

1 Estimating π

Note 19

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.

- What is the value of p ?
- 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} (in other words, create a random variable X such that $\mathbb{E}[X] = \pi$).
- Using Chebyshev's inequality, find a value of N such that your estimate X 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, nor use any bounds on the value of π .

Solution:

- The total area is 4, and the area of the circle is π . The throw is uniform, so $p = \frac{\pi}{4}$.
- 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 .
- We have

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

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 $\frac{\text{Var}(\hat{p})}{\left(\frac{1}{4}\varepsilon\right)^2} < \delta$. Since $N\hat{p}$ is a 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. However, it is true that $p(1-p)$ is bounded above by $1/4$ (which is achieved when $p = 1/2$), so we have

$$\frac{p(1-p)}{\left(\frac{1}{4}\varepsilon\right)^2} < \delta,$$

which is true as long as

$$N > \frac{16p(1-p)}{\delta\epsilon^2} = \frac{4}{\delta\epsilon^2}.$$

2 Tightness of Inequalities

Note 17

- (a) Give an example where Markov's inequality is tight; that is, show that given any fixed $k > 0$, there exists a discrete non-negative random variable X which is **not always zero** such that $\mathbb{P}[X \geq k] = \mathbb{E}[X]/k$.
- (b) Give an example where Chebyshev's inequality is tight; that is, show that given any fixed $k \geq 1$, there exists a random variable X such that $\mathbb{P}[|X - \mathbb{E}[X]| \geq k\sigma] = 1/k^2$, where $\sigma^2 = \text{Var}(X)$.

Solution:

- (a) In the proof of Markov's Inequality ($\mathbb{P}[X \geq \alpha] \leq \frac{\mathbb{E}[X]}{\alpha}$), the first time we lose equality is at this step:

$$\mathbb{E}[X] = \sum_a (a \cdot \mathbb{P}[X = a]) \geq \sum_{a \geq \alpha} (a \cdot \mathbb{P}[X = a])$$

We get an inequality because we drop all $a \cdot \mathbb{P}[X = a]$ terms where $a < \alpha$. Thus, we can only maintain equality if all of these dropped terms were actually 0. This would mean either $a = 0$ or $\mathbb{P}[X = a] = 0$ for $a > 0$, which means X can put probability on 0, but should put no probability on any other value $< \alpha$.

The next time we lose equality in the proof is the step following the one above:

$$\sum_{a \geq \alpha} (a \cdot \mathbb{P}[X = a]) \geq \alpha \cdot \sum_{a \geq \alpha} \mathbb{P}[X = a]$$

We get an inequality because we treat all $a \geq \alpha$ in the summation as just α , so we can pull out the α term. The only way for us to maintain equality is if we never have to substitute α for some larger a . This tells us that X should not put probability on any value $> \alpha$.

Both of these facts drive the intuition behind our example: that X can only take values 0 and α .

Let X be the random variable which is 0 with probability $1 - p$ and k with probability p , where $k > 0$. Then, $\mathbb{E}[X] = kp$, and Markov's inequality says

$$\mathbb{P}(X \geq k) \leq \frac{\mathbb{E}[X]}{k} = \frac{kp}{k} = p,$$

which is tight.

Note that for a complete solution, any probability $p < 1$ gives a valid example.

- (b) The proof of Chebyshev's Inequality ($\mathbb{P}[|X - \mathbb{E}[X]| \geq \alpha] \leq \frac{\text{Var}(X)}{\alpha^2}$) comes from an application of Markov's Inequality to the variable $Y = (X - \mathbb{E}[X])^2$ being $\geq \alpha^2$. The only ways we can

lose equality in the proof of Chebyshev's is if we lose equality in the application of Markov! Therefore, we need the variable Y to satisfy the conditions from Part (a) that ensure the application of Markov will be tight. To recap, we would need Y to only take values 0 and α^2 . Thus, $(X - \mathbb{E}[X])$ can take on the values $\{-\alpha, 0, \alpha\}$.

Let

$$X = \begin{cases} -a & \text{with probability } k^{-2}/2 \\ a & \text{with probability } k^{-2}/2 \\ 0 & \text{with probability } 1 - k^{-2} \end{cases}$$

for $a > 0$. Note that $\mathbb{E}[X] = 0$ and $\text{Var}(X) = a^2k^{-2}$, so $k\sigma = a$, so Chebyshev's inequality gives

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq k\sigma) = \mathbb{P}(|X| \geq a) \leq \frac{1}{k^2},$$

which is tight.

There are other correct examples as well, which are all linear transformations of the example above.

3 Max of Uniforms

Note 20

Let X_1, \dots, X_n be independent $\text{Uniform}(0, 1)$ random variables (i.e., they are each uniformly distributed on $[0, 1]$), and let $X = \max(X_1, \dots, X_n)$. Compute each of the following in terms of n .

- What is the cdf of X ?
- What is the pdf of X ?
- What is $\mathbb{E}[X]$?
- What is $\text{Var}(X)$?

Solution:

- $\mathbb{P}[X \leq x] = x^n$ for $0 \leq x \leq 1$ (and 0 for $x < 0$, 1 for $x > 1$), since in order for $\max(X_1, \dots, X_n) < x$, 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.
- 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.
- To find the expectation, we integrate $xf_X(x)$ over all values of x :

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

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

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

4 Short Answer

Note 20

(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 the 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 $i = 1, \dots, n$ be mutually independent. It is a (very nice) fact that $\min(X_1, \dots, X_n) \sim \text{Exp}(\mu)$. Find μ .

Solution:

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

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

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

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

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

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

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 $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, so $\mathbb{E}[X^2] = \text{Var}(X) + \mathbb{E}[X]^2$. Since $X \sim \text{Uniform}[0, 2]$, we know $\text{Var}(X) = \frac{1}{3}$ and $\mathbb{E}[X] = 1$; this means

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

Another way is to use LOTUS and integrate directly:

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

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

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

Putting this together, we have

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

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

$$1 = \int_1^2 \int_0^2 (cxy + 1/4) dy dx = 3c + \frac{1}{2},$$

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

$$f_X(x) = \int_0^2 f(x,y) dy = 1/2 + x/3$$

$$f_Y(y) = \int_1^2 f(x,y) dx = 1/4 + y/4.$$

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

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

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

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

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

(d) Since the X_i are independent,

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

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

5 Darts with Friends

Note 20

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 (independent of Michelle's distribution).

- (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:

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

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

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

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

$$f_X(x) = \frac{d}{dx} \mathbb{P}[X \leq x] = 2x, \quad x \in [0, 1].$$

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

$$f_Y(y) = \frac{d}{dy} \mathbb{P}[Y \leq y] = \frac{y}{2}, \quad y \in [0, 2].$$

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

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

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:

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

- (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 $u \in [0, 1]$:

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

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.