EECS 126 Probability and Random Processes University of California, Berkeley: Spring 2019
Kannan Ramchandran February 19, 2019

Midterm 1

Last Name	First Name	SID
Zhang	Albert	3034019226

Rules.

- No form of collaboration between the students is allowed. If you are caught cheating, you may fail the course and face disciplinary consequences.
- You have 10 minutes to read the exam and 90 minutes to complete it.
- The exam is not open book; we are giving you a cheat sheet. No calculators or phones allowed.
- Unless otherwise stated, all your answers need to be justified. Show all your work to get partial credit.
- Maximum you can score is 114 but 100 points is considered perfect.

Problem	points earned	out of
Problem 1		54
Problem 2		10
Problem 3		25
Problem 4		25
Total		114

Problem 1: Answer these questions briefly but clearly.

(a) [6] Show that if A and B are events with $\mathbb{P}(A) = \mathbb{P}(B) = 1/2$, then $\mathbb{P}(A \cap B) = \mathbb{P}(A^c \cap B^c)$

The probability of $A^c \cap A^c$ is 1 minus the probability that A or B happens. In particular,

$$\mathbb{P}[A^c \cap A^c] = 1 - \mathbb{P}[A \cup B] = 1 - \underbrace{\left(\overset{1/2}{\mathbb{P}[A]} + \overset{1/2}{\mathbb{P}[B]} - \mathbb{P}[A \cap B]\right)}_{\text{principle, inclusion, exclusion}} = \mathbb{P}[A \cap B].$$

(b) [6] Bob plays the following game: He tosses independent coins (which have probability of heads as p) until he gets a heads. Let N be the number of coins he tosses. He then re-tosses all those coins, and counts the number of heads among them, let that be X. What is $\mathbb{E}X$? Write your final answer in the box.

$N \sim \text{geom}(p)$, $B = \text{Bernoulli}(p)$

$$\mathbb{E}[X] = \mathbb{E}[NB] \underset{\substack{\uparrow \\ \text{since } N \\ \text{and } B \text{ are} \\ \text{independent}}}{=} \mathbb{E}[N]\mathbb{E}[B] = \frac{1}{p} \cdot p = 1$$

$\boxed{1}$

(c) [6] Find the distribution of Y (denoted $f_Y(y)$ for $y \in (-1,1)$) if X is uniform on $[-\frac{3\pi}{2}, \frac{3\pi}{2}]$ and $Y = \sin(X)$. Note that the inverse of the sine function ($\arcsin(x) : [-1,1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$) has a derivative of $\frac{1}{\sqrt{1-x^2}}$.

$$F_Y(t) = \mathbb{P}[\sin(X) \le t] = \mathbb{P}\left[X \in \left[-\frac{\pi}{2} - \arcsin(t), -\frac{\pi}{2}\right] \cup \left[-\frac{\pi}{2}, \arcsin(t)\right] \cup \left[\pi - \arcsin(t), \frac{3\pi}{2}\right]\right]$$

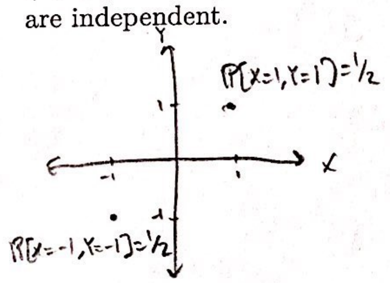
These 3 intervals are all the same size, and since X is uniform, we have:

$$3 \cdot \mathbb{P}\left[X \in \left[-\frac{\pi}{2}, \arcsin(t)\right]\right] = 3\left(\frac{\arcsin(t) - (-\frac{\pi}{2})}{\frac{3\pi}{2} - (-\frac{3\pi}{2})}\right) = 3\left(\frac{\arcsin(t) + \pi/2}{3\pi}\right)$$

$$= \frac{\arcsin(t) + \pi/2}{\pi}$$

So $f_Y(t) = \frac{dF_Y}{dt}(t) = \boxed{\frac{1}{\pi\sqrt{1-t^2}}}$ for $t \in (-1,1)$.

(d) [6] Find a pair of random variables X, Y such that X, Y are dependent, but X^2 and Y^2 are independent.

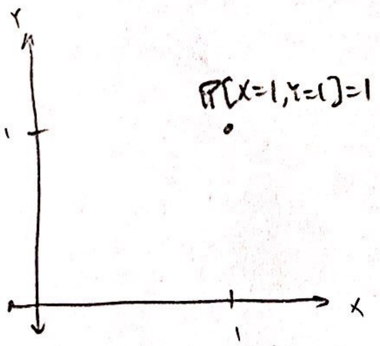


Consider R.V.'s X and Y to the left.

$$X = \begin{cases} 1 & \text{prob. } 1/2 \\ -1 & \text{prob. } 1/2 \end{cases} \quad \text{and} \quad Y = X.$$

$\mathbb{P}[X=1, Y=1] = 1/2 \neq 1/4 = \mathbb{P}[X=1] \cdot \mathbb{P}[Y=1]$

So X and Y are dependent.



But X^2 and Y^2 are independent, since both take on the value 1 with probability 1, and

$$\mathbb{P}[X^2=1, Y^2=1] = 1 = \mathbb{P}[X^2=1] \cdot \mathbb{P}[Y^2=1].$$

(e) **[6]** There are N (≥ 2) servers behind a counter at a store, having independent service times that are exponentially distributed with rates $\lambda_1, \lambda_2, \ldots, \lambda_N$ respectively. There is a single line (queue) for service to all N servers. Alice and Bob are customers who come into the store when all servers are busy serving their respective customers. Alice is first in line to be served next, followed by Bob. What is the probability that Alice and Bob get served by different servers? What are the expected departure times for Alice and Bob?

Part 1: We want $1-\mathbb{P}[A \text{ and } B \text{ same server}] = 1-\sum_{i=1}^{N}\mathbb{P}[A=\text{Server } i]\,\mathbb{P}[B=\text{server } i \mid A=\text{server } i]$

$$= 1-\sum_{i=1}^{N}\frac{\lambda_i}{(\lambda_1+\cdots+\lambda_N)}\cdot\frac{\lambda_i}{(\lambda_1+\cdots+\lambda_N)}$$

$$= \boxed{1-\frac{\lambda_1^2+\cdots+\lambda_N^2}{(\lambda_1+\cdots+\lambda_N)^2}}$$

Part 2: $\mathbb{E}[\text{Alice}] = \sum_{i=1}^{N}\mathbb{P}[A=\text{server } i]\left(\mathbb{E}[\text{Exp}(\lambda_i)]+\mathbb{E}[\text{Exp}(\lambda_i)]\right) = \sum_{i=1}^{N}\frac{2}{\lambda_1+\cdots+\lambda_N} = \boxed{\frac{2N}{\lambda_1+\cdots+\lambda_N}}$

$\mathbb{E}[\text{Bob}] = \mathbb{E}[\text{time of 1st server}]+\mathbb{E}[\text{time of 2nd server}] = \sum_{i=1}^{N}\frac{\lambda_i}{\lambda_1+\cdots+\lambda_N}\cdot\frac{1}{\lambda_i} + \sum_{i=1}^{N}\frac{\lambda_i}{\lambda_1+\cdots+\lambda_N}\cdot\frac{1}{\lambda_i}$

$$= \boxed{\frac{2N}{\lambda_1+\cdots+\lambda_N}}$$

(f) **[6]** I have 10 independent draws from a U(0,1) distribution (uniform distribution between 0 and 1). What is the expected difference between the largest and the smallest draws? What is the probability that the third largest draw is smaller than 0.8 given that the largest draw is 0.9?

Part 1: $f_{X^{(1)}}(x) = 10\binom{9}{0}\frac{1}{1}\cdot x^0(1-x)^9 = 10(1-x)^9$

$f_{X^{(10)}}(x) = 10\binom{9}{9}\frac{1}{1}\cdot x^9(1-x)^0 = 10x^9$

$$\mathbb{E}[X^{(10)}-X^{(1)}] = \mathbb{E}[X^{(10)}]-\mathbb{E}[X^{(1)}] = \overbrace{\int_0^1 x\,10x^9\,dx}^{\left(\frac{10x^{11}}{11}\right)\Big|_{0\text{ to }1} = \frac{10}{11}} - \underbrace{\int_0^1 x\,10(1-x)^9\,dx}_{-x(1-x)^{10}\Big|_{0\text{ to }1} - \int_0^1 -(1-x)^{10}dx = \left(\frac{-(1-x)^{11}}{11}\right)_{0\text{ to }1} = \frac{1}{11}} = \boxed{\frac{9}{11}}$$

Part 2: $\mathbb{P}[X^{(8)}<0.8 \mid X^{(10)}=0.9]$; if we truncate the largest draw, we are left with 9 independent draws from $U(0, 0.9)$

so we want $\sum_{k=8}^{9}\binom{9}{k}\left(\frac{8}{9}\right)^k\left(\frac{1}{9}\right)^{9-k} = 9\left(\frac{8}{9}\right)^8\left(\frac{1}{9}\right)+1\left(\frac{8}{9}\right)^9\left(\frac{1}{9}\right)^0 = \boxed{\left(\frac{8}{9}\right)^8+\left(\frac{8}{9}\right)^9}$

(g) [6] You have N items, G of which are good and B of which are bad. You start to draw items without replacement, and suppose that the first good item appears on draw X. Find $\mathbb{P}(X > k)$.

We see that the sample space is uniform in the arrangements of G's and B's.
In particular, $\mathbb{P}[X>k] = \mathbb{P}[\text{first } k \text{ items are bad}]$

$$= \frac{\#\text{ ways to choose } k \text{ B's in beginning}}{\#\text{ ways to choose any } k \text{ in beginning}} = \begin{cases} \dfrac{\binom{B}{k}}{\binom{N}{k}} & \text{if } k \leq B, \\ 0 & \text{if } k > B \end{cases}$$

Alternatively, we also have $\mathbb{P}[X>k] = \dfrac{\#\text{ ways to arrange G in last N-k}}{\#\text{ ways to arrange G anywhere}}$

$$= \begin{cases} \dfrac{\binom{N-k}{G}}{\binom{N}{G}} & \text{if } N-k \geq G, \\ 0 & \text{else.} \end{cases}$$

We can easily verify that both these answers are actually the same numerically.

(h) [6] Show that for a random variable that is upper bounded by B, the following probability bound holds:

$$P(X \leq t) \leq \frac{B - \mathbb{E}[X]}{B - t} \quad \text{for all } t \leq B$$

We have that $\mathbb{E}[X] \leq B(1 - \mathbb{P}[X \leq t]) + t\mathbb{P}[X \leq t]$ since B is an upper bound on values in the interval $(t, B]$ and t is an upper bound on the values in the interval $(-\infty, t]$.

Thus, since $t \leq B$, we get that

$$\mathbb{E}[X] \leq B - (B-t)\mathbb{P}[X \leq t]$$
$$\Rightarrow (B-t)\mathbb{P}[X \leq t] \leq B - \mathbb{E}[X]$$
$$\Rightarrow \mathbb{P}[X \leq t] \leq \frac{B - \mathbb{E}[X]}{B-t}$$

(i) [6] Consider the r.v.s $Y \sim \mathcal{N}(0,1)$ and $Z \sim Pois(\lambda)$.

Calculate $\mathbb{E}[Y^Z]$

$$\mathbb{E}\left[Y^z\right] = \int_{-\infty}^{\infty} \frac{e^{-y^2/2}}{\sqrt{2\pi}} \sum_{z=0}^{\infty} \frac{e^{-\lambda}(y\lambda)^z}{z!}\, dy$$

$$= \int_{-\infty}^{\infty} \frac{e^{-y^2/2}}{\sqrt{2\pi}}\, e^{-\lambda}\left(e^{y\lambda}\right) dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2}\, e^{\lambda(y-1)}\, dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y-\lambda)^2 + \frac{\lambda^2}{2} - \lambda}\, dy$$

$$= \frac{e^{\frac{\lambda^2}{2}-\lambda}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(y-\lambda)^2}{2}}\, dy \leftarrow \text{just } \mathcal{N}(\lambda,1)$$

$$= \frac{e^{\frac{\lambda^2}{2}-\lambda}}{\sqrt{2\pi}} \cdot \sqrt{2\pi}$$

$$= \boxed{e^{\frac{\lambda^2}{2}-\lambda}}$$

Problem 2: Basic probability

(a) [2] Say you have n balls and n buckets numbered 1 to n. Now, you are going to put one ball in each bucket by picking a permutation uniformly at random. Let X be the number of balls that have the same number as their bin. What is $\mathbb{E}X$? Write your final answer in the box.

$$\mathbb{E}[X] \underset{\text{by linearity}}{=} \mathbb{E}[I_1]+\cdots+\mathbb{E}[I_n] = n\left(\frac{1}{n}\right) = 1$$

$$\boxed{1}$$

(b) [3] Same set-up, but now instead you toss the balls uniformly at random at each bucket (now buckets can have more than one ball). Let Y be the number of balls that have the same number as their bin. What is $\mathbb{P}(Y = 0)$? Write your final answer in the box.

$\mathbb{P}[Y=0] = \left(\frac{n-1}{n}\right)^n$ since each toss is independent

$$\boxed{\left(\frac{n-1}{n}\right)^n}$$

(c) [5] Compare the variances of X and Y. Which one is bigger? Justify briefly but rigorously. (You do not need to compute the variances of X and Y to answer this problem.)

$\operatorname{Var} X > \operatorname{Var} Y$, splitting X and Y into indicator RVs $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ for each ball, then

$$\operatorname{Var} X = \sum_{i=1}^{n} \operatorname{Var} X_i + 2\underbrace{\sum_{k<j} \operatorname{cov} X_k X_j}_{>0 \text{ since } X_k \text{ not indep from } X_j}$$

$$||$$

$$\operatorname{Var} Y = \sum_{i=1}^{n} \operatorname{Var} Y_i + 2\underbrace{\sum_{k<j} \operatorname{cov} Y_k Y_j}_{=0 \text{ since } Y_k \text{ indep from } Y_j}$$

Thus $\operatorname{Var} X > \operatorname{Var} Y$ since the second term for $\operatorname{Var} X$ is >0 whereas the second term for $\operatorname{Var} Y$ is $=0$.

Problem 3: Apples!

You are starting a business of shipping apples. On each day, you get X requests, where X is a Poisson random variable with a rate of 6 requests/day (so $\lambda = 6$).

(a) [3] For each request, you make a box and put Y_i apples in it where Y_i is uniform in the set $\{1, 2, ..., 6\}$ (and i goes from 1 to X) and all Y_is are independent of each other. What is the distribution of the number of boxes that you ship that have 6 apples in it?

$$\mathbb{P}[N=k] = \sum_{n=k}^{\infty} \mathbb{P}[X=n] \cdot \mathbb{P}[N=k \mid X=n] = \sum_{n=k}^{\infty} \binom{n}{k}\left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k} e^{-6} \frac{6^n}{n!}$$

$$= \frac{e^{-6}}{k!} \sum_{n=k}^{\infty} \frac{5^{n-k}}{(n-k)!}$$

$$= \frac{e^{-6}}{k!} \sum_{i=0}^{\infty} \frac{5^i}{i!}$$

$$= \frac{e^{-6}}{k!} \cdot e^5$$

$$= \frac{e^{-1} \cdot 1^k}{k!} \Rightarrow N \sim \boxed{\text{Poisson}(1)}$$

(b) [5] When your customers start complaining that they got a different number of apples than their neighbor, you decide to change your business model a little bit. Every day, you decide to toss a fair 6 sided die, and put that many apples inside each of the shipments. (So, if you have 10 requests, and you roll a 4, you are going to give away 40 apples total.)

Now you start to get scared: It is possible that you give away way too many apples because on some day, you roll big and you have a lot of requests. More formally, let Z be the number of apples you give away. Using a Markov bound, upper bound the probability that Z is greater than or equal to 70.

$$\mathbb{P}[Z \geq 70] \leq \frac{\mathbb{E}[Z]}{70} = \frac{\mathbb{E}[X \cdot Y_1]}{70} \underset{\uparrow}{=} \frac{\mathbb{E}[X]\mathbb{E}[Y_1]}{70} = \frac{\lambda \cdot 7/2}{70} = \boxed{\frac{3}{10}}$$

follows from

$X \perp\!\!\!\perp Y_1$

(c) [5] Using the Chebyshev bound, find the ~~probability~~ bound that $Z \geq 70$? (For simplicity, you can use the fact that the variance of a die roll comes down to $\frac{35}{12}$)

$$\mathbb{P}[Z \geq 70] = \mathbb{P}[Z - \mathbb{E}[Z] \geq 70 - \mathbb{E}[Z]]$$
$$= \mathbb{P}[Z - 21 \geq 49]$$
$$\leq \mathbb{P}[|Z-21| \geq 49]$$
$$\leq \frac{Var(Z)}{49^2}$$

$\mathbb{E}[Z] = 21$ from part B.

$$Var(Z) = \mathbb{E}[Var(Z|X)] + Var(\mathbb{E}[Z|X])$$
$$= \mathbb{E}[X^2 \cdot Var(Y)] + Var(X\mathbb{E}[Y])$$
$$= \frac{35}{12}\mathbb{E}[X^2] + \left(\frac{7}{2}\right)^2 Var(X)$$
$$= \frac{35}{\cancel{12}_{2}} \cdot \cancel{42}^{7} + \frac{49}{\cancel{4}_{2}} \cdot \cancel{6}^{3}$$
$$= \frac{392}{2}$$
$$= 196$$

so $\mathbb{P}[Z \geq 70] \leq \frac{196}{49^2} = \boxed{\frac{4}{49}}$

to find $\mathbb{E}[X^2]$:

$$\mathbb{E}[X(X-1)] = \sum_{k=0}^{\infty} \frac{e^{-6}6^k}{k!} \cdot k(k-1)$$
$$= \sum_{k=2}^{\infty} \frac{e^{-6}6^k}{(k-2)!}$$
$$= 36 \sum_{k=0}^{\infty} \frac{e^{-6}6^k}{k!}$$
$$= 36 \cdot 1 = 36$$

$$\mathbb{E}[X^2] - \mathbb{E}[X] = 36$$
$$\Rightarrow \mathbb{E}[X^2] = 36 + \mathbb{E}[X] = 42$$

$\mathbb{E}[X] = \lambda = 6$

(d) [5] Now, say that you are a customer of this business instead. You receive orders from this business every day; how many days will it take in expectation for you to receive all 6 different types of boxes? Leave it as a sum of fractions.

X_i = # of boxes after the i-1th new type to get the ith type.

$$\mathbb{E}[\text{days}] = \mathbb{E}[\#\text{ boxes it takes}] = \mathbb{E}[X_1 + \cdots + X_6]$$

$$= \sum_{i=1}^{6} \mathbb{E}[X_i]$$

$$= \sum_{i=0}^{5} \frac{6}{6-i} = \boxed{6\left(\frac{1}{6}+\frac{1}{5}+\frac{1}{4}+\frac{1}{3}+\frac{1}{2}+\frac{1}{1}\right)}$$

(e) [7] Say that you are trying to deduce how many apples were shipped in your box based on how long the shipment takes. In particular, say that the arrival time of a box containing k apples is exponential with parameter k per hour (So, on average, it takes $\frac{1}{k}$ hours for a box of size k apples to arrive). What is the probability that your box contains 5 apples given that you have already waited 5 hours?

$T \sim \text{Exp}(k)$ 5

$$\mathbb{P}[X=5 \mid T>5] = \frac{\mathbb{P}[T>5 \mid X=5]\cdot\mathbb{P}[X=5]}{\mathbb{P}[T>5]}$$

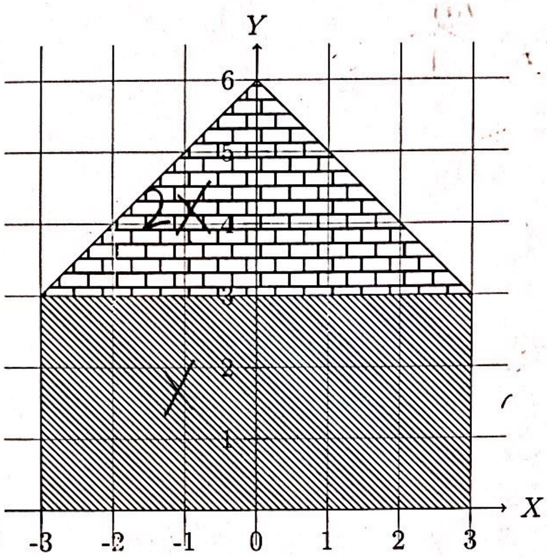
$$= \frac{\left(\int_5^\infty 5e^{-5x}dx\right)\frac{1}{6}}{\frac{1}{6}\sum_{k=1}^{6}\int_5^\infty ke^{-kx}dx}$$

$$= \frac{e^{-25}\cdot\cancel{\frac{1}{6}}}{\cancel{\frac{1}{6}}\sum_{k=1}^{6} e^{-5k}}$$

$$= \boxed{\frac{e^{-25}}{e^{-5}+e^{-10}+e^{-15}+e^{-20}+e^{-25}+e^{-30}}} = \frac{e^{-20}-e^{-25}}{1-e^{-30}}$$

Problem 4: Chimney construction

Shown below is a density profile of a house from which you are going to sample a point (X, Y). *Note that the density $f_{(X,Y)}$ of the brick roof ($\mathcal{B}$) is twice that of the wall of the house ($\mathcal{W}$) (check the shading).*



(a) [5] What is the joint PDF of (X, Y)?

$[\text{roof}] = \frac{6 \cdot 3}{2} = 9$

$[\text{wall}] = 6 \cdot 3 = 18$

$9 \cdot 2x + 18 \cdot x = 1$

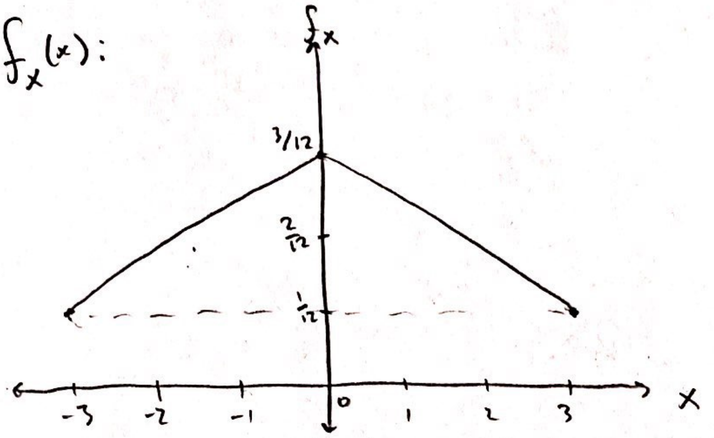
$36x = 1$

$x = \frac{1}{36}, \quad 2x = \frac{1}{18}$

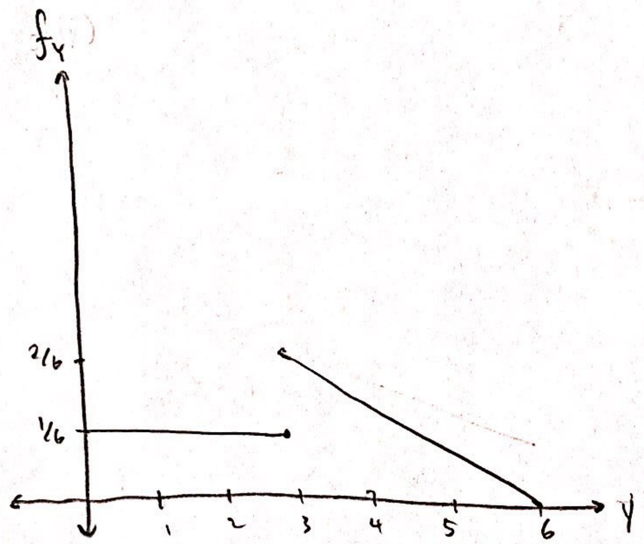
$$\Rightarrow f_{X,Y}(x,y) = \begin{cases} \frac{1}{18} & \text{for } (x,y) \in \mathcal{B}, \\ \frac{1}{36} & \text{for } (x,y) \in \mathcal{W} \end{cases}$$

(b) [10] Plot the marginal PDFs of $f(x)$ and $f(y)$.

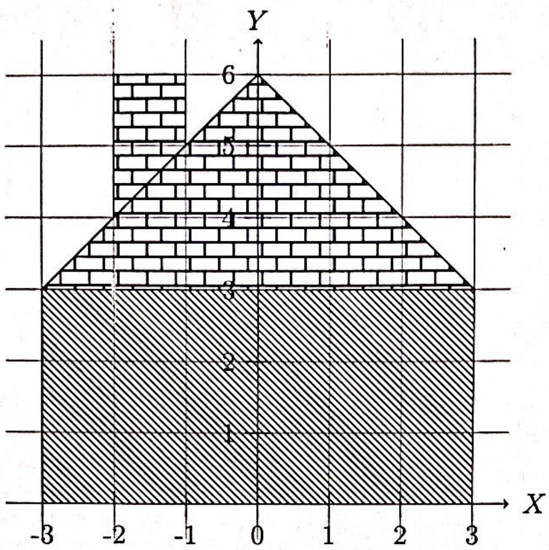
$f_X(x)$:



$f_Y(y)$:

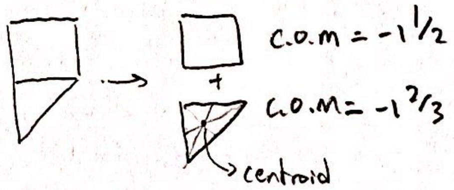


Now, you want to install a little chimney to your house as follows:



(c) **[10]** Find the expectation of X for the house with chimney.

$[\text{chimney} + \text{roof}] = 9 + \frac{3}{2} = \frac{21}{2}$



$\frac{21}{2} \cdot 2x + 18x = 1$

$39x = 1$

$x = \frac{1}{39}$

$$\mathbb{E}[x] = \frac{1}{39} \cdot 18 \cdot 0 + \frac{2}{39} \cdot 9 \cdot 0 + \frac{2}{39}[\text{Chimney}] \cdot (\text{center of mass of chimney})$$

$$= \frac{2}{39} \cdot \frac{3}{2} \cdot \left(\frac{2}{3}\left(-1\tfrac{1}{2}\right) + \frac{1}{3}\left(-1\tfrac{2}{3}\right)\right)$$

$$= \frac{1}{13}\left(-1 - \frac{5}{9}\right) = \boxed{\frac{-14}{117}}$$