# Homework 13

Michael Pham

Spring 2023

# Contents

1	Just One Tail, Please	3
2	Coupon Collector Variance	7
3	Estimating $\pi$	8
4	Short Answer	9
5	Useful Uniforms	12
6	Waiting For the Bus	13

## 1 Just One Tail, Please

Let X be some random variable with finite mean and variance which is not necessarily non-negative. The extended version of Markov's Inequality states that for a non-negative function  $\varphi(x)$  which is monotonically increasing for x>0 and some constant  $\alpha>0$ ,

$$\Pr[X \ge \alpha] \le \frac{\mathbb{E}[\varphi(X)]}{\varphi(\alpha)}$$

Suppose  $\mathbb{E}[X] = 0$ ,  $\operatorname{Var}(X) = \sigma^2 < \infty$ , and  $\alpha > 0$ .

**Problem 1.1.** Use the extended version of Markov's Inequality stated above with  $\varphi(x)=(x+c)^2$ , where c is some positive constant, to show that:

$$\Pr\left[X \ge \alpha\right] \le \frac{\sigma^2 + c^2}{(\alpha + c)^2}$$

*Solution.* We observe that for  $\varphi = (x+c)^2$ , we have the following:

$$\begin{aligned} \Pr[X \geq \alpha] &\leq \frac{\mathbb{E}[\varphi(X)]}{\varphi(\alpha)} \\ &= \frac{\mathbb{E}\left[(X+c)^2\right]}{(\alpha+c)^2} \\ &= \frac{\mathbb{E}\left[X^2 + 2Xc + c^2\right]}{(\alpha+c)^2} \\ &= \frac{\mathbb{E}\left[X^2\right] + \mathbb{E}\left[2Xc\right] + \mathbb{E}\left[c^2\right]}{(\alpha+c)^2} \\ &= \frac{\mathbb{E}[X^2] + 2c\mathbb{E}[X] + c^2\mathbb{E}[1]}{(\alpha+c)^2} \end{aligned}$$

From here, take note that  $2c\mathbb{E}[X] = 0 = -\mathbb{E}[X]^2$ . Therefore, we get:

$$\frac{\mathbb{E}[X^2] + 2c\mathbb{E}[X] + c^2\mathbb{E}[1]}{(\alpha + c)^2} = \frac{\mathbb{E}[X^2] - \mathbb{E}[X]^2 + c^2}{(\alpha + c)^2}$$
$$= \frac{\sigma^2 + c^2}{(\alpha + c)^2}$$
$$\therefore \Pr[X \ge \alpha] \le \frac{\sigma^2 + c^2}{(\alpha + c)^2}$$

**Problem 1.2.** Note that the above bound applies for all positive c, so we can choose a value of c to minimize the expression, yielding the best possible bound. Find the value for c which will minimize the RHS expression (you may assume that the expression has a unique minimum).

*Solution.* In order to find the minimum for  $\frac{\sigma^2+c^2}{(\alpha+c)^2}$ , we can first differentiate it as such:

$$\frac{d}{dc} \frac{\sigma^2 + c^2}{(\alpha + c)^2} = \frac{d}{dc} \frac{\sigma^2}{(\alpha + c)^2} + \frac{d}{dc} \frac{c^2}{(\alpha + c)^2}$$

$$= \frac{d}{dc} \sigma^2 (\alpha + c)^{-2} + \frac{d}{dc} c^2 (\alpha + c)^{-2}$$

$$= -2\sigma^2 (\alpha + c)^{-3} + 2c(\alpha + c)^{-2} - 2c^2 (\alpha + c)^{-3}$$

$$= \frac{2c(\alpha + c) - 2\sigma^2 - 2c^2}{(\alpha + c)^3}$$

Now, we will set this equal to zero to find the minimum:

$$\frac{2c(\alpha+c) - 2\sigma^2 - 2c^2}{(a+c)^3} = 0$$
$$2c(\alpha+c) - 2\sigma^2 - 2c^2 = 0$$
$$c\alpha + c^2 - \sigma^2 - c^2 = 0$$
$$c\alpha - \sigma^2 = 0$$
$$c = \frac{\sigma^2}{\alpha}$$

Thus, we have that  $c = \frac{\sigma^2}{\alpha}$ .

We can plug in the minimizing value of c you found in part (b) to prove the following bound:

$$\Pr\left[X \ge \alpha\right] \le \frac{\sigma^2}{\alpha^2 + \sigma^2}.$$

This bound is also known as Cantelli's inequality.

Now, recall that Chebyshev's inequality provides a two-sided bound. That is, it provides a bound on

$$\Pr[|X - \mathbb{E}[X]| \ge \alpha] = \Pr[X \ge \mathbb{E}[X] + \alpha] + \Pr[X \le \mathbb{E}[X] - \alpha]$$

If we only wanted to bound the probability of one of the tails, e.g. if we wanted to bound  $\Pr[X \ge \mathbb{E}[X] + \alpha]$ , it is tempting to just divide the bound we get from Chebyshev's by two.

#### Problem 1.3. Why is this not always correct in general?

Solution. This is because we can't just assume that the distribution is symmetric; if a distribution isn't symmetric, then halving our bound from Chebyshev would result in an incorrect number.

Problem 1.4. Provide an example of a random variable X (does not have to be zero-mean) and a constant  $\alpha$  such that using this method (dividing by two to bound one tail) is not correct, that is,  $\Pr\left[X \geq \mathbb{E}[X] + \alpha\right] > \frac{\operatorname{Var}(X)}{2\alpha^2}$  or  $\Pr\left[X \leq \mathbb{E}[X] - \alpha\right] > \frac{\operatorname{Var}(X)}{2\alpha^2}$ .

Solution. Suppose we have a random variable X which takes on values  $\{-1,0,1\}$  with probability  $\frac{1}{2}$  for 1 and  $\frac{1}{4}$  for -1,0. We observe that  $\mathbb{E}[X]=\frac{1}{4}$ , and  $\mathrm{Var}(X)=1$ .

Now, we observe that for  $\alpha=-\frac{3}{4}$ , we get  $\Pr[X\leq \mathbb{E}[X]-\alpha]=\Pr[X\leq 1]=1$ . Meanwhile, we see that  $\frac{\operatorname{Var}(X)}{2\alpha^2}=\frac{1}{2(\frac{3}{4})^2}=\frac{8}{9}$ .

We see then that 
$$\Pr[X \leq \mathbb{E}[X] - \alpha] > \frac{\operatorname{Var}(X)}{2\alpha^2}$$
.

Now we see the use of the bound proven in part (b) - it allows us to bound just one tail while still taking variance into account, and does not require us to assume any property of the random variable. Note that the bound is also always guaranteed to be less than 1 (and therefore at least somewhat useful), unlike Markov's and Chebyshev's inequality!

Let's try out our new bound on a simple example. Suppose X is a positively-valued random variable with  $\mathbb{E}[X]=3$  and  $\mathrm{Var}(X)=2$ .

**Problem 1.5.** What bound would Markov's inequality give for  $Pr[X \ge 5]$ ?

Solution. We observe that, using Markov's Inequality, we have:

$$\Pr[X \ge \alpha] \le \frac{\mathbb{E}[X]}{\alpha}$$
$$\Pr[X \ge 5] \le \frac{3}{5}$$

#### **Problem 1.6.** What bound would Chebyshev's inequality give for $Pr[X \ge 5]$ ?

Solution. We observe that, using Chebyshev's inequality, we get:

$$Pr[X \ge 5] = Pr[|X - \mathbb{E}[X]| \ge 5 - \mathbb{E}[X]]$$
$$= Pr[|X - 3| \ge 2]$$
$$\le \frac{1}{2}$$

### **Problem 1.7.** What bound would Cantelli's Inequality give for $Pr[X \ge 5]$ ?

Solution. Since Cantelli's Inequality works for zero-mean only, we have to first shift it (but note that variance doesn't change when we shift). This yields us:

$$\Pr[X - 3 \ge 5 - 3] \Pr[X - 3 \ge 2]$$

$$leq \frac{2}{4 + 2}$$

$$= \frac{2}{6}$$

$$= \frac{1}{3}$$

# 2 Coupon Collector Variance

It's that time of the year again—Safeway is offering its Monopoly Card promotion. Each time you visit Safeway, you are given one of n different Monopoly Cards with equal probability. You need to collect them all to redeem the grand prize.

Let X be the number of visits you have to make before you can redeem the grand prize. Remember that we've previously shown  $\operatorname{Var}(X) = n^2 \left( \sum_{i=1}^n i^{-2} \right) - \mathbb{E}[X]$ .

**Problem 2.1.** The series  $\sum_{i=1}^{\infty} i^{-2}$  converges to the constant value  $\pi^2/6$ . Using this fact and Chebyshev's Inequality, find a lower bound on  $\beta$  for which the probability you need to make more than  $\mathbb{E}[X] + \beta n$  visits is less than 1/100, for large n.

Solution. To begin with, we will use Chebyshev's Inequality as follow:

$$\begin{split} \Pr[X \geq \mathbb{E}[X] + \beta n] &= \Pr[|X - \mathbb{E}[X]| \geq \beta n] \\ &\leq \frac{\operatorname{Var}(X)}{(\beta n)^2} \\ &= \frac{\operatorname{Var}(X)}{\beta^2 n^2} \end{split}$$

We now want  $\frac{\text{Var}(X)}{\beta^2 n^2}$  to be less than  $\frac{1}{100}$  for large n:

$$\lim_{n \to \infty} \frac{\operatorname{Var}(X)}{\beta^2 n^2} = \lim_{n \to \infty} \frac{n^2 \left(\sum_{i=1}^{\infty} i^{-2}\right) - \sum_{i=1}^{n} i^{-1}}{\beta^2 n^2}$$

$$= \lim_{n \to \infty} \frac{\pi^2}{6\beta^2} - \lim_{n \to \infty} \frac{\ln n}{n^2}$$

$$= \frac{\pi^2}{6\beta^2} - \lim_{n \to \infty} \frac{1}{2n}$$

$$= \frac{\pi^2}{6\beta^2} - \lim_{n \to \infty} \frac{1}{2n^2}$$

$$= \frac{\pi^2}{6\beta^2}$$

$$= \frac{\pi^2}{6\beta^2}$$

$$\frac{\pi^2}{6\beta^2} < \frac{1}{100}$$

$$\beta > \frac{10\pi}{\sqrt{6}}$$

7

### 3 Estimating $\pi$

In this problem, we discuss some interesting ways that you could probabilistically estimate  $\pi$ , and see how good these techniques are at estimating  $\pi$ .

**Technique 1:** Buffon's needle is a method that can be used to estimate the value of  $\pi$ . There is a table with infinitely many parallel lines spaced a distance 1 apart, and a needle of length 1. It turns out that if the needle is dropped uniformly at random onto the table, the probability of the needle intersecting a line is  $\frac{2}{\pi}$ . We have seen a proof of this in the notes.

**Technique 2:** Consider a square dartboard, and a circular target drawn inscribed in the square dartboard. A dart is thrown uniformly at random in the square. The probability the dart lies in the circle is  $\frac{\pi}{4}$ .

**Technique 3:** Pick two integers x and y independently and uniformly at random from 1 to M, inclusive. Let  $p_M$  be the probability that x and y are relatively prime. Then

$$\lim_{M \to \infty} p_M = \frac{6}{\pi^2}.$$

Let  $p_1=\frac{2}{\pi}, p_2=\frac{\pi}{4}$ , and  $p_3=\frac{6}{\pi^2}$  be the probabilities of the desired events of **Technique 1**, **Technique 2**, and **Technique 3**, respectively. For each technique, we apply each technique N times, then compute the proportion of the times each technique occurred, getting estimates  $\hat{p_1}, \hat{p_2}$ , and  $\hat{p_3}$ , respectively.

**Problem 3.1.** For each  $\hat{p_i}$ , compute an expression  $X_i$  in terms of  $\hat{p_i}$  that would be an estimate of  $\pi$ .

Solution. We observe that for each  $X_i$ , we get the following:

$$X_1 = \frac{2}{\hat{p}_1}$$

$$X_2 = \frac{4\pi^2}{\hat{p}_2}$$

$$X_3 = \frac{6}{\pi \hat{p}_3}$$

**Problem 3.2.** Using Chebyshev's Inequality, compute the minimum value of N such that  $X_2$  is within  $\varepsilon$  of  $\pi$  with  $1-\delta$  confidence. Your answer should be in terms of  $\varepsilon$  and  $\delta$ .

Solution. Using Chebyshev's Inequality, we get the following expression:

$$\Pr[|X_2 - \pi| \ge \varepsilon] \le \frac{\operatorname{Var}(X_2)}{(\varepsilon)^2}$$

$$= \frac{\frac{\pi}{4} \left(1 - \frac{\pi}{4}\right)}{\frac{N}{(\varepsilon)^2}}$$

$$= \frac{3\pi^2}{16N(\varepsilon)^2}$$

$$\le 1 - \delta$$

$$N \ge \frac{3\pi^2}{16(1 - \delta)(\varepsilon)^2}$$

8

### 4 Short Answer

**Problem 4.1.** Let X be uniform on the interval [0,2], and define  $Y=4X^2+1$ . Find the PDF, CDF, expectation, and variance of Y.

Solution. For  $Y=4X^2+1$ , we get  $\Pr[Y\leq y]=\Pr[4X^2+1\leq y]=\Pr\left[X\leq \frac{\sqrt{y-1}}{2}\right]$ . We also note that our bound is now [1,17].

Now, since we know that X is uniform on the interval [0, 2], we get for the CDF:

$$\int_{0}^{\frac{\sqrt{y-1}}{2}} \frac{1}{2} dy = \frac{\sqrt{y-1}}{4}$$

Now, to get the PDF, we simply differentiate the CDF to get:

$$f(x) = \begin{cases} \frac{1}{8\sqrt{y-1}} & 1 \le y \le 17 \\ 0 & \text{elsewhere} \end{cases}$$

For the expected value, we see that

$$\mathbb{E}[Y] = \int_{1}^{17} y f(y) dy$$
$$= \int_{1}^{17} \frac{y}{8\sqrt{y-1}} dy$$
$$= \frac{19}{3}$$

For the variance, we observe the following:

$$\begin{aligned} \operatorname{Var}(Y) &= \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \\ &= \int_{1}^{17} \frac{y^2}{8\sqrt{y-1}} \mathrm{d}y - \left( \int_{1}^{17} \frac{y}{8\sqrt{y-1}} \mathrm{d}y \right)^2 \\ &= \frac{943}{15} - \frac{361}{9} \\ &= \frac{1024}{45} \end{aligned}$$

**Problem 4.2.** Let *X* and *Y* have joint distribution

$$f(x,y) = \begin{cases} cxy + \frac{1}{4} & x \in [1,2] \text{ and } y \in [0,2] \\ 0 & \text{otherwise.} \end{cases}$$

Find the constant c. Are X and Y independent?

Solution. To find  $c_i$ , we have to make it so that the following holds:

$$\int_{1}^{2} \int_{0}^{2} cxy + \frac{1}{4} dy dx = 1$$

After integrating the expression, we get:

$$\frac{1}{4}cx^2y^2 = 3c + \frac{1}{2}$$

Then, we see:

$$c = \frac{1}{6}$$

Now, to see if they are independent or not, we first find the marginal distributions:

$$f_x = \int_0^2 \frac{xy}{6} + \frac{1}{4} dy$$
$$= \frac{x}{3} + \frac{1}{2}$$
$$f_y = \int_1^2 \frac{xy}{6} + \frac{1}{4} dx$$
$$= \frac{y}{4} + \frac{1}{4}$$

From here, we see that:

$$f_x f_y = \left(\frac{x}{3} + \frac{1}{2}\right) \left(\frac{y}{4} + \frac{1}{4}\right)$$

$$= \frac{xy}{12} + \frac{x}{12} + \frac{y}{8} + \frac{1}{8}$$

$$= \frac{2xy + x + 3y + 3}{24}$$

$$\neq \frac{xy}{6} + \frac{1}{4}$$

Therefore, we see that they're not independent.

For the following two questions, let Let  $X \sim \text{Exp}(3)$ 

**Problem 4.3.** Find probability that  $X \in [0, 1]$ .

*Solution.* We want to find  $\Pr[0 \le X \le 1]$ , and since X is an exponential function with mean  $\mu = 3$ , we get:

$$\int_0^1 3e^{-3x} \mathrm{d}x = 1 - e^{-3}$$

**Problem 4.4.** Let  $Y = \lfloor X \rfloor$ . For each  $k \in \mathbb{N}$ , what is the probability that Y = k? Write the distribution of Y in terms of one of the famous distributions; provide that distribution's name and parameters.

Solution. We observe that for this question, we essentially want to find  $\Pr[Y=k] = \Pr[\lfloor X \rfloor = k] = \Pr[k \le X < k+1]$ . Now, since X is an exponential function, we get the following:

$$\Pr[k \le X < k+1] = \int_{k}^{k+1} 3e^{-3x} dx$$
$$= e^{-3k} (1 - e^{-3})$$

We observe then that this is a geometric random variable, with  $p=1-e^{-3}$ . Therefore, we have that  $Y\sim {\rm Geometric}(1-e^{-3})$ .

**Problem 4.5.** Let  $X_i \sim \operatorname{Exp}(\lambda_i)$  for  $i=1,\ldots,n$  be mutually independent. It is a (very nice) fact that  $\min(X_1,\ldots,X_n) \sim \operatorname{Exp}(\mu)$ . Find  $\mu$ .

Solution. We observe that since  $X_i$  is an exponential distribution, we get:

$$\Pr[X_i \le x] = 1 - e^{-\lambda_i x}$$

From here, we observe the following:

$$\Pr[Y \le y] = 1 - \Pr[Y > y]$$

$$= 1 - \Pr[\min(X_1, \dots, X_n) > y]$$

$$= 1 - \Pr[X_1 > y, \dots, X_n > y]$$

$$= 1 - \Pr[X_1 > y] \cdots \Pr[X_n > y]$$

$$= 1 - (e^{-\lambda_1 y} \cdots e^{-\lambda_n y})$$

$$= 1 - e^{-y} \sum_{i=1}^n \lambda_i$$

Therefore, we see that  $\mu = \sum_{i=1}^{n} \lambda_i$ .

### 5 Useful Uniforms

Let X be a continuous random variable such that  $\Pr[X \in (a,b)] > 0$  for all  $a,b \in \mathbb{R}$  and a < b.

Problem 5.1. Give an example of a distribution that X could have, and one that it could not.

Solution. An example of a distribution which X could have is a normal distribution N(0,1), while a distribution which it can't have would be a uniform distribution U(0,1).

**Problem 5.2.** Show that the CDF F of X is strictly increasing. That is,  $F(x + \varepsilon) > F(x)$  for any  $\varepsilon > 0$ . Argue why this implies that  $F : \mathbb{R} \to (0,1)$  must be invertible.

Solution. Recall that the slope of the CDF is the PDF. Now, we observe that because  $\Pr[X \in (a,b)]$  is always positive, meaning that the PDF is always positive, we observe then the slope of the CDF must always be positive as well; in other words, it's strictly increasing.

Now, because the CDF is strictly increasing, it means then that it passes the horizontal line test; in other words, a horizontal line intersects the CDF in all places at exactly one point. As such, we see then that the CDF F must be invertible.

**Problem 5.3.** Let U be a uniform random variable on (0,1). What is the distribution of  $F^{-1}(U)$ ?

Solution. Keeping in mind the fact that U is a random variable, we observe the following:

$$\Pr[F^{-1}(U) \le x] = \Pr[U \le F(x)]$$
$$= \frac{F(x) - 0}{1 - 0}$$
$$= F(x)$$

**Problem 5.4.** Your work in part (c) shows that in order to sample X, it is enough to be able to sample U. If X was a discrete random variable instead, taking finitely many values, can we still use U to sample X?

Solution. No; this is because we're going from an uncountable set U to a countable set X.

# 6 Waiting For the Bus

Edward and Jerry are waiting at the bus stop outside of Soda Hall.

Like many bus systems, buses arrive in periodic intervals. However, the Berkeley bus system is unreliable, so the length of these intervals are random, and follow Exponential distributions.

Edward is waiting for the 51B, which arrives according to an Exponential distribution with parameter  $\lambda$ . That is, if we let the random variable  $X_i$  correspond to the difference between the arrival time ith and (i-1)st bus (also known as the inter-arrival time) of the 51B,  $X_i \sim \text{Exp}(\lambda)$ .

Jerry is waiting for the 79, whose inter-arrival times also follows Exponential distributions with parameter  $\mu$ . That is, if we let  $Y_i$  denote the inter-arrival time of the 79,  $Y_i \sim \text{Exp}(\mu)$ . Assume that all inter-arrival times are independent.

Problem 6.1. What is the probability that Jerry's bus arrives before Edward's bus?

Solution. We want to find  $Pr[Y_i < X_i]$ . Keeping in mind the fact that  $X_i$  and  $Y_i$  are exponential distributions which are independent of each other, we see that this is the same as asking the following:

$$\int_0^\infty \Pr[X_i = k \cap Y_i < k] dk = \int_0^\infty \Pr[X_i = k] \Pr[Y_i < k] dk$$

$$= \int_0^\infty \lambda e^{-\lambda k} \left(1 - e^{-\mu k}\right) dk$$

$$= \int_0^\infty \lambda e^{-\lambda k} dk - \int_0^\infty e^{-\mu k} dk$$

$$= 1 - \frac{\lambda}{\lambda + \mu}$$

$$= \frac{\mu}{\lambda + \mu}$$

**Problem 6.2.** After 20 minutes, the 79 arrives, and Jerry rides the bus. However, the 51B still hasn't arrived yet. Let D be the additional amount of time Edward needs to wait for the 51B to arrive. What is the distribution of D?

Solution. We observe that, by the exponential distribution's memoryless property, we get that  $D \sim \text{Exp}(\mu)$ .

**Problem 6.3.** Lavanya isn't picky, so she will wait until either the 51B or the 79 bus arrives. Find the distribution of Z, the amount of time Lavanya will wait before catching her bus.

*Solution.* Let  $Z = \min(X_i, Y_i)$ . We observe then that:

$$\begin{aligned} \Pr[Z > z] &= \Pr[\min(X_i, Y_i) > z] \\ &= \Pr[X_i > z, Y_i > z] \\ &= \Pr[X_i > z] \Pr[Y_i > z] \\ &= \left(e^{-\mu t}\right) \left(e^{-\lambda t}\right) \\ &= e^{-(\mu + \lambda)t} \end{aligned}$$

Then, from here, we see that the CDF  $\Pr[Z \leq z] = 1 - e^{-(\mu + \lambda)t}$ ; therefore, we observe that  $Z \exp{(\mu + \lambda)}$ .

**Problem 6.4.** Khalil doesn't feel like riding the bus with Edward. He decides that he will wait for the second arrival of the 51B to ride the bus. Find the distribution of  $T = X_1 + X_2$ , the amount of time that Khalil will wait to ride the bus.

Solution. To begin with, we want to find the CDF of T; that is, we want to find  $\Pr[T \leq x]$ . To do this, we do the following:

$$\Pr[T \le x] = \Pr[X_1 + X_2 \le x]$$

$$= \int_0^x \int_0^{x - x_2} \Pr[X_1 = x_1] \Pr[X_2 = x_2] dx_1 dx_2$$

$$= \int_0^x \int_0^{x - x_2} \lambda e^{-\lambda x_1} \lambda e^{-\lambda x_2} dx_1 dx_2$$

$$= 1 - e^{-\lambda x} - \lambda x e^{-\lambda x}$$

Then, to get the PDF, we differentiate our CDF, giving us the following:

$$f_t = \lambda^2 x e^{-\lambda x}, \quad t \ge 0$$

14