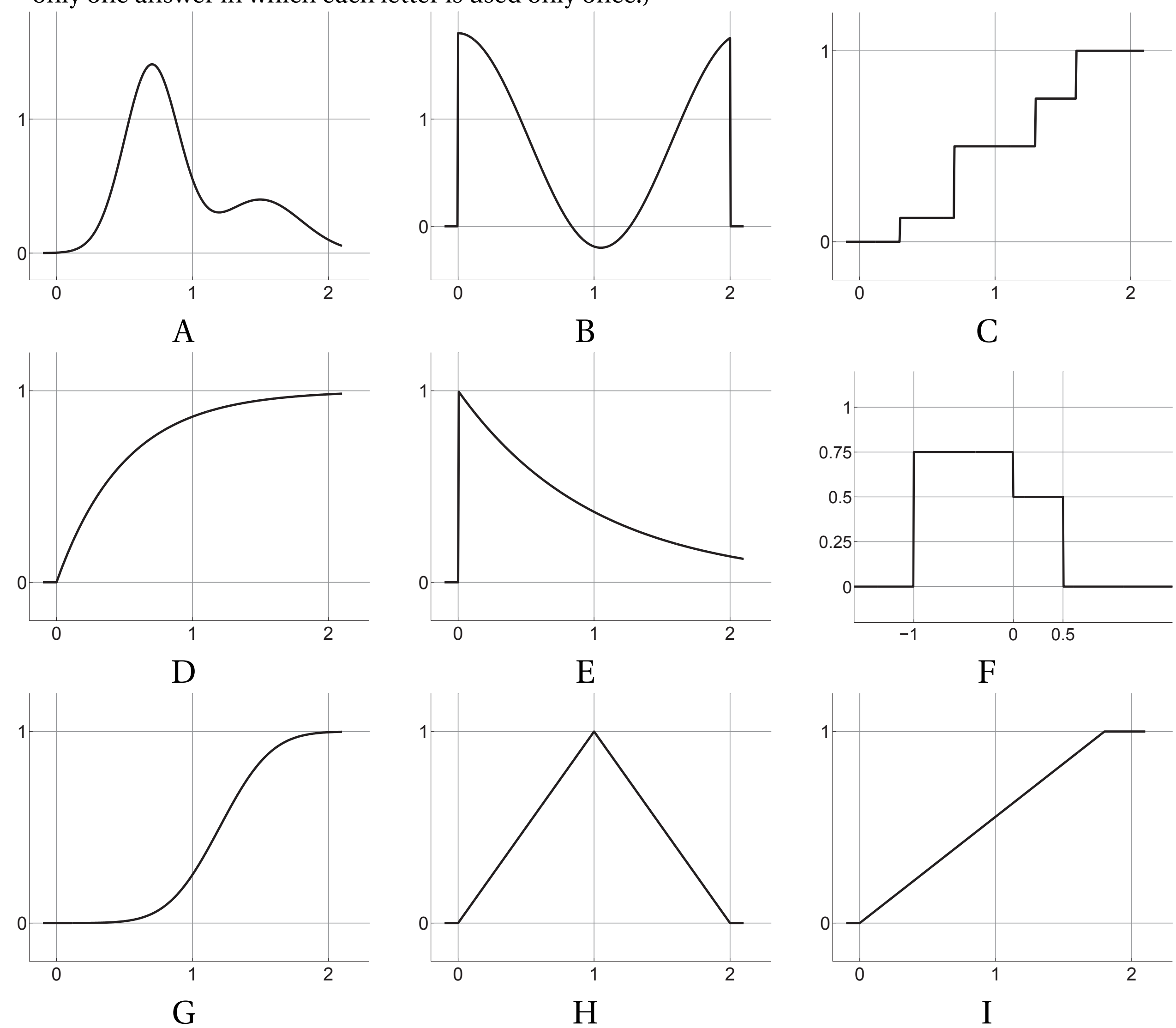
Rensselaer Polytechnic Institute
Department of Electrical, Computer, and Systems Engineering
ECSE-2500: Engineering Probability, Spring 2023
Exam 2 Solutions

1. (10 points.) Match each graph to the corresponding description by writing its letter in the box. Use each letter only once. No work is required! (Note: some graphs may fit more than one description but there's only one answer in which each letter is used only once.)



$\boxed{\text{I}}$ is the CDF of a uniform random variable. This one ranges from 0 to 1.8. The key concept is that the CDF is a straight line.

$\boxed{\text{D}}$ is the CDF of an exponential random variable. It grows quickly from 0 (since most of the probability is concentrated at small numbers) and then tapers off, reaching 1 in the limit.

$\boxed{\text{G}}$ is the CDF of a Gaussian random variable. The CDF grows slowly at either end and rapidly in the middle (around the mean).

$\boxed{\text{C}}$ is the CDF of a discrete random variable. It has the characteristic stair-step pattern that comes from integrating a sum of delta functions.

$\boxed{\text{E}}$ is the PDF of inter-arrival times generated by a Poisson random variable. These inter-arrival times have an exponential distribution, and this is an exponential PDF.

$\boxed{\text{F}}$ is the PDF for a random variable with $E(X) < 0$. This is the only one that has non-zero probability for negative values of x. We can compute from the graph that

$$E(X) = \int_{-1}^{0} \frac{3}{4}\, dx + \int_{0}^{\frac{1}{2}} \frac{1}{2}\, dx = -\frac{3}{4} + \frac{1}{4} = -\frac{1}{2}$$

$\boxed{\text{H}}$ is a valid PDF for some random variable. It satisfies the requirements: the function is nonnegative and it integrates to 1.

$\boxed{\text{B}}$ can't be a valid PDF for any random variable. It has negative "probabilities"!

$\boxed{\text{A}}$ could be a valid PDF for a random variable, but more information is needed. We can't really tell without knowing what the area under the curve is.

2. (15 points.) The number of hours per night a student spends playing Vampire Survivors has a PDF of the following form:

$$f_X(x) = \begin{cases} 2x - ax^3 & x \in [0,2] \\ 0 & \text{otherwise} \end{cases}$$

(a) (3 points.) Determine the value of a that makes $f_X(x)$ a valid PDF. Double check your answer since the rest of the problem depends on it!

We must have

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f_X(x)\, dx \\ &= \int_0^2 2x - ax^3\, dx \\ &= x^2 - \frac{a}{4}x^4 \Big]_{x=0}^{x=2} \\ &= 4 - 4a \end{aligned}$$

Thus $a = \frac{3}{4}$.

(b) (3 points.) Determine the CDF $F_X(x)$. Be sure to specify the value of the CDF at all values of x.

Assuming that $x \in [0,2]$,

$$\begin{aligned} F_X(x) &= P(X \le x) \\ &= \int_0^x f_X(u)\, du \\ &= \int_0^x 2u - \frac{3}{4}u^3\, du \\ &= u^2 - \frac{3}{16}u^4 \Big]_{u=0}^{u=x} \\ &= x^2 - \frac{3}{16}x^4 \end{aligned}$$

Thus

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x^2 - \frac{3}{16}x^4 & x \in [0,2] \\ 1 & x > 2 \end{cases}$$

(c) (3 points.) Compute $P(X < 1)$.

This is easily computed as $F_X(1) = 1 - \frac{3}{16} = \frac{13}{16}$.

(d) (3 points.) Compute $E(X)$.

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f_X(x)\ dx \\ &= \int_0^2 2x^2 - \frac{3}{4}x^4\ dx \\ &= \frac{2}{3}x^3 - \frac{3}{20}x^5 \Big]_{x=0}^{x=2} \\ &= \frac{16}{3} - \frac{24}{5} \\ &= \frac{8}{15} \end{aligned}$$

(e) (3 points.) Compute $\mathrm{Var}(X)$ (a decimal value is OK).

First we compute

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f_X(x)\ dx \\ &= \int_0^2 2x^3 - \frac{3}{4}x^5\ dx \\ &= \frac{1}{2}x^4 - \frac{3}{24}x^6 \Big]_{x=0}^{x=2} \\ &= 8 - \frac{24}{3} \\ &= 0 \end{aligned}$$

Thus

$$\mathrm{Var}(X) = E(X^2) - E(X)^2$$

There is an error in defining pdf so it will have negative value at part of the interval. Every student will get 3 points in this question (e) since the variance becomes negative.

3. (15 points.) Let X be the range in astronomical units (AU) of the Rocinante, a Corvette-class ship with a full tank of fuel. X is modeled as a Gaussian random variable with mean 278 AU and standard deviation 36 AU.

   (a) (5 points.) Compute the probability that $X > 332$ AU.

   The mean $\mu_X = 278$ and the standard deviation $\sigma_X = 36$. Thus the desired probability is:

$$\begin{aligned} P(X > 332) &= Q\left(\frac{332 - 278}{36}\right) \\ &= Q(1.5) \\ &= 0.0668 \end{aligned}$$

   (b) (5 points.) Naomi finds a new engine that has range Y on a full tank of fuel, where Y is modeled as a Gaussian random variable with mean 350 AU and standard deviation 9 AU. Compute $Y > 332$ AU.

   The mean $\mu_Y = 350$ and the standard deviation $\sigma_Y = 9$. Thus the desired probability is:

$$\begin{aligned} P(X > 332) &= Q\left(\frac{332 - 350}{9}\right) \\ &= Q(-2) \\ &= 1 - Q(2) \\ &= 0.9772 \end{aligned}$$

   (c) (5 points.) Suppose we think about the upgrade from the first engine to the second engine as a linear transformation between random variables, $Y = aX + b$, where a is a positive number. Determine the constants a and b.

   By the linearity of the expected value we have

$$\begin{aligned} \mu_Y &= a\mu_X + b \\ \sigma_Y &= a\sigma_X \end{aligned}$$

   The second equation gives us that $a = \frac{1}{4}$, and plugging this into the first equation gives $b = 280.5$.

4. (15 points.) Let X be a continuous uniform random variable on the interval $[-1,1]$. Let Y be another random variable computed as $Y = 3X^2$.

First, we note that the PDF of X is $\begin{cases} \frac{1}{2} & \text{if } x \in [-1,1] \\ 0 & \text{otherwise} \end{cases}$.

(a) (7 points.) Determine the CDF of Y, $F_Y(y)$. Be sure to specify the values of the CDF for all values of y.

Since X is between -1 and 1, $Y = 3X^2$ must be between 0 and 3. Proceeding from first principles, the CDF of Y is

$$\begin{aligned}
F_Y(y) &= P(Y \le y) \\
&= P(3X^2 \le y) \\
&= P\left(X \in [-\sqrt{y/3}, \sqrt{y/3}]\right) \\
&= \begin{cases} \frac{2\sqrt{y/3}}{2} & \text{if } y/3 \in [0,1] \\ 0 & y/3 < 0 \\ 1 & y/3 > 1 \end{cases} \\
&= \begin{cases} \sqrt{y/3} & \text{if } y \in [0,3] \\ 0 & y < 0 \\ 1 & y > 3 \end{cases}
\end{aligned}$$

(b) (8 points.) Determine the PDF of Y, $f_Y(y)$. Be sure to specify the values of the PDF for all values of y.

The PDF is the derivative of the CDF:

$$\begin{aligned}
f_Y(y) &= \frac{d}{dy} F_Y(y) \\
&= \begin{cases} \frac{1}{6\sqrt{y/3}} & \text{if } y \in [0,3] \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

5. (20 points.) Let X be a random variable representing the length of an episode of Andor, in minutes. Initially we don't know anything about the PDF of X, other than that it is symmetric about its mean.

   (a) (5 points.) The Markov bound says that $P(X \geq 55) \leq \frac{50}{55} = 0.91$.

   (b) (5 points.) The Chebyshev bound says that $P(|X-50| \geq 5) \leq \left(\frac{4}{5}\right)^2 = 0.64$, which, assuming the distribution is symmetric, would bound $P(X \geq 55) \leq 0.32$ (a lot less than the previous estimate).

   (c) (5 points.) If the random variable is in fact Gaussian, we can compute from the Q table that $P(X \geq 55) = Q(\frac{5}{4}) = Q(1.25) = 0.106$. So while both bounds are satisfied, neither is very tight. (The bounds work better when we ask for values really far from the mean.)

6. (15 points.) A variety of people come to office hours with questions. The time required to deal with a first-year advisee is an exponential random variable with mean 4 minutes. The time required to deal with a probability student is a continuous uniform random variable on [1, 11] minutes. The time required to deal with a grad student is always exactly 9 minutes. The probabilities that a visitor is a first-year, probability student, or grad student are 0.2, 0.7, and 0.1 respectively. Let X be the service time of a random visitor.

   (a) (5 points.) Compute $E(X)$.

   Recall that

$$\begin{aligned} E(X) &= E(X \mid \text{first}-\text{year})P(\text{first}-\text{year}) + E(X \mid \text{prob})P(\text{prob}) + E(X \mid \text{grad})P(\text{grad}) \\ &= 4(0.2) + 6(0.7) + 9(0.1) \\ &= 5.9 \end{aligned}$$

   A common error here was to use $\frac{1}{4}$ for the mean of the exponential random variable instead of the given mean of 4. Also, some people computed the mean of the uniform distribution as 5, not 6.

   (b) (5 points.) Use the Markov inequality to bound $P(X > 8)$. Is this an upper or lower bound?

   Markov gives us an upper bound on the tail probability, which in this case is

$$P(X > 8) \leq \frac{E(X)}{8} = \frac{5.8}{8} = 0.7375$$

   (c) (5 points.) Compute the exact value of $P(X > 8)$. Is the Markov bound tight?

$$\begin{aligned} P(X > 8) &= P(X > 8 \mid \text{first}-\text{year})P(\text{first}-\text{year}) + P(X > 8 \mid \text{prob})P(\text{prob}) + P(X > 8 \mid \text{grad})P(\text{grad}) \\ &= (e^{-\frac{1}{4}8})(0.2) + \frac{3}{10}(0.7) + 1(0.1) \\ &= 0.027 + 0.21 + 0.1 \\ &= 0.337 \end{aligned}$$

   In this case (and in general) the Markov bound isn't tight. That is, there's less probability in the tail than was predicted by Markov.

7. (15 points.) Vi and Jinx attempt to destroy the Nexuses in each others' bases. Let V represent the damage Vi can do in one attempt, and J represent the the damage Jinx can do in one attempt. We model V as a Gaussian with mean 404 and variance 4, and J as a Gaussian with mean 400 and variance 16. (Note these are variances, not standard deviations.)

   (a) (2 points.) Who is more likely to do at least 407 damage?

   We can compute that

$$P(\text{damage } > 407 \mid \text{Vi}) = Q\left(\frac{407-404}{2}\right) = Q(1.5) \tag{1}$$

$$P(\text{damage } > 407 \mid \text{Jinx}) = Q\left(\frac{407-400}{4}\right) = Q(1.75) \tag{2}$$

   It is obvious that $Q(1.5) > Q(1.75)$. So Vi is more likely to do more than 407 damage.

   (b) (3 points.) Who is more likely to do at least 416 damage?

$$P(\text{damage } > 416 \mid \text{Vi}) = Q\left(\frac{416-404}{2}\right) = Q(6) \tag{3}$$

$$P(\text{damage } > 416 \mid \text{Jinx}) = Q\left(\frac{416-400}{4}\right) = Q(4) \tag{4}$$

   It is obvious that $Q(4) > Q(6)$. So Jinx is more likely to do more than 416 damage. (Note that we don't actually have to compute the Q values in either problem, just compare the number of standard deviations away from the mean.)

   (c) (5 points.) Compute the value D for which P (V < D) = P (J < D). That is, if d is below D, Vi is more likely than Jinx to do more than d damage (since Jinx has a lower mean). If d is above D, Jinx is more likely than Vi to do more than d damage (since Vi has a lower variance).

   Basically, we're trying to find D that is the same number of standard deviations away from the respective means, so that the areas in the two tails of the distributions are the same. This results in the simple equation

$$\frac{D-404}{2} = \frac{D-400}{4}$$

   which means that $D = 408$, i.e., that 408 is 2 standard deviations away from each of the means.

   (d) (5 points.) Compute the probability that Jinx did more than 418 damage, given that she has done at least 408 damage.

   This is a conditional probability problem.

$$
\begin{aligned}
P(J > 418 \mid J > 408) &= \frac{P(J > 418 \cap J > 408)}{P(J > 408)} \\
&= \frac{P(J > 418)}{P(J > 408)} \\
&= \frac{P(Z > 4.5)}{P(Z > 2)} \\
&= \frac{Q(4.5)}{Q(2)} \\
&= \frac{3.4e-6}{2.275e-2} \\
&= 1.495e-4
\end{aligned}
$$