CS 70	Discrete Mathematics and Probability Theory	
Spring 2025	Rao	HW 13

Due: Saturday, 4/26, 4:00 PM
Grace period until Saturday, 4/26, 6:00 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

Zachary Brandt
zbrandt@berkeley.edu

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 probability that the dart lands in the circle is $p = \frac{\pi \cdot 1^2}{2^2} = \frac{\pi}{4}$.

(b) An unbiased estimator X for π using $\hat{p}$ would be $X = 4\hat{p}$. The expected value of X is $\mathbb{E}[X] = 4\,\mathbb{E}[\hat{p}] = 4p = \pi$ since, $\hat{p} = \frac{1}{N}\sum_{i=0}^{N} X_i$, where X_i is an indicator variable for if the ith dart is in the circle or not, with probability p.

(c) From Chebyshev's Inequality, $\mathbb{P}[|X-\pi| \geq \varepsilon] \leq \frac{\operatorname{Var}(X)}{\varepsilon^2}$. The variance of X is

$$\operatorname{Var}(X) = \operatorname{Var}(4\hat{p}) = 16\operatorname{Var}(\hat{p}).$$

The variance of $\hat{p}$ is

$$\text{Var}(\hat{p}) = \frac{1}{N^2}\sum_{i=1}^{N}\text{Var}(I_i) = \frac{p(1-p)}{N}.$$

The answer cannot have π in it, so, to get rid of p, the worst-case scenario for $p(1-p)$ is $\frac{1}{4}$. Then, the probability that X deviates from π by some ε is

$$\frac{\text{Var}(X)}{\varepsilon^2} = \frac{16 \times 1/4N}{\varepsilon^2} = \frac{4}{N\varepsilon^2} \leq \delta,$$

which is set to less than some δ for the confidence level. Therefore, the answer is $N \geq \frac{4}{\delta\varepsilon^2}$.

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) X can be rewritten as a sum of Bernoulli random variables, e.g., $X = X_1 + X_2 + \cdots + X_n$, where X_i is the outcome of the ith trial with probability p. Then, e^{tX} equals

$$\begin{aligned} e^{tX} &= e^{t\times(X_1+X_2+\cdots+X_n)} \\ &= e^{tX_1} \cdot e^{tX_2} \cdots. \end{aligned}$$

The expected value of which is

$$\begin{aligned} \mathbb{E}[e^{tX}] &= \mathbb{E}[e^{tX_1} \cdot e^{tX_2} \cdots] \\ \text{independence} \quad &= \mathbb{E}[e^{tX_1}] \cdot \mathbb{E}[e^{tX_2}] \cdots \\ &= (1 - p\times 1 + p\times e^t) \cdot (1 - p\times 1 + p\times e^t) \cdots \\ &= (1 - p\times 1 + p\times e^t)^n. \end{aligned}$$

(b) First, I say $\alpha = p - pe^t$. Then,

$$\begin{aligned}
1-\alpha &\leq e^{-\alpha} \\
1-p+pe^t &\leq e^{-p+pe^t} \\
(1-p+pe^t)^n &\leq e^{-np+npe^t}
\end{aligned}$$

(c) Using Markov's inequality, an upper bound is

$$\begin{aligned}
\mathbb{P}[e^{tX} \geq e^{t(1+\delta)\mu}] &\leq \frac{\mathbb{E}[e^{tx}]}{e^{t(1+\delta)\mu}} \\
&\leq \frac{(1-p+pe^t)^n}{e^{t(1+\delta)\mu}} \\
\text{upper bound} \quad &\leq \frac{e^{-np+npe^t}}{e^{t(1+\delta)\mu}} \\
&\leq \frac{e^{-\mu(1-e^t)}}{e^{t(1+\delta)\mu}}
\end{aligned}$$

This then becomes the $f(t)$ which bounds $\mathbb{P}[X \geq (1+\delta)\mu]$:

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

for $t > 0$. This is probability we solved for before and therefore $f(t) = \frac{e^{-\mu(1-e^t)}}{e^{t(1+\delta)\mu}}$.

(d) The derivative of this expression with respect to t is

$$\begin{aligned}
\frac{d}{dt}f(t) &= \frac{d}{dt}\left(e^{\mu e^t - \mu - \mu t(1+\delta)}\right) \\
&= (\mu e^t - \mu(1+\delta))e^{\mu e^t - \mu - \mu t(1+\delta)} \\
\text{minimize} \quad 0 &= (\mu e^t - \mu(1+\delta))e^{\mu e^t - \mu - \mu t(1+\delta)} \\
\text{exp never zero} \quad 0 &= (\mu e^t - \mu(1+\delta)) \\
\mu(1+\delta) &= \mu e^t \\
(1+\delta) &= e^t \\
t &= \ln(1+\delta).
\end{aligned}$$

Now using this value of t verifies

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

(e) Using Markov's inequality, the upper bound for (i) is

$$\begin{aligned}\mathbb{P}[X \geq 30] &\leq \frac{\mathbb{E}[X]}{30} \\ &\leq \frac{100 \times 0.2}{30} \\ &\leq \frac{2}{3}.\end{aligned}$$

Using Chebyshev's inequality, the upper bound for (ii) is

$$\begin{aligned}\mathbb{P}[X \geq 30] &= \mathbb{P}[X - 20 \geq 10] \\ &= \mathbb{P}[X - \mu \geq 10] \\ &\leq \frac{\text{Var}(X)}{10^2} \\ &\leq \frac{100 \times 0.2(0.8)}{100} \\ &\leq \frac{16}{100}\end{aligned}$$

Using the Chernoff bound, the upper bound for (iii) is

$$\begin{aligned}\mathbb{P}[X \geq 30] &= \mathbb{P}[X \geq (1+0.5)20] \\ \delta = 0.5 \qquad &= e^{-20(1.5)\ln(1.5)+10} \\ &= e^{-30\ln(1.5)+10} \\ &\approx 0.11487\end{aligned}$$

Using an 'exact calculator', the value is the following (from the binomial):

$$\mathbb{P}[X \geq 30] = \sum_{x=30}^{100} \mathbb{P}[X = x] = \sum_{x=30}^{100} \binom{100}{x} 0.2^x \cdot 0.8^{100-x} \approx 0.00112.$$

The Chernoff bound is closest and the Markoff bound is furthest.

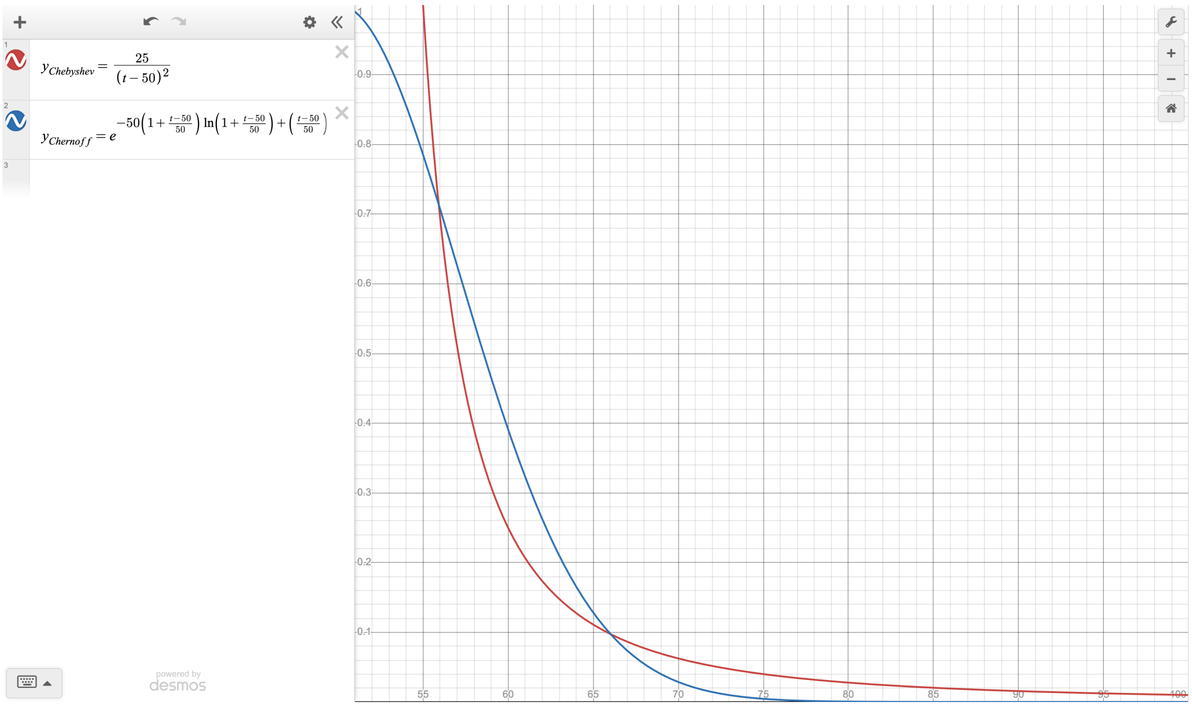

Figure 1: Part (f) graph

(f) Chebyshev is below Chernoff between roughly $56 \leq t \leq 66$, and above elsewhere. The δ value is $\frac{t-50}{50}$. For Chebyshev,

$$\mathbb{P}[X \geq t] = \mathbb{P}[X]$$

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) The cumulative distribution function of X is

$$\begin{aligned}
F_X(t) &= \mathbb{P}[X < t] \\
&= \mathbb{P}[X_1 < t] \cdot \mathbb{P}[X_2 < t] \cdots \mathbb{P}[X_n < t] \\
&= F_{X_1}(t) \cdot F_{X_2}(t) \cdots F_{X_n}(t) \\
&= \begin{cases} 0, & \text{for } t < a \\ \frac{(t-a)^n}{(b-a)^n}, & \text{for } a \le t \le b \\ 1, & \text{for } b < t \end{cases} \\
&= \begin{cases} 0, & \text{for } t < 0 \\ t^n, & \text{for } 0 \le t \le 1 \\ 1, & \text{for } 1 < t \end{cases}
\end{aligned}$$

since all of $X_1, \ldots X_n$ must be less than t for X to be less than t, and the variables are independent and uniform over (0, 1).

(b) From the fundamental thereom of calculus,

$$\begin{aligned}
\frac{d}{dt} F_X(t) &= \begin{cases} 0, & \text{for } t < 0 \\ n \cdot t^{n-1}, & \text{for } 0 \le t \le 1 \\ 0, & \text{for } 1 < t \end{cases} \\
&= f_X(t)
\end{aligned}$$

and is the probability density function.

(c) The expected value of X is

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} t \cdot f_X(t)\, dt \\
&= \int_{-\infty}^{0} 0\, dt + \int_{0}^{1} n \cdot t^{n-1}\, dt + \int_{1}^{\infty} 0\, dt \\
&= \frac{n}{n+1} (t^{n+1})\Big|_0^1 \\
&= \frac{n}{n+1}.
\end{aligned}$$

(d) To find the variance of X, I will first find the expected value of X^2, which is

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_{-\infty}^{\infty} t^2 \cdot f_X(t)\, dt \\
&= n \int_{0}^{1} t^2 \cdot t^{n-1}\, dt \\
&= \frac{n}{n+2} (t^{n+1})\Big|_0^1 \\
&= \frac{n}{n+2}.
\end{aligned}$$

Then, the variance of X is

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{n}{n+2} - \left(\frac{n}{n+1}\right)^2.$$

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 $X \sim \mathrm{Exp}(3)$.

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

Solution:

(a) The PDF and CDF of X are

$$f_X = \begin{cases} 0, & \text{for } t < 0 \\ \frac{1}{2}, & \text{for } 0 \le t \le 2 \\ 0, & \text{for } 2 < t \end{cases} \qquad F_X = \begin{cases} 0, & \text{for } t < 0 \\ \frac{t}{2}, & \text{for } 0 \le t \le 2 \\ 1, & \text{for } 2 < t \end{cases}$$

which can be used to find those for Y. The CDF of Y in terms of X is

$$\begin{aligned} F_Y(t) &= \mathbb{P}[Y < t] = \mathbb{P}[4X^2 + 1 < t] = \mathbb{P}[X < \sqrt{t-1}/2] \\ &= \begin{cases} 0, & \text{for } \sqrt{t-1}/2 < 0 \\ \frac{\sqrt{t-1}}{4}, & \text{for } 0 \le \sqrt{t-1}/2 \le 2 \\ 1, & \text{for } 2 < \sqrt{t-1}/2 \end{cases} \\ &= \begin{cases} 0, & \text{for } t < 1 \\ \frac{\sqrt{t-1}}{4}, & \text{for } 1 \le t \le 17 \\ 1, & \text{for } 17 < t \end{cases} \end{aligned}$$

from which the PDF can be found by taking the derivative:

$$\frac{d}{dt}F_Y(t) = f_Y(t)$$
$$= \begin{cases} 0, & \text{for } t < 1 \\ \frac{1}{8\sqrt{t-1}}, & \text{for } 1 \le t \le 17 \\ 0, & \text{for } 17 < t. \end{cases}$$

The expected value and the variance of Y are

$$\begin{aligned} \mathbb{E}[Y] &= \mathbb{E}[4X^2+1] = 4\,\mathbb{E}[X^2]+1 \\ &= 4\int_0^2 t^2 \cdot \frac{1}{2}\,dt + 1 \\ &= \frac{2}{3}\cdot 8 + 1 \\ &= \frac{19}{3} \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(Y) &= \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 \\ &= 16\,\mathbb{E}[X^4] + 8\,\mathbb{E}[X^2] + 1 - \left(\frac{19}{3}\right)^2 \\ &= 16\cdot\frac{32}{10} + 8\cdot\frac{4}{3} + 1 - \left(\frac{19}{3}\right)^2 \\ &\approx 22.756 \end{aligned}$$

(b) The double integral of the joint distribution evaluates to 1, which solves for the unknown c:

$$\begin{aligned} 1 &= \int_1^2 \int_0^2 (cxy + \frac{1}{4})\,dy dx \\ &= \int_1^2 (2cx + \frac{1}{2})\,dx \\ &= 4c + 1 - \left(c + \frac{1}{2}\right) \\ 1 &= 3c + \frac{1}{2} \\ c &= \frac{1}{6}. \end{aligned}$$

For independence, the product of the marginal PDFs for both X and Y must equal the joint PDF above. The marginal PDFs are below:

$$\begin{aligned} f_X &= \int_0^2 (\frac{1}{6}xy + \frac{1}{4})\,dy \\ &= (\frac{1}{12}xy^2 + \frac{1}{4}y)\Big|_0^2 \\ &= \frac{1}{3}x + \frac{1}{2} \end{aligned} \qquad \begin{aligned} f_Y &= \int_1^2 (\frac{1}{6}xy + \frac{1}{4})\,dx \\ &= (\frac{1}{12}x^2y + \frac{1}{4}x)\Big|_1^2 \\ &= \frac{1}{4}y + \frac{1}{4}. \end{aligned}$$

The product of the two is not equal to the original:

$$f_X \cdot f_Y = \left(\frac{1}{3}x + \frac{1}{2}\right)\left(\frac{1}{4}y + \frac{1}{4}\right) \neq \frac{1}{6}xy + \frac{1}{4} = f_{X,Y}$$

so they are not independent.

(c) The probability that $X \in [0,1]$ is

$$\begin{aligned}\mathbb{P}[0 \le X \le 1] &= \int_0^1 3e^{-3t}\,dt \\ &= 3 \cdot -\frac{1}{3}(e^{-3t})\Big|_0^1 \\ &= 1 - e^{-3}.\end{aligned}$$

The probability that $Y = k$ is the exact same as the probability that $X \in [k, k+1)$ or $k \le X < k+1$. So,

$$\begin{aligned}\mathbb{P}[Y = k] &= \mathbb{P}[k \le X < k+1] \\ &= \mathbb{P}[X < k+1] - \mathbb{P}[X < k] \\ &= F_X(k+1) - F_X(k) \\ &= (1 - e^{-3(k+1)}) - (1 - e^{-3k}) \\ &= e^{-3k+3} + e^{-3k} \\ &= e^{-3k}(-e^3 + 1)\end{aligned}$$

and Y is a random variable following a geometric distribution with probability of success $p = 1 - e^3$.

(d) We can find the parameter by finding the CDF of $\min(X_1, \dots, X_n)$:

$$\begin{aligned}\mathbb{P}[\min(X_1, \dots, X_n)] &= 1 - \mathbb{P}[X_1 > k, \dots, X_n > k] \\ &= 1 - \mathbb{P}[X_1 > k] \cdots \mathbb{P}[X_n > k] \\ &= 1 - e^{-\lambda_1 k} \cdots e^{-\lambda_n k} \\ &= 1 - e^{-(\lambda_1 + \cdots + \lambda_n)k}\end{aligned}$$

]

which takes on the same form as $F_X(t) = 1 - e^{-\lambda t}$, and 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) The CDF of X is $F_X(r) = \mathbb{P}[X < r] = \frac{\pi r^2}{\pi \cdot 1^2} = r^2$. The CDF of Y is $F_Y(r) = \mathbb{P}[Y < r] = \frac{\pi r^2}{\pi \cdot 4} = \frac{r^2}{4}$. The PDF of X is then $2r$, and the PDF of Y is then $\frac{r}{2}$.

(b) The probability that Michelle throws closer to Alex is

$$\begin{aligned}
\mathbb{P}[X < Y] &= \int_0^2 \mathbb{P}[X < Y \mid Y = y]\,\mathbb{P}[Y = y]\,dy \\
&= \int_0^1 y^2 \cdot \frac{y}{2}\,dy + \int_1^2 1 \cdot \frac{y}{2}\,dy \\
&= \frac{1}{2} \cdot \frac{1}{4} + \frac{1}{2} \cdot \left(2 - \frac{1}{2}\right) \\
&= \frac{7}{8}
\end{aligned}$$

since it is guaranteed Michelle's throw will be closer than Alex's after $r = 1$, and before it is the CDF given Alex's throw. The probability that Alex's throw is closer to the center is the complement of this probability so $\mathbb{P}[Y < X] = \frac{1}{8}$.

(c) The CDF of U is $\mathbb{P}[U < r] = \mathbb{P}[X < r, Y < r]$ since both need to be less than to evade the maximum selector function. So, $\mathbb{P}[X < r, Y < r] = \mathbb{P}[X < r] \cdot \mathbb{P}[Y < r]$ given independence. Although this varies depending on what r is in question, i.e.,

$$F_U(r) = \begin{cases} r^2 \times \frac{r^2}{4}, & \text{for } r < 1 \\ \frac{r^2}{4}, & \text{for } 1 \leq r \leq 2 \\ 1, & \text{for } 2 < r \end{cases}$$