CS 70 Discrete Mathematics and Probability Theory
Fall 2021 HW 14

1 Final Exam Format

Please fill out this form to choose how you will take the final exam.

Solution: Please fill out the form here.

2 Continuous Intro

(a) Is

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

a valid density function? Why or why not? Is it a valid CDF? Why or why not?

(b) Calculate $\mathbb{E}[X]$ and $\mathrm{Var}(X)$ for X with the density function

$$f(x) = \begin{cases} 1/\ell, & 0 \le x \le \ell, \\ 0, & \text{otherwise.} \end{cases}$$

(c) Suppose X and Y are independent and have densities

$$f_X(x) = \begin{cases} 2x, & 0 \le x \le 1, \\ 0, & \text{otherwise,} \end{cases}$$

$$f_Y(y) = \begin{cases} 1, & 0 \le y \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

What is their joint distribution? (Hint: for this part and the next, we can use independence in much the same way that we did in discrete probability)

(d) Calculate $\mathbb{E}[XY]$ for the above X and Y.

Solution:

(a) Yes; it is non-negative and integrates to 1. No; a CDF should go to 1 as x goes to infinity and be non-decreasing.

(b) $\mathbb{E}[X] = \int_{x=0}^{\ell} x \cdot (1/\ell)\,\mathrm{d}x = \ell/2$. $\mathbb{E}[X^2] = \int_{x=0}^{\ell} x^2 \cdot (1/\ell)\,\mathrm{d}x = \ell^2/3$.
$\mathrm{Var}(X) = \ell^2/3 - \ell^2/4 = \ell^2/12$.

This is known as the continuous uniform distribution over the interval $[0,\ell]$, sometimes denoted $\mathrm{Uniform}[0,\ell]$.

(c) Note that due to independence,

$$\begin{aligned} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \mathbb{P}(X \in [x, x+\mathrm{d}x], Y \in [y, y+\mathrm{d}y]) = \mathbb{P}(X \in [x, x+\mathrm{d}x])\mathbb{P}(Y \in [y, y+\mathrm{d}y]) \\ &\approx f_X(x) f_Y(y)\,\mathrm{d}x\,\mathrm{d}y \end{aligned}$$

so their joint distribution is $f(x,y) = 2x$ on the unit square $0 \le x \le 1$, $0 \le y \le 1$.

(d) $\mathbb{E}[XY] = \int_{x=0}^{1}\int_{y=0}^{1} xy \cdot 2x\,\mathrm{d}y\,\mathrm{d}x = \int_{x=0}^{1} x^2\,\mathrm{d}x = 1/3$.
Alternatively, since X and Y are independent, we can compute $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$. Note that

$$\mathbb{E}[X] = \int_0^1 x \cdot 2x\,\mathrm{d}x = \frac{2}{3}x^3\bigg|_0^1 = \frac{2}{3},$$

and $\mathbb{E}[Y] = 1/2$ since the density of Y is symmetric around $1/2$. Hence,

$$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y] = \frac{1}{3}.$$

3 Max of Uniforms

Let $X_1, \ldots X_n$ be independent $U[0,1]$ 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 $Var[X]$?

Solution:

(a) $Pr[X \le x] = x^n$ 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) Taking the derivative of the cdf, we have $f_X(x) = nx^{n-1}$

(c)

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

(d)

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

4 Darts with Friends

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.

- What's the cumulative distribution function of X?
- What's the cumulative distribution function of Y?
- What's the probability density function of X?
- 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 $U = \max\{X, Y\}$?

(d) What's the cumulative distribution function of $V = \min\{X, Y\}$?

(e) What is the expectation of the absolute difference between Michelle's and Alex's distances from the center, that is, what is $\mathbb{E}[|X - Y|]$? [*Hint*: Use parts (c) and (d), together with the continuous version of the tail sum formula, which states that $\mathbb{E}[Z] = \int_0^\infty P(Z \geq z) dz$.]

Solution:

(a) • 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:

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

• Using the same approach as the previous part:

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

• We'll take the derivative of the CDF to get the following:

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

• Using the same approach as the previous part:

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

(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 $\mathbb{P}(X \le Y) = 1$. This allowed us to separate the integral to simplify our solution. Using this, we can get $\mathbb{P}(Y \le X)$ by the following:

$$\mathbb{P}(Y \le X) = 1 - \mathbb{P}(X \le Y) = \frac{1}{8}$$

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

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

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

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

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

$$\mathbb{P}(U \le u) = \mathbb{P}(Y \le u) = \frac{u^2}{4}.$$

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

(d) Getting the CDF of V relies on a similar insight that for the minimum of two random variables to be greater than a value, they both need to be greater than that value. Taking the complement of this will give us the CDF of V. This allows us to get the following result. For $v \in [0,1]$:

$$\begin{aligned}\mathbb{P}(V \le v) &= 1 - \mathbb{P}(V \ge v) = 1 - \mathbb{P}(X \ge v)\mathbb{P}(Y \ge v) = 1 - \big(1 - \mathbb{P}(X \le v)\big)\big(1 - \mathbb{P}(Y \le v)\big) \\ &= 1 - \left(1 - v^2\right)\left(1 - \frac{v^2}{4}\right) = \frac{5v^2}{4} - \frac{v^4}{4}.\end{aligned}$$

For $v > 1$, we get $\mathbb{P}(X > v) = 0$, making $\mathbb{P}(V \le v) = 1$.

(e) We can subtract V from U to get this difference. Using the tail-sum formula to calculate the expectation, we can get the following result:

$$\begin{aligned}\mathbb{E}[|X - Y|] = \mathbb{E}[U - V] = \mathbb{E}[U] - \mathbb{E}[V] &= \int_0^2 \mathbb{P}(U \ge u)\,\mathrm{d}u - \int_0^1 \mathbb{P}(V \ge v)\,\mathrm{d}v \\ &= \int_0^1 \left(1 - \frac{u^4}{4}\right)\mathrm{d}u + \int_1^2 \left(1 - \frac{u^2}{4}\right)\mathrm{d}u - \int_0^1 \left(1 - \frac{5v^2}{4} + \frac{v^4}{4}\right)\mathrm{d}v \\ &= \frac{19}{20} + \frac{5}{12} - \frac{19}{30} = \frac{11}{15}.\end{aligned}$$

Alternatively, you could derive the density of U and V and use those to calculate the expectation. For $u \in [0,1]$:

$$f_U(u) = \frac{\mathrm{d}\mathbb{P}(U \le u)}{\mathrm{d}u} = u^3.$$

For $u \in [1,2]$:

$$f_U(u) = \frac{\mathrm{d}\mathbb{P}(U \le u)}{\mathrm{d}u} = \frac{u}{2}.$$

Using this we can calculate $\mathbb{E}[U]$ as:

$$\mathbb{E}[U] = \int_0^2 u f_U(u)\,\mathrm{d}u = \int_0^1 u^4\,\mathrm{d}u + \frac{1}{2}\int_1^2 u^2\,\mathrm{d}u = \frac{1}{5} + \frac{7}{6} = \frac{41}{30}.$$

To calculate $\mathbb{E}[V]$ we will use the following PDF for $v \in [0,1]$:

$$f_V(v) = \frac{\mathrm{d}\mathbb{P}(V \le v)}{\mathrm{d}v} = \frac{5v}{2} - v^3.$$

We can get the $\mathbb{E}[V]$ by the following:

$$\mathbb{E}[V] = \int_0^1 v f_V(v)\,\mathrm{d}v = \int_0^1 \left(\frac{5v^2}{2} - v^4\right)\mathrm{d}v = \frac{5}{6} - \frac{1}{5} = \frac{19}{30}.$$

Combining the two results gives us the same result as above:

$$\mathbb{E}[|X - Y|] = \mathbb{E}[U - V] = \mathbb{E}[U] - \mathbb{E}[V] = \frac{41}{30} - \frac{19}{30} = \frac{11}{15}.$$

5 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 λ. 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, $X_i \sim \text{Expo}(\lambda)$.

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

(a) What is the probability that Jerry's bus arrives before Edward's bus?

(b) 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?

(c) 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.

(d) 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:

(a) Let f_{Y_i} be the pdf of Y_i. By total probability,

$$\begin{aligned}
\mathbb{P}(X_i > Y_i) &= \int_{t=0}^{\infty} f_{Y_i}(t) \cdot \mathbb{P}(X_i > Y_i | Y_i = t)\,\mathrm{d}t \\
&= \int_{t=0}^{\infty} f_{Y_i}(t) \cdot \mathbb{P}(X_i > t)\,\mathrm{d}t \\
&= \int_{t=0}^{\infty} f_{Y_i}(t) \cdot (1 - F_{X_i}(t))\,\mathrm{d}t \\
&= \int_{t=0}^{\infty} \mu e^{-\mu t}(e^{-\lambda t})\,\mathrm{d}t \\
&= \mu \int_{t=0}^{\infty} e^{-(\lambda+\mu)t}\,\mathrm{d}t \\
&= \frac{\mu}{\lambda+\mu} \int_{t=0}^{\infty} (\lambda+\mu)e^{-(\lambda+\mu)t}\,\mathrm{d}t \\
&= \frac{\mu}{\lambda+\mu},
\end{aligned}$$

where the integral in the second-to-last line evaluates to 1, since it is the total integral of the Exponential$(\lambda+\mu)$ density.

(b) We observe that $\mathbb{P}(D > d) = \mathbb{P}(X > 20 + d \mid X \geq 20)$. Then, we apply Bayes Rule:

$$\begin{aligned}
\mathbb{P}(X > 20 + d \mid X \geq 20) &= \frac{\mathbb{P}(X > 20 + d)}{\mathbb{P}(X \geq 20)} \\
&= \frac{1 - F_X(20 + d)}{1 - F_X(20)} \\
&= \frac{e^{-\lambda(20+d)}}{e^{-20\lambda}} \\
&= e^{-\lambda d}
\end{aligned}$$

Thus, the CDF of D is given by $\mathbb{P}(D \leq d) = 1 - \mathbb{P}(D > d) = 1 - e^{-\lambda d}$. This is the CDF of an exponential, so D is exponentially distributed with parameter λ.

One can also directly apply the memoryless property of the exponential distribution to arrive at this answer.

(c) Lavanya's waiting time is the minimum of the time it takes for the 51B and the time it takes for the 79 to arrive. Thus, $Z = \min(X, Y)$.

$$\begin{aligned}
\mathbb{P}(Z > t) &= \mathbb{P}(X > t \cap Y > t) \\
&= \mathbb{P}(X > t) \cdot \mathbb{P}(Y > t) \\
&= (1 - F_X(t))(1 - F_Y(t)) \\
&= (1 - (1 - e^{-\mu t}))(1 - (1 - e^{-\lambda t})) \\
&= e^{-\mu t} e^{-\lambda t} \\
&= e^{-(\mu+\lambda)t}
\end{aligned}$$

It follows that the CDF is Z, $\mathbb{P}(Z \leq t) = 1 - e^{-(\mu+\lambda)t}$. Thus, Z is exponentially distributed with parameter $\mu + \lambda$.

(d) Let $t > 0$. By total probability,

$$\begin{aligned}
\mathbb{P}(T \le t) &= \mathbb{P}(X_1 + X_2 \le t) \\
&= \int_0^\infty \mathbb{P}(X_1 + X_2 \le t \mid X_1 \in \mathrm{d}x) \cdot \mathbb{P}(X_1 \in \mathrm{d}x) \\
&= \int_0^t \mathbb{P}(X_1 + X_2 \le t \mid X_1 \in \mathrm{d}x) \cdot \mathbb{P}(X_1 \in \mathrm{d}x) + \int_t^\infty 0 \cdot \mathbb{P}(X_1 \in \mathrm{d}x) \\
&= \int_0^t \mathbb{P}(X_2 \le t - X_1 \mid X_1 \in \mathrm{d}x) \cdot \mathbb{P}(X_1 \in \mathrm{d}x) + 0 \\
&= \int_0^t \mathbb{P}(X_2 \le t - x) \cdot \mathbb{P}(X_1 \in \mathrm{d}x) \\
&= \int_0^t F_{X_2}(t - x) \cdot f_{X_1}(x)\,\mathrm{d}x \\
&= \int_0^t \left(1 - e^{-\lambda(t-x)}\right) \cdot \lambda e^{-\lambda x}\,\mathrm{d}x \\
&= \int_0^t \lambda e^{-\lambda x} - \lambda e^{-\lambda t}\,\mathrm{d}x \\
&= \int_0^t \lambda e^{-\lambda x} - \lambda e^{-\lambda t} \int_0^t \mathrm{d}x \\
&= F_{X_1}(t) - \lambda e^{-\lambda t} \cdot t \\
&= 1 - e^{-\lambda t} - \lambda t e^{-\lambda t}
\end{aligned}$$

Upon differentiating the CDF, we have

$$\begin{aligned}
f_T(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}(T \le t) = \lambda e^{-\lambda t} - \lambda e^{-\lambda t} + \lambda^2 t e^{-\lambda t} \\
&= \lambda^2 t e^{-\lambda t}, \qquad t > 0.
\end{aligned}$$

6 Chebyshev's Inequality vs. Central Limit Theorem

Let n be a positive integer. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with the following distribution:

$$\mathbb{P}[X_i = -1] = \frac{1}{12}; \qquad \mathbb{P}[X_i = 1] = \frac{9}{12}; \qquad \mathbb{P}[X_i = 2] = \frac{2}{12}.$$

(a) Calculate the expectations and variances of X_1, $\sum_{i=1}^n X_i$, $\sum_{i=1}^n (X_i - \mathbb{E}[X_i])$, and

$$Z_n = \frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{\sqrt{n/2}}.$$

(b) Use Chebyshev's Inequality to find an upper bound b for $\mathbb{P}[|Z_n| \ge 2]$.

(c) Can you use b to bound $\mathbb{P}[Z_n \geq 2]$ and $\mathbb{P}[Z_n \leq -2]$?

(d) As $n \to \infty$, what is the distribution of Z_n?

(e) We know that if $Z \sim \mathcal{N}(0,1)$, then $\mathbb{P}[|Z| \leq 2] = \Phi(2) - \Phi(-2) \approx 0.9545$. As $n \to \infty$, can you provide approximations for $\mathbb{P}[Z_n \geq 2]$ and $\mathbb{P}[Z_n \leq -2]$?

Solution:

(a) $\mathbb{E}[X_1] = -1/12 + 9/12 + 4/12 = 1$, and

$$\operatorname{Var} X_1 = \frac{1}{12} \cdot 2^2 + \frac{9}{12} \cdot 0^2 + \frac{2}{12} \cdot 1^2 = \frac{1}{2}.$$

Using linearity of expectation and variance (since $X_1, \ldots, X_n$ are independent), we find that $\mathbb{E}[\sum_{i=1}^n X_i] = n$ and $\operatorname{var}(\sum_{i=1}^n X_i) = n/2$.

Again, by linearity of expectation, $\mathbb{E}[\sum_{i=1}^n X_i - n] = n - n = 0$. Subtracting a constant does not change the variance, so $\operatorname{var}(\sum_{i=1}^n X_i - n) = n/2$, as before.

Using the scaling properties of the expectation and variance, $\mathbb{E}[Z_n] = 0/\sqrt{n/2} = 0$ and $\operatorname{Var} Z_n = (n/2)/(n/2) = 1$.

(b)

$$\mathbb{P}[|Z_n| \geq 2] \leq \frac{\operatorname{Var} Z_n}{2^2} = \frac{1}{4}$$

(c) $1/4$ for both, since $\mathbb{P}[Z_n \geq 2] \leq \mathbb{P}[|Z_n| \geq 2]$ and $\mathbb{P}[Z_n \leq -2] \leq \mathbb{P}[|Z_n| \geq 2]$.

(d) By the Central Limit Theorem, we know that $Z_n \to \mathcal{N}(0,1)$, the standard normal distribution.

(e) Since $Z_n \to \mathcal{N}(0,1)$, we can approximate $\mathbb{P}[|Z_n| \geq 2] \approx 1 - 0.9545 = 0.0455$. By the symmetry of the normal distribution, $\mathbb{P}[Z_n \geq 2] = \mathbb{P}[Z_n \leq -2] \approx 0.0455/2 = 0.02275$.

It is interesting to note that the CLT provides a much smaller answer than Chebyshev. This is due to the fact that the CLT is applied to a particular kind of random variable, namely the (scaled) sum of a bunch of random variables. Chebyshev's inequality, however, holds for any random variable, and is therefore weaker.