

Quiz Section 8 – Solutions

Review

- 1) **Markov's Inequality:** Let X be a non-negative random variable, and $\alpha > 0$. Then, $\mathbb{P}(X \geq \alpha) \leq \frac{\mathbb{E}[X]}{\alpha}$.
- 2) **Chebyshev's Inequality:** Suppose Y is a random variable with $\mathbb{E}[Y] = \mu$ and $\text{Var}(Y) = \sigma^2$. Then, for any $\alpha > 0$, $\mathbb{P}(|Y - \mu| \geq \alpha) \leq \frac{\sigma^2}{\alpha^2}$.
- 3) **Chernoff Bound:** Suppose $X = X_1 + \dots + X_n$ where the X_i are independent and in $[0, 1]$. Let $\mu = \mathbb{E}[X]$. Then, for any $0 < \delta \leq 1$, $\mathbb{P}(|X - \mu| \geq \delta\mu) \leq e^{-\delta^2\mu/4}$ and for any $\delta > 0$, $\mathbb{P}(X - \mu \geq \delta\mu) \leq e^{-\delta^2\mu/4}$.
- 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	$F_{X,Y}(x, y) = \sum_{t \leq x, s \leq y} p_{X,Y}(t, s)$	$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(t, s) ds dt$
Normalization	$\sum_{x,y} p_{X,Y}(x, y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$
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	$\mathbb{E}[g(X, Y)] = \sum_{x,y} g(x, y) p_{X,Y}(x, y)$	$\mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy$
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 – 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 $X \sim \text{Unif}(0, 10)$ 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 $\text{Unif}(x, 5x)$ 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

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

(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 $x \leq 0$ or $x \geq 10$. However, we can't just plug in $\frac{1}{40x}$ for $f_{XY}(x, y)$ because we also need to satisfy that $x \leq y \leq 5x$ for that value to be correct. For a fixed value of y , which values of x could work? We need to have $x \leq y$, but we also need to have $y \leq 5x$ or in other words $x \geq y/5$. In particular, this is equivalent to $y/5 \leq x \leq y$. Therefore we need $\max(0, y/5) \leq x \leq \min(10, y)$. 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 $0 \leq y \leq 50$ and would be a big mess to calculate since the integral involves $1/x$ which would have logarithms in it...

Task 2 – Tail bounds

Suppose $X \sim \text{Bin}(6, 0.4)$. We will bound $\mathbb{P}(X \geq 4)$ 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: $\mathbb{P}(X \geq 4) \leq \frac{\mathbb{E}[X]}{4} = \frac{2.4}{4} = 0.6$.
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.

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

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

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

- 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 $X \geq 4$ are 4, 5, or 6. We plug in the PMF for each to get: $\mathbb{P}(X \geq 4) = \mathbb{P}(X = 4) + \mathbb{P}(X = 5) + \mathbb{P}(X = 6) = \binom{6}{4}(0.4)^4(0.6)^2 + \binom{6}{5}(0.4)^5(0.6) + \binom{6}{6}0.4^6 \approx 0.1792$

Task 3 – 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 $\mathbb{P}(X \geq k)$.

$$\mathbb{P}(X \geq k) \leq \frac{1}{\lambda k}$$

- b) Use Chebyshev's inequality to bound $\mathbb{P}(X \geq k)$.

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

- c) What is the exact formula for $\mathbb{P}(X \geq k)$?

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

- d) For $\lambda k \geq 3$, 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 4 – How many samples?

Let $X = X_1 + \dots + X_n$ 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 $\mathbb{P}(X \leq \mathbb{E}[X]/2) \leq 0.01$?

We have

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

Applying Chebyshev's inequality we have

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

In order for this to be at most 0.01, we require $n \geq 400/\lambda$.

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

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