CS 70 Discrete Mathematics and Probability Theory
Spring 2025 Rao

HW 13

1 Estimating π

Note 17

In this problem, we discuss one way that you could probabilistically estimate π. We'll use a square dartboard of side length 2, and a circular target drawn inscribed in the square dartboard with radius 1. A dart is then thrown uniformly at random in the square. Let p be the probability that the dart lands inside the circle.

(a) What is p?

(b) Suppose we throw N darts uniformly at random in the square. Let $\hat{p}$ be the proportion of darts that land inside the circle. Create an unbiased estimator X for π using $\hat{p}$.

(c) Using Chebyshev's Inequality, compute the minimum value of N such that your estimate is within ε of π with $1-\delta$ confidence. Your answer should be in terms of ε and δ. Note that since we are estimating π, your answer should not have π in it.

Solution:

(a) The total area is 4, and the area of the circle is π. The throw is uniform, so $p = \frac{\pi}{4}$.

(b) We have that $\mathbb{E}[\hat{p}] = p = \frac{\pi}{4}$, so we also have that $\mathbb{E}[4\hat{p}] = \pi$. Thus, $X = 4\hat{p}$ is an unbiased estimator for π.

(c) We have

$$\begin{aligned} \mathbb{P}[|X-\pi| \geq \varepsilon] &= \mathbb{P}\left[\left|\hat{p}-\frac{\pi}{4}\right| \geq \frac{1}{4}\varepsilon\right] \\ &\leq \frac{\mathrm{Var}(\hat{p})}{\left(\frac{1}{4}\varepsilon\right)^2} \end{aligned}$$

by Chebyshev's Inequality and using the fact that $X = 4\hat{p}$. We want our estimate to have confidence $1-\delta$, so we want $\frac{\mathrm{Var}(\hat{p})}{\left(\frac{1}{4}\varepsilon\right)^2} < \delta$. Since $N\hat{p}$ is a $\mathrm{Binomial}(N,p)$ variable, it has variance $Np(1-p)$ and therefore $\hat{p}$ has variance $\frac{Np(1-p)}{N^2} = \frac{p(1-p)}{N}$. Since we are estimating π, we should not assume anything about the value of p in our calculations. Thus, we should use the greatest possible value of the variance, which is $\frac{1}{4}$ (when $p = \frac{1}{2}$). Then

$$\frac{\frac{p(1-p)}{N}}{\left(\frac{1}{4}\varepsilon\right)^2} < \delta \implies N > \frac{16p(1-p)}{\delta\varepsilon^2} = \frac{4}{\delta\varepsilon^2}.$$

2 Deriving the Chernoff Bound

Note 17

We've seen the Markov and Chebyshev inequalities already, but these inequalities tend to be quite loose in most cases. In this question, we'll derive the *Chernoff bound*, which is an *exponential* bound on probabilities.

The Chernoff bound is a natural extension of the Markov and Chebyshev inequalities: in Markov's inequality, we utilize only information about $\mathbb{E}[X]$; in Chebyshev's inequality, we utilize only information about $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$ (in the form of the variance). In the Chernoff bound, we'll end up using information about $\mathbb{E}[X^k]$ for *all* k, in the form of the *moment generating function* of X, defined as $\mathbb{E}[e^{tX}]$. (It can be shown that the kth derivative of the moment generating function evaluated at $t = 0$ gives $\mathbb{E}[X^k]$.)

Here, we'll derive the Chernoff bound for the binomial distribution. Suppose $X \sim \text{Binomial}(n, p)$.

(a) We'll start by computing the *moment generating function* of X. That is, what is $\mathbb{E}[e^{tX}]$ for a fixed constant $t > 0$? (Your answer should have no summations.)

Hint: It can be helpful to rewrite X as a sum of Bernoulli RVs.

(b) A useful inequality that we'll use is that

$$1 - \alpha \leq e^{-\alpha},$$

for any α. Since we'll be working a lot with exponentials here, use the above to find an upper bound for your answer in part (a) as a single exponential function. (This will make the expressions a little nicer to work with in later parts.)

(c) Use Markov's inequality to give an upper bound for $\mathbb{P}[e^{tX} \geq e^{t(1+\delta)\mu}]$, for $\mu = \mathbb{E}[X] = np$ and a constant $\delta > 0$.

Use this to deduce an upper bound on $\mathbb{P}[X \geq (1+\delta)\mu]$ for any constant $\delta > 0$. (Your bound should be a single exponential of the form $\exp(f(t))$, for a function f that should also depend on $\mu = np$ and δ.)

(d) Notice that so far, we've kept this new parameter t in our bound—the last step is to optimize this bound by choosing a value of t that minimizes our upper bound.

Take the derivative of your expression with respect to t to find the value of t that minimizes the bound. Note that from part (a), we require that $t > 0$; make sure you verify that this is the case!

Use your value of t to verify the following Chernoff bound on the binomial distribution:

$$\mathbb{P}[X \geq (1+\delta)\mu] \leq \exp(-\mu(1+\delta)\ln(1+\delta) + \delta\mu).$$

Note: As an aside, if we carried out the computations without using the bound in part (b), we'd get a better Chernoff bound, but the math is a lot uglier. Furthermore, instead of looking at the binomial distribution (i.e. the sum of independent and identical Bernoulli trials), we could have also looked at the sum of independent but not necessarily identical *Bernoulli trials as well; this would give a more general but very similar Chernoff bound.*

(e) Let's now look at how the Chernoff bound compares to the Markov and Chebyshev inequalities. Let $X \sim \text{Binomial}(n = 100, p = \frac{1}{5})$. We'd like to find $\mathbb{P}[X \geq 30]$.

  (i) Use Markov's inequality to find an upper bound on $\mathbb{P}[X \geq 30]$.

  (ii) Use Chebyshev's inequality to find an upper bound on $\mathbb{P}[X \geq 30]$.

  (iii) Use the Chernoff bound from part (d) to find an upper bound on $\mathbb{P}[X \geq 30]$.

  (iv) Now use a calculator to find the exact value of $\mathbb{P}[X \geq 30]$. How did the three bounds compare? That is, which bound was the closest and which bound was the furthest from the exact value?

(f) Let $X \sim \text{Binomial}(n = 100, p = \frac{1}{2})$. We'll look at upper bounds on the probability $\mathbb{P}[X \geq k]$ for a few values of $k > np = 50$, using Chebyshev's inequality and using the Chernoff bound, comparing the two results.

In particular, there are three regions of $k \in [51, 100]$ that are interesting to note, where the best bound swaps between Chebyshev's inequality and the Chernoff bound. Describe these three regions, and indicate which bound is best in each region (you don't need to give the exact intervals; a high level description suffices).

Solution:

(a) Note that we can write $X = \sum_{i=1}^{n} X_i$, where each X_i is an independent and identical Bernoulli trial with probability p. This means that we have

$$\begin{aligned}
\mathbb{E}[e^{tX}] &= \mathbb{E}\left[e^{t\sum_{i=1}^{n} X_i}\right] \\
&= \mathbb{E}\left[\prod_{i=1}^{n} e^{tX_i}\right] \\
&= \prod_{i=1}^{n} \mathbb{E}[e^{tX_i}] && \text{(independence)} \\
&= \prod_{i=1}^{n} \left(e^t \cdot \mathbb{P}[X_i = 1] + e^0 \cdot \mathbb{P}[X_i = 0]\right) && \text{(LOTUS)} \\
&= \prod_{i=1}^{n} \left(pe^t + 1 - p\right) \\
&= (pe^t + 1 - p)^n
\end{aligned}$$

Alternate Solution: We can also evaluate the expectation directly; using LOTUS, we have

$$\begin{aligned}
\mathbb{E}[e^{tX}] &= \sum_{k=0}^{n} e^{tk} \cdot \mathbb{P}[X=k] \\
&= \sum_{k=0}^{n} e^{tk} \cdot \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k (1-p)^{n-k} \\
&= (pe^t + 1 - p)^n
\end{aligned}$$

In the last step, we used the binomial theorem in reverse:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k},$$

for $a = pe^t$ and $b = 1-p$.

(b) With $\alpha = p - pe^t = p(1-e^t)$, we have

$$(pe^t + 1 - p)^n = (1 - p(1-e^t))^n \leq \exp(-np(1-e^t)) = \exp(-\mu(1-e^t)).$$

(c) By Markov's inequality on the RV e^{tX} (which is always nonnegative), we have

$$\begin{aligned}
\mathbb{P}[e^{tX} \geq e^{t(1+\delta)\mu}] &\leq \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu)}} \\
&\leq e^{-t(1+\delta)\mu)} e^{-\mu(1-e^t)} \\
&= \exp(-t(1+\delta)\mu - \mu(1-e^t))
\end{aligned}$$

where the second inequality comes from plugging in our answer from part (b).

As such, we have

$$\mathbb{P}[X \geq (1+\delta)\mu] = \mathbb{P}[e^{tX} \geq e^{(1+\delta)\mu}] \leq \exp(-t(1+\delta)\mu - \mu(1-e^t)).$$

(d) Taking the derivative of the exponential, we have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left[\exp(-t(1+\delta)\mu - \mu(1-e^t)\right] \\
&= \left[\exp(-t(1+\delta)\mu - \mu(1-e^t)\right] \cdot \left(-(1+\delta)\mu + \mu e^t\right)
\end{aligned}$$

This quantity is equal to zero when the last term is equal to zero (we can ignore the exponential, since it'll never be equal to 0). As such,

$$\begin{aligned}
-(1+\delta)\mu + \mu e^t &= 0 \\
\mu e^t &= (1+\delta)\mu \\
e^t &= 1+\delta \\
t &= \ln(1+\delta)
\end{aligned}$$

Since $\delta > 0$, we have that $t > 0$ as well, which satisfies our conditions on t.

Plugging this back in to our bound in part (c), we have

$$\begin{aligned}\mathbb{P}[X \geq (1+\delta)\mu] &\leq \exp(-t(1+\delta)\mu - \mu(1-e^t)) \\ &= \exp(-\mu(1+\delta)\ln(1+\delta) - \mu(1-(1+\delta))) \\ &= \exp(-\mu(1+\delta)\ln(1+\delta) + \delta\mu)\end{aligned}$$

as desired.

(e) Firstly, we'll compute a few statistics of X, which will be useful in these subparts:

$$\mathbb{E}[X] = np = 100 \cdot \frac{1}{5} = 20$$
$$\mathrm{Var}(X) = np(1-p) = 100 \cdot \frac{1}{5} \cdot \frac{4}{5} = 16$$

(i) Using Markov's inequality, we have

$$\mathbb{P}[X \geq 30] \leq \frac{\mathbb{E}[X]}{30} = \frac{20}{30} = \frac{2}{3} \approx 0.6666.$$

(ii) Using Chebyshev's inequality, we have

$$\begin{aligned}\mathbb{P}[X \geq 30] &= \mathbb{P}[X - 20 \geq 10] \\ &= \mathbb{P}[X - \mathbb{E}[X] \geq 10] \\ &\leq \mathbb{P}[|X - \mathbb{E}[X]| \geq 10] \\ &\leq \frac{\mathrm{Var}(X)}{10^2} \\ &= \frac{16}{100} = 0.16\end{aligned}$$

(iii) Using the Chernoff bound, we have

$$\begin{aligned}\mathbb{P}[X \geq 30] &= \mathbb{P}\left[X \geq \left(1 + \frac{1}{2}\right) \cdot 20\right] \\ &\leq \exp\left(-\mu\left(1+\frac{1}{2}\right)\ln\left(1+\frac{1}{2}\right) + \frac{1}{2}\mu\right) \qquad \text{(Chernoff with } \delta = \tfrac{1}{2}\text{)} \\ &= \exp(-30 \cdot \ln(1.5) + 10) \\ &\approx 0.1148\end{aligned}$$

(iv) The exact value is

$$\mathbb{P}[X \geq 30] = \sum_{k=30}^{100} \mathbb{P}[X = k] = \sum_{k=30}^{100} \binom{100}{k} \left(\frac{1}{5}\right)^k \left(\frac{4}{5}\right)^{100-k} \approx 0.01124.$$

The Chernoff bound is the closest, followed by Chebyshev's inequality, and Markov's inequality is the furthest.

As an aside, this should be expected—the Markov bound utilizes the least amount of information, while the Chernoff bound utilizes the most. In particular, Markov's inequality only requires the expectation $\mathbb{E}[X]$, Chebyshev's requires the variance (which includes information about both $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$), and the Chernoff bound requires the moment generating function (which contains information about all *moments* of X, i.e. all $\mathbb{E}[X^k]$ for $k \geq 1$).

(f) In terms of k, Chebyshev's inequality gives

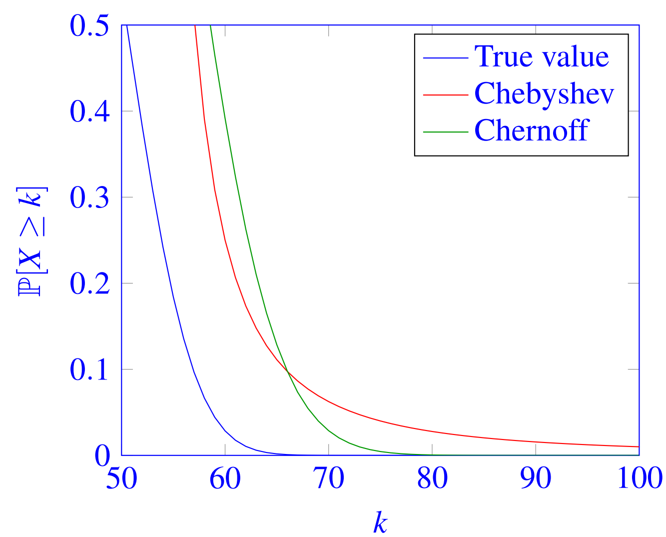
$$\mathbb{P}[X \geq k] = \mathbb{P}[X - 50 \geq k - 50] \leq \mathbb{P}[|X - 50| \geq k - 50] \leq \frac{\mathrm{Var}(X)}{(k-50)^2} = \frac{25}{(k-50)^2}.$$

Similarly, for Chernoff bound, note that we can rewrite

$$\mathbb{P}[X \geq k] = \mathbb{P}\left[X \geq \left(1 + \frac{k-50}{50}\right) \cdot 50\right],$$

so we can just plug in $\delta = \frac{k-50}{50}$ into the bound.

Using this information, we can actually plot out the results from Chebyshev's inequality and the Chernoff bound for various values of $k > np = 50$.



Here, we can see that for values of k closer to the mean, Chebyshev generally does better, while for values of k further away from the mean, Chernoff generally does better. Specifically, for $51 \leq k \leq 55$, the Chernoff bound does better; for $56 \leq k \leq 66$, Chebyshev's inequality does better; and for $k \geq 67$, the Chernoff bound does better again. (The values close to the mean are cut off from the plot, since they are significant outliers.)

In fact, the Chernoff bound does incredibly well for these larger values of k: it gives an absolute error of about 0.0005 for $k = 80$, an absolute error of about 2.5×10^{-6} for $k = 90$, and an absolute error of about 10^{-7} for $k = 95$. (In comparison, Chebyshev's inequality has an absolute error of 0.03 for $k = 80$, 0.016 for $k = 90$, and 0.0123 for $k = 95$.)

It is also worthwhile to note that although Chebyshev's inequality gives better results for values closer to the mean, there are a few values *very* close to the mean in which it gives meaningless bounds (worse than what we'd get from Chernoff); specifically, for $51 \leq k \leq 55$, Chebyshev's inequality gives probability bounds of ≥ 1, while the Chernoff bound always gives values < 1 as long as $k > 50$ (i.e. $\delta > 0$, as required).

Further note that since these specific parameters of the binomial distribution makes it symmetric, so we can actually divide Chebyshev's bound by 2 to get a tighter upper bound on the tail probability—it'll do better than the Chernoff bound for a couple more values of k, but the overall comparison still holds.

In general, the Chernoff bound is used a lot more for larger deviation bounds; in circumstances where we want to bound a tail probability that is relatively far away from the mean, the Chernoff bound gives an extraordinarily tight bound on probabilities that are very hard to compute directly.

3 Max of Uniforms

Note 21

Let $X_1, \ldots X_n$ be independent Uniform$(0,1)$ random variables, and let $X = \max(X_1, \ldots X_n)$. Compute each of the following in terms of n.

(a) What is the cdf of X?

(b) What is the pdf of X?

(c) What is $\mathbb{E}[X]$?

(d) What is $\mathrm{Var}(X)$?

Solution:

(a) $\mathbb{P}[X \leq x] = x^n$ for $0 \leq x \leq 1$ (and 0 for $x < 0$, 1 for $x > 1$), since in order for $\max(X_1, \ldots X_n) < x$, we must have $X_i < x$ for all i. Since they are independent, we can multiply together the probabilities of each of them being less than x, which is x itself, as their distributions are uniform.

(b) The pdf is the derivative of the cdf, so we have $f_X(x) = nx^{n-1}$ for $0 \leq x \leq 1$ and 0 elsewhere.

(c) To find the expectation, we integrate $x f_X(x)$ over all values of x:

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^1 x f_X(x) \\
&= \int_0^1 n x^n dx \\
&= \frac{n}{n+1}
\end{aligned}$$

(d) First, we calculate $\mathbb{E}[X^2]$, then apply the formula $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

$$\mathbb{E}[X^2] = \int_0^1 x^2 f_X(x) = \int_0^1 nx^{n+1} dx = \frac{n}{n+2}$$
$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{n}{n+2} - \frac{n^2}{(n+1)^2}$$

4 Short Answer

Note 21

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

(b) 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 (Hint: remember that the PDF must integrate to 1). Are X and Y independent?

(c) Let $X \sim \text{Exp}(3)$.

(i) Find probability that $X \in [0,\, 1]$.

(ii) Let $Y = \lfloor X \rfloor$, where the floor operator is defined as: $(\forall x \in [k, k+1))(\lfloor x \rfloor = k)$. 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.

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

Solution:

(a) Let's begin with the CDF. It will first be useful to recall that

$$F_X(t) = \mathbb{P}[X \le t] = \begin{cases} 0 & t \le 0 \\ \frac{t}{2} & t \in [0,2] \\ 1 & t \ge 2 \end{cases}.$$

Since Y is defined in terms of X, we can compute that

$$\begin{aligned}
F_Y(t) = \mathbb{P}[Y \le t] &= \mathbb{P}[4X^2 + 1 \le t] \\
&= \mathbb{P}\left[X^2 \le \frac{t-1}{4}\right] \\
&= \mathbb{P}\left[X \le \frac{1}{2}\sqrt{t-1}\right] \\
&= F_X\left(\frac{1}{2}\sqrt{t-1}\right) \\
&= \begin{cases} 0 & t \le 1 \\ \frac{1}{4}\sqrt{t-1} & t \in [1,17] \\ 1 & t \ge 17 \end{cases}
\end{aligned}$$

where in the third line we use that $X \in [0,2]$, and in the final line we have used the PDF for X. We know that the PDF can be found by taking the derivative of the CDF, so

$$f_Y(t) = \frac{\mathrm{d}}{\mathrm{d}t} F_Y(t) = \begin{cases} \frac{1}{8\sqrt{t-1}} & t \in [1,17] \\ 0 & \text{else} \end{cases}.$$

By linearity of expectation, we have $\mathbb{E}[Y] = \mathbb{E}[4X^2 + 1] = 4\,\mathbb{E}[X^2] + 1$. There are a couple ways to compute $\mathbb{E}[X^2]$.

One way is to use the fact that $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, so $\mathbb{E}[X^2] = \mathrm{Var}(X) + \mathbb{E}[X]^2$. Since $X \sim \mathrm{Uniform}[0,2]$, we know $\mathrm{Var}(X) = \frac{1}{3}$ and $\mathbb{E}[X] = 1$; this means

$$\mathbb{E}[X^2] = \mathrm{Var}(X) + \mathbb{E}[X]^2 = \frac{1}{3} + 1^2 = \frac{4}{3}.$$

Another way is to use LOTUS and integrate directly:

$$\mathbb{E}[X^2] = \int_0^2 t^2 f_X(t)\,\mathrm{d}t = \int_0^2 t^2 \cdot \frac{1}{2}\,\mathrm{d}t = \frac{1}{2}\left(\frac{1}{3}2^3\right) = \frac{4}{3}.$$

Plugging this in, we have $\mathbb{E}[Y] = 4\,\mathbb{E}[X^2] + 1 = 4 \cdot \frac{4}{3} + 1 = \frac{19}{3}$.

For the variance, we have $\mathrm{Var}(Y) = \mathrm{Var}(4X^2 + 1) = 16\,\mathrm{Var}(X^2) = 16(\mathbb{E}[X^4] - \mathbb{E}[X^2]^2)$. Here, we already know $\mathbb{E}[X^2] = \frac{4}{3}$, so we only need to compute $\mathbb{E}[X^4]$:

$$\mathbb{E}[X^4] = \int_0^2 t^4 f_X(t)\,\mathrm{d}t = \int_0^2 t^4 \cdot \frac{1}{2}\,\mathrm{d}t = \frac{1}{2}\left(\frac{1}{5}2^5\right) = \frac{16}{5}.$$

Putting this together, we have

$$\mathrm{Var}(Y) = 16(\mathbb{E}[X^4] - \mathbb{E}[X^2]^2) = 16\left(\frac{16}{5} - \frac{16}{9}\right) = \frac{1024}{45}.$$

(b) To find the correct constant, we use the fact that a PDF must integrate to one. In particular,

$$1 = \int_1^2 \int_0^2 (cxy + 1/4)\,\mathrm{d}y\,\mathrm{d}x = 3c + \frac{1}{2},$$

so $c = 1/6$. In order to check independence, we need to first find the marginal distributions of X and Y:

$$\begin{aligned} f_X(x) &= \int_0^2 f(x,y)\,\mathrm{d}y = 1/2 + x/3 \\ f_Y(y) &= \int_1^2 f(x,y)\,\mathrm{d}x = 1/4 + y/4. \end{aligned}$$

Since

$$f_X(x) f_Y(y) = \frac{1}{8} + \frac{y}{8} + \frac{x}{12} + \frac{xy}{12} \neq \frac{1}{4} + \frac{xy}{6} = f(x,y),$$

the random variables are not independent.

(c) (i) Since $X \sim \mathrm{Exp}(3)$, the CDF of X is $F(x) = 1 - e^{-3x}$. Thus we have

$$\mathbb{P}[X \in [0,1]] = \int_0^1 f(x)\,\mathrm{d}x = F(1) - F(0) = (1 - e^{-3}) - (1 - e^0) = 1 - e^{-3}.$$

(ii) Similarly, if $Y = \lfloor X \rfloor$, then $Y = k$ exactly when $X \in [k, k+1)$, so

$$\begin{aligned} \mathbb{P}[Y = k] &= \mathbb{P}[X \in [k, k+1)] \\ &= \int_k^{k+1} f(x)\,\mathrm{d}x \\ &= F(k+1) - F(k) \\ &= (1 - e^{-3(k+1)}) - (1 - e^{-3k}) \\ &= e^{-3k} - e^{-3(k+1)} \\ &= e^{-3k}(1 - e^{-3}) = (e^{-3})^k (1 - e^{-3}). \end{aligned}$$

In other words, $Y = W - 1$ for $W \sim \mathrm{Geometric}(1 - e^{-3})$.

(d) Since the X_i are independent,

$$\begin{aligned} \mathbb{P}[\min(X_1, \ldots, X_n) \le t] &= 1 - \mathbb{P}[X_1 > t, X_2 > t, \cdots \cdot X_n > t] \\ &= 1 - \mathbb{P}[X_1 > t] \cdot \mathbb{P}[X_2 > t] \cdot \cdots \cdot \mathbb{P}[X_n > t] \qquad \text{(by independence)} \\ &= 1 - e^{-\lambda_1 t} e^{-\lambda_2 t} \cdot \cdots \cdot e^{-\lambda_n t} \\ &= 1 - e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_n)t}. \end{aligned}$$

This is exactly the CDF of an $\mathrm{Exp}(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$ random variable, so $\mu = \lambda_1 + \cdots + \lambda_n$.

5 Darts with Friends

Note 21

Michelle and Alex are playing darts. Being the better player, Michelle's aim follows a uniform distribution over a disk of radius 1 around the center. Alex's aim follows a uniform distribution over a disk of radius 2 around the center.

(a) Let the distance of Michelle's throw from the center be denoted by the random variable X and let the distance of Alex's throw from the center be denoted by the random variable Y.

(i) What's the cumulative distribution function of X?

(ii) What's the cumulative distribution function of Y?

(iii) What's the probability density function of X?

(iv) What's the probability density function of Y?

(b) What's the probability that Michelle's throw is closer to the center than Alex's throw? What's the probability that Alex's throw is closer to the center?

(c) What's the cumulative distribution function of $U = \max(X,Y)$?

Solution:

(a) (i) To get the cumulative distribution function of X, we'll consider the ratio of the area where the distance to the center is less than x, compared to the entire available area. This gives us the following expression:

$$\mathbb{P}[X \le x] = \frac{\pi x^2}{\pi} = x^2, \qquad x \in [0,1].$$

(ii) Using the same approach as the previous part:

$$\mathbb{P}[Y \le y] = \frac{\pi y^2}{\pi \cdot 4} = \frac{y^2}{4}, \qquad y \in [0,2].$$

(iii) We'll take the derivative of the CDF to get the following:

$$f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}[X \le x] = 2x, \qquad x \in [0,1].$$

(iv) Using the same approach as the previous part:

$$f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y}\mathbb{P}[Y \le y] = \frac{y}{2}, \qquad y \in [0,2].$$

(b) We'll condition on Alex's outcome and then integrate over all the possibilities to get the marginal $\mathbb{P}[X \le Y]$ as following:

$$\begin{aligned}\mathbb{P}[X \le Y] &= \int_0^2 \mathbb{P}[X \le Y \mid Y = y] f_Y(y)\,\mathrm{d}y = \int_0^1 y^2 \times \frac{y}{2}\,\mathrm{d}y + \int_1^2 1 \times \frac{y}{2}\,\mathrm{d}y \\ &= \frac{1}{8} + \frac{3}{4} = \frac{7}{8}.\end{aligned}$$

Note the range within which $\mathbb{P}[X \leq Y] = 1$. This allowed us to separate the integral to simplify our solution. Using this, we can get $\mathbb{P}[Y \leq X]$ by the following:

$$\mathbb{P}[Y \leq X] = 1 - \mathbb{P}[X \leq Y] = \frac{1}{8}$$

A similar approach to the integral above could be used to verify this result:

$$\mathbb{P}[Y \leq X] = \int_0^1 \mathbb{P}[Y \leq X \mid X = x] f_X(x)\, \mathrm{d}x = \int_0^1 \frac{x^2}{4} 2x\, \mathrm{d}x = \frac{1}{2} \int_0^1 x^3\, \mathrm{d}x = \frac{1}{8}.$$

(c) Getting the CDF of U relies on the insight that for the maximum of two random variables to be smaller than a value, they both need to be smaller than that value. Using this we can get the following result for $u \in [0,1]$:

$$\mathbb{P}[U \leq u] = \mathbb{P}[X \leq u]\, \mathbb{P}[Y \leq u] = \left(u^2\right)\left(\frac{u^2}{4}\right) = \frac{u^4}{4}.$$

For $u \in [1,2]$ we have $\mathbb{P}[X \leq u] = 1$; this makes

$$\mathbb{P}[U \leq u] = \mathbb{P}[Y \leq u] = \frac{u^2}{4}.$$

For $u > 2$ we have $\mathbb{P}[U \leq u] = 1$ since CDFs of both X and Y are 1 in this range.