$ Unif$(0, 10)$ be the amount of minutes Morty is going to be late. Rick has cameras around the meeting spot and will observe Morty's arrival time $X = x$. Then, he will arrive at the meeting spot Unif$(x, 5x)$ minutes late. Let Y be the random variable indicating how late Rick will be.

a) Using the above definitions determine f_X, $f_{Y|X}$, and f_{XY}.

Since X is a uniform RV on $(0, 10)$, we have

$$f_X(x) = \begin{cases} \frac{1}{10} & x \in (0, 10) \\ 0 & \text{otherwise} \end{cases}$$

Also given that $X = x$, Y is also uniform on $(x, 5x)$ so

$$f_{Y|X}(y|x) = \begin{cases} \frac{1}{4x} & y \in (x, 5x) \\ 0 & \text{otherwise} \end{cases}$$

Since $f_{Y|X}(y|x) = \frac{f_{YX}(y,x)}{f_X(x)}$, we have

$$f_{YX}(y, x) = f_{Y|X}(y|x) \cdot f_X(x) = \begin{cases} \frac{1}{40x} & x \in (0, 10) \text{ and } y \in (x, 5x) \\ 0 & \text{otherwise} \end{cases}$$

b) Compute $\mathbb{E}[Y]$.

By definition since Y conditioned on $X = x$ is a uniform RV on $(x, 5x)$, we have

$$\mathbb{E}[Y \mid X = x] = \frac{x + 5x}{2} = 3x.$$

(Alternatively, we could compute it from first principles using $f_{Y|X}$ and the definition

$$\mathbb{E}[Y|X = x] = \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y|x)dy = \int_{x}^{5x} \frac{y}{4x}dy = \frac{y^2}{8x}\Big|_{y=x}^{y=5x} = \frac{25x^2}{8x} - \frac{x^2}{8x} = \frac{24x}{8} = 3x.)$$

Then, using the Law of Total Expectation we get:

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} \mathbb{E}[Y|X = x] \cdot f_X(x)dx = \int_{0}^{10} \frac{3x}{10}dx = \frac{3}{10} \cdot \frac{x^2}{2}\Big|_0^{10} = \frac{300}{20} - 0 = 15$$

Note : This is a place where the Law of Total Expectation makes things **much** easier than figuring out the PDF of Y and doing direct calculation of the expectation:
Just to see how bad it would get... by definition we have:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)dx = \int_{0}^{10} f_{XY}(x, y)dx$$

since we know that $f_{XY}(x, y)$ is 0 when $x \leqslant 0$ or $x \geqslant 10$. However, we can't just plug in $\frac{1}{40x}$ for $f_{XY}(x, y)$ because we also need to satisfy that $x \leqslant y \leqslant 5x$ for that value to be correct. For a fixed value of y, which values of x could work? We need to have $x \leqslant y$, but we also need to have $y \leqslant 5x$ or in other worse $x \geqslant y/5$. In particular, this is equivalent to $y/5 \leqslant x \leqslant y$. Therefore we need $\max(0, y/5) \leqslant x \leqslant \min(10, y)$. Therefore

$$f_Y(y) = \int_{0}^{10} f_{XY}(x, y)dx = \int_{\max(0,y/5)}^{\min(10,y)} \frac{1}{40x}dx.$$

This would have non-zero contributions for all y with $0 \leqslant y \leqslant 50$ and would be a big mess to calculate since the integral involves $1/x$ which would have logarithms in it...

Task 5 – Tail bounds

Suppose $X \sim \text{Bin}(6, 0.4)$. We will bound $\mathbb{P}(X \geqslant 4)$ using the tail bounds we've learned, and compare this to the true result.

a) Give an upper bound for this probability using Markov's inequality. Why can we use Markov's inequality?

We know that the expected value of a binomial distribution is np, so: $\mathbb{P}(X \geqslant 4) \leqslant \frac{\mathbb{E}[X]}{4} = \frac{2.4}{4} = 0.6$.
We can use it since X is nonnegative.

b) Give an upper bound for this probability using Chebyshev's inequality. You may have to rearrange algebraically and it may result in a weaker bound.

$\mathbb{P}(X \geqslant 4) = \mathbb{P}(X - 2.4 \geqslant 1.6) \leqslant \mathbb{P}(|X - 2.4| \geqslant 1.6)$ we can add those absolute value signs because that only adds more possible values, so it is an upper bound on the probability of $X - 2.4 \geqslant 1.6$.
Then, using Chebyshev's inequality we get:
$\mathbb{P}(|X - 2.4| \geqslant 1.6) \leqslant \frac{Var(X)}{1.6^2} = \frac{1.44}{1.6^2} = 0.5625$

c) Give an upper bound for this probability using the Chernoff bound.

$$\mathbb{P}(X \geqslant 4) = \mathbb{P}(X \geqslant (1+\tfrac{2}{3})2.4) \leqslant e^{-(\frac{2}{3})^2\mathbb{E}[X]/4} = e^{-4\times 2.4/36} \approx 0.77$$

d) Give the exact probability.

Since X is a binomial, we know it has a range from 0 to n (or in this case 0 to 6). Thus, the possible values to satisfy $X \geqslant 4$ are 4, 5, or 6. We plug in the PMF for each to get: $\mathbb{P}(X \geqslant 4) = \mathbb{P}(X=4) + \mathbb{P}(X=5) + \mathbb{P}(X=6) = \binom{6}{4}(0.4)^4(0.6)^2 + \binom{6}{5}(0.4)^5(0.6) + \binom{6}{6}0.4^6 \approx 0.1792$

Task 6 – Exponential Tail Bounds

Let $X \sim \text{Exp}(\lambda)$ and $k > 1/\lambda$. Recall that $\mathbb{E}[X] = \frac{1}{\lambda}$ and $\text{Var}(X) = \frac{1}{\lambda^2}$.

a) Use Markov's inequality to bound $\text{P}(X \geqslant k)$.

$$\mathbb{P}(X \geqslant k) \leqslant \frac{1}{\lambda k}$$

b) Use Chebyshev's inequality to bound $\text{P}(X \geqslant k)$.

$$\mathbb{P}(X \geqslant k) = \mathbb{P}\left(X - \frac{1}{\lambda} \geqslant k - \frac{1}{\lambda}\right) \leqslant \mathbb{P}\left(\left|X - \frac{1}{\lambda}\right| \geqslant k - \frac{1}{\lambda}\right) \leqslant \frac{1}{\lambda^2(k-1/\lambda)^2} = \frac{1}{(\lambda k - 1)^2}$$

c) What is the exact formula for $\text{P}(X \geqslant k)$?

$$\mathbb{P}(X \geqslant k) = e^{-\lambda k}$$

d) For $\lambda k \geqslant 3$, how do the bounds given in parts (a), (b), and (c) compare?

$$e^{-\lambda k} < \frac{1}{(\lambda k - 1)^2} < \frac{1}{\lambda k}$$

so Markov's inequality gives the worst bound.

Task 7 – How many samples?

Let $X = X_1 + \ldots + X_n$ be the sum of n independent $\text{Poi}(\lambda)$ random variables. Recall that the Poisson distribution has expectation and variance both equal to λ and has the summation property that X is a $\text{Poi}(n\lambda)$ random variable.

a) How large a value of n would Chebyshev's inequality need to guarantee that $\mathbb{P}(X \leqslant \mathbb{E}[X]/2) \leqslant 0.01$?

We have

$$\mathbb{P}(X \leqslant \mathbb{E}[X]/2) = \mathbb{P}(X - \mathbb{E}[X] \leqslant -\mathbb{E}[X]/2) \leqslant \mathbb{P}(|X - \mathbb{E}[X]| \geqslant \mathbb{E}[X]/2).$$

Applying Chebyshev's inequality we have

$$\mathbb{P}(X \leqslant \mathbb{E}[X]/2) \leqslant \mathbb{P}(|X - \mathbb{E}[X]| \geqslant \mathbb{E}[X]/2) \leqslant \frac{4\text{Var}(X)}{\mathbb{E}[X]^2} = \frac{4n\lambda}{n^2\lambda^2} = \frac{4}{n\lambda}.$$

In order for this to be at most 0.01, we require $\boxed{n \geqslant 400/\lambda}$.

b) How large a value of n would Markov's inequality need to guarantee that $\mathbb{P}(X \leqslant \mathbb{E}[X]/2) \leqslant 0.01$?

X is non-negative so Markov's inequality applies to X, but no value of n will guarantee any probability less than 1.