Problemset 13

Albert Ye

November 25, 2022

1 Continuous Intro

(a) It is a valid density function, since all values of $f(x)$ are nonnegative, and

$$\begin{aligned}\int_{-\infty}^{\infty} f(x) &= \int_{-\infty}^{0} 0 + \int_{0}^{1} 2x + \int_{1}^{\infty} 0 \\ &= 0 + (1^2 - 0^2) + 0 = 1.\end{aligned}$$

However, it isn't a valid CDF since $\lim_{x\to\infty} f(x) = 0$, while it should equal 1.

(b)

$$\begin{aligned}f_x(x) &= \frac{d}{dx} F_x(x) \\ &= \begin{cases} \frac{1}{l} & \text{for } 0 \leq x \leq l \\ 0 & \text{otherwise} \end{cases}.\end{aligned}$$

(c) Because X, Y are independent, we see that

$$\begin{aligned}f(x,y)dxdy &= \mathbb{P}[X \in [x, x+dx], Y \in [y, y+dy]] \\ &= \mathbb{P}[X \in [x, x+dx]]\mathbb{P}[Y \in [y, y+dy]] \\ &\approx f_X(x) f_Y(y) dxdy.\end{aligned}$$

Therefore, the joint distribution is $f(x,y) = f_x(x)f_y(y) = 2x$ for $0 \leq x, y \leq 1$.

(d)

$$\begin{aligned}\mathbb{E}[XY] &= \int_0^1 \int_0^1 xyf(x,y)dxdy \\ &= \int_0^1 \int_0^1 2x^2 ydxdy \\ &= \int_0^1 (\frac{2}{3}(1)^3 y) - 0dy \\ &= \int_0^1 \frac{2}{3} ydy \\ &= (\frac{2}{3}\frac{1}{2}(1)) - 0 \\ &= \boxed{\frac{1}{3}}.\end{aligned}$$

2 Lunch Meeting

As the distribution of Alice and Bob's arrival times are uniform, we can think of the distribution as a uniform square with opposing corners $(0,0)$ and $(1,1)$. Then, if Alice arrives first at time a, she will leave at time $a+0.25$, and if Bob arrives first at time b, then he will leave at time $b+0.25$.

Therefore, we would want $\mathbb{P}[|a-b|<0.25]$, where a,b are Alice and Bob's times of arrival, respectively. Ignoring the constraints of $0 \le a,b \le 1$, we would have a total probability of $\int_0^1 0.50da = 0.50$. However, we need to remove cases of $b<0$ or $b>1$, which happens with probability

$$\int_0^{0.25} 0.25-a\,da + \int_{0.75}^1 a-0.75\,da = \left(0.25a - \frac{a^2}{2}\right|_0^{0.25} + \left(\frac{a^2}{2}-0.75a\right|_{0.75}^1 = \frac{1}{16}-\frac{1}{32}+\frac{1}{16}-\frac{1}{32} = \frac{1}{16}.$$

Therefore, our probability is $\frac{1}{2}-\frac{1}{16} = \boxed{\frac{7}{16}}$.

3 Darts with Friends

(a) The cumulative distribution function $F_X(x)$ for X is the probability of getting the dart in a disk of radius x. The larger disk has an area of π and the smaller disk has an area of πx^2, so the cumulative distribution function $\boxed{F_X(x)=x^2}$. Similarly, the CDF $F_Y(y)$ for Y is the probability of getting the dart in a disk of radius y, but the larger disk has an area of 4π. Therefore, $\boxed{F_Y(y)=\frac{y^2}{4}}$.

Using this, we can get $f_X(x) = \frac{d}{dx}F_X(x) = 2x$ and $f_Y(y) = \frac{d}{dx}F_Y(y) = \frac{y}{2}$.

(b) Because X,Y are independent, the probability that $X=x, Y=y$ is $f_X(x)f_Y(y)$. So

$$\begin{aligned}
\mathbb{P}[X \le Y] &= \int_0^1 \int_x^2 f_X(x)f_Y(y)dydx \\
&= \int_0^1 \left(f_X(x)F_Y(y)\right|_x^2 dx \\
&= \int_0^1 f_X(x)\left(1-\frac{x^2}{4}\right)dx \\
&= \int_0^1 2x\left(1-\frac{x^3}{2}\right) \\
&= \left(x^2-\frac{x^2}{8}\right|_0^1 \\
&= \boxed{\frac{7}{8}}.
\end{aligned}$$

The probability that Alex's throw is closer to the center must thus be $1-\frac{7}{8} = \boxed{\frac{1}{8}}$.

(c) Note that $\mathbb{P}[U \le k] = \mathbb{P}[X \le k \cap Y \le k] = \mathbb{P}[X \le k]\mathbb{P}[Y \le k]$ because X,Y are independent. Therefore, this equals $f_X(k)f_Y(k) = \frac{k^4}{4}$ assuming $0 \le k \le 1$. If $k<0$, then $\mathbb{P}[U \le k]=0$. If $1 \le k \le 2$, then $\mathbb{P}[\max(X,Y) \le k] = \mathbb{P}[Y \le k] = \frac{k^2}{4}$. And if $k>2$, then $\mathbb{P}[U \le k]=1$. Therefore, the cumulative distribution function for U is

$$F_U(c_k) = \begin{cases} 0 & \text{for } c_k \le 0 \\ \frac{c_k^4}{4} & \text{for } 0 < c_k \le 1 \\ \frac{c_k^2}{4} & \text{for } 1 < k \le 2 \\ 1 & \text{for } c_k > 2. \end{cases}$$

(d) Note that $\mathbb{P}[V \leq k] = \mathbb{P}[X \leq k] + \mathbb{P}[Y \leq k] - \mathbb{P}[X \leq k \cap Y \leq k]$. The third term is just $F_U(k)$, so this because $\mathbb{P}[X \leq k] + \mathbb{P}[Y \leq k] + F_U(k)$. For $k \leq 0$, we have that $\mathbb{P}[V \leq k] = 0$ as before. For $k > 1$, we have that $\mathbb{P}[V \leq k] = 1$ because $X \leq 1$. So all that's left to consider is $0 \leq k \leq 1$, where $F_U(k) = \frac{k^4}{4}$. So in this range, $\mathbb{P}[V \leq k] = F_X(k) + F_Y(k) - F_U(k) = k^2 + \frac{k^2}{4} - \frac{k^4}{4} = \frac{5}{4}k^2 + \frac{k^4}{4}$.

Therefore,

$$F_V(k) = \begin{cases} 0 & \text{for } k < 0 \\ \frac{5}{4}k^2 - \frac{k^4}{4} & \text{for } 0 \leq k \leq 1 \\ 1 & \text{for } k > 1 \end{cases} .$$

(e) Note that $\mathbb{E}[|X - Y|] = \mathbb{E}[U - V]$ for U, V from parts (c) and (d).We can use the tail sum formula to get this expected value.

$$\begin{aligned}
\mathbb{E}[U - V] = \mathbb{E}[U] - \mathbb{E}[V] &= \int_0^\infty y\,\mathrm{pr}[U \geq u]du - \int_0^\infty v\mathbb{P}[V \geq v]dv \\
&= \int_0^\infty u(1 - F_U(u))du - \int_0^\infty v(1 - F_V(v))dv \\
&= \int_0^1 u(1 - \frac{u^4}{4})du + \int_1^2 u(1 - \frac{u^2}{4})du - \int_0^1 v(1 - \frac{5}{4}v^2 + \frac{v^4}{4})dv \\
&= \int_0^1 u - \frac{u^5}{4} + \int_1^2 u - \frac{u^3}{4}du - \int_0^1 v - \frac{5}{4}v^3 + \frac{v^5}{4}dv \\
&= \left(\frac{u^2}{2} - \frac{u^6}{4}du\right|_0^1 \left(\frac{u^2}{2} - \frac{u^4}{16}\right|_1^2 - \left(\frac{v^2}{2} - \frac{5}{16}v^4 + \frac{v^6}{24}\right|_0^1 \\
&= \frac{1}{2} - \frac{1}{24} + \frac{4}{2} - \frac{16}{16} - \frac{1}{2} + \frac{1}{16} - \frac{1}{2} + \frac{5}{16} - \frac{1}{24} \\
&= \frac{1}{2} + \frac{3}{8} - \frac{1}{2} = \frac{12 + 9 - 2}{24} = \boxed{\frac{19}{24}}.
\end{aligned}$$

4 Waiting for the Bus

(a) We want to find $\mathbb{P}[Y_1 \leq X_1]$, or, in other words,

$$\int_0^\infty \int_0^x f(x)F(y)dydx,$$

where $f(x)$ is the probability distribution function at $X = x$ and $F(y)$ is the cumulative distribution function $Y \leq y$. We find that $f(x) = \lambda e^{-\lambda i}$ by definition, and

$$\begin{aligned} F(y) = \mathbb{P}[Y \leq y] &= \int_0^x \mu e^{-\mu y} dy \\ &= \left(-e^{-\mu x}\right|_0^x = (-e^{-\mu x}) - (-e^{\mu}(0)) \\ &= 1 - e^{-\mu x}. \end{aligned}$$

From this, we can find that

$$\begin{aligned} \mathbb{P}[Y_1 \leq X_1] &= \int_0^\infty \lambda e^{-\lambda x}(1 - e^{-\mu x})dx \\ &= \int_0^\infty \lambda e^{-\lambda x} - \lambda e^{-(\lambda+\mu)x} dx \\ &= \left(e^{-\lambda x}\right|_0^\infty - \left(\frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)x}\right|_0^\infty \\ &= (1-0) - \left(\frac{\lambda}{\lambda+\mu} - 0\right) \\ &= \boxed{\frac{\mu}{\lambda+\mu}}. \end{aligned}$$

(b) We claim that the distribution of D is identical to the distribution of X.

Note that we want to find $\mathbb{P}[D = k]$, or essentially, that we want to find the probability of $X = 20 + k$ given that $X \geq 20$. Obviously, note that if $k < 0$ then the probability is just 0. However, if $k \geq 0$, then we have $\mathbb{P}[D = k] = \frac{\mathbb{P}[X_i = 20+k]}{\int_{20}^\infty \mathbb{P}[X_i = x]dx}$. The top is evidently equal to $\lambda e^{-\lambda(20+k)}$ from our initial formula, and the bottom must equal

$$\begin{aligned} \int_{20}^\infty \mathbb{P}[X = x]dx &= \int_{20}^\infty \lambda e^{-\lambda x} dx \\ &= \left(e^{-\lambda x}\right|_{20}^\infty \\ &= 0 - (-e^2 0\lambda) = e^{-20\lambda}. \end{aligned}$$

Thus, we can find $\mathbb{P}[D = k]$ to be

$$\frac{\lambda e^{-\lambda k + 20}}{e^{-20\lambda}} = \lambda e^{\lambda(20-(k+20)} = \boxed{\lambda e^{-\lambda k}},$$

which is equal to the exponential distribution for λ. □

(c) We have

$$\begin{aligned} \mathbb{P}[Z \leq k] &= \mathbb{P}[\min(X, Y) \leq k] \\ &= \mathbb{P}[X \leq k \cup Y \leq k] \\ &= \mathbb{P}[X \leq k] + \mathbb{P}[Y \leq k] - \mathbb{P}[X \leq k \cap Y \leq k]. \end{aligned}$$

Because X and Y are indepedent, we have that

$$\begin{aligned} \mathbb{P}[Z \leq k] &= \mathbb{P}[X \leq k] + \mathbb{P}[Y \leq k] - \mathbb{P}[X \leq k]\mathbb{P}[Y \leq k] \\ &= (1 - e^{-\lambda k}) + (1 - e^{-\mu k}) - (1 - e^{-\lambda k})(1 - e^{-\mu k}) \\ &= 2 - e^{-\lambda k} - e^{-\mu k} - 1 + e^{-\lambda k} + e^{-\mu k} - e^{-\lambda k} e^{-\mu k} \\ &= 1 - e^{-(\lambda+\mu)k}. \end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{P}[Z = k] = \frac{d}{dk}\mathbb{P}[Z \le k] &= \frac{d}{dk} 1 - e^{-(\lambda+\mu)k} \\ &= (\lambda + \mu)e^{\lambda+\mu}.\end{aligned}$$

This means that $\boxed{Z \sim \text{Expo}(\lambda + \mu).}$

(d) We can think of $T = k$ as $X_1 = x$ and $X_2 = k - x$. We will have to integrate over all possible such x.

$$\begin{aligned}\mathbb{P}[T = k] &= \int_0^k \mathbb{P}[X_1 = x]\mathbb{P}[X_2 = k - x]dx \\ &= \int_0^k (\lambda e^{-\lambda x})(\lambda e^{-\lambda(k-x)})dx \\ &= \left(x\lambda^2 e^{-\lambda k}\right|_0^k \\ &= k\lambda^2 e^{-\lambda k}\end{aligned}$$

5 Chebyshev's Inequality vs. Central Limit Theorem

(a) •

$$\begin{aligned}\mathbb{E}[X_1] &= -1\left(\frac{1}{12}\right) + 1\left(\frac{9}{12}\right) + 2\left(\frac{2}{12}\right)\\ &= -\frac{1}{12} + \frac{9}{12} + \frac{4}{12} = \boxed{1}.\\ \text{Var}(X_1) &= \mathbb{E}[(X - \mathbb{E}(X))^2]\\ &= (-2)^2\left(\frac{1}{12}\right) + (0)^2\left(\frac{9}{12}\right) + (1)^2\left(\frac{2}{12}\right)\\ &= \frac{4}{12} + \frac{2}{12} = \frac{6}{12} = \boxed{\frac{1}{2}}.\end{aligned}$$

• Let Y be $\sum_{i=0}^{n} X_i$.

$$\mathbb{E}[Y] = \sum_{i=0}^{n} \mathbb{E}[X_i] = \sum_{i=0}^{n} 1 = \boxed{n}.$$

$$\text{Var}(Y) = \sum_{i=0}^{n} \text{Var}(X_i) = \sum_{i=0}^{n} \frac{1}{2} = \boxed{\frac{n}{2}}.$$

•

$$\mathbb{E}\left[\sum_{i=0}^{n} X_i - \mathbb{E}[X_i]\right] = \mathbb{E}[Y] - \mathbb{E}[\mathbb{E}[X_i] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n]] = n - \mathbb{E}[n] = \boxed{0}.$$

$$\text{Var}\left(\sum_{i=0}^{n} X_i - \mathbb{E}[X_i]\right) = \mathbb{E}\left[\sum_{i=0}^{n}(X_i - \mathbb{E}[X_i] - 0)^2\right] = \text{Var}(Y) = \boxed{\frac{n}{2}}.$$

• We know that $\mathbb{E}[Z_n] = 0$ because it is simply the previous expression divided by $\sqrt{n/2}$. Moreover, $\text{Var}(Z_n) = 1$ because it is once again the previous expression divided by $\sqrt{n/2}$, so we must divide the variance by $\frac{n}{2}$.

(b) From Chebyshev's Inequality, we have

$$\begin{aligned}\mathbb{P}[|Y - \mathbb{E}[Y]| \geq 2] &\leq \frac{\text{Var}(Y)}{c^2}\\ \mathbb{P}\left[\left|\frac{Y}{\sqrt{n/2}}\right| \geq 2\right] &\leq \frac{\frac{\text{Var}(Y)}{n/2}}{4}\\ \mathbb{P}[|Z_n| \geq 2] &\leq \frac{\frac{n/2}{n/2}}{4}\\ \mathbb{P}[|Z_n] \geq 2] &\leq \frac{1}{4}.\end{aligned}$$

Therefore, $b = \boxed{\frac{1}{4}}$.

(c) No, because there is no guarantee that Z_n is symmetric.

(d) From the Central Limit Theorem, it becomes the normal distribution $N(0, 1)$.

(e) Yes, because a normal distribution is symmetric and as $n \to \infty$ Z_n is normal.