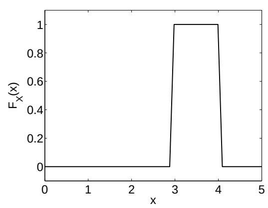
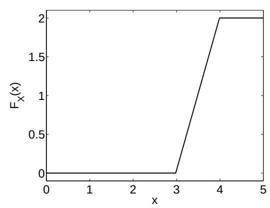
6.041/6.431 Spring 2010 Quiz 2 Solutions

Question 1

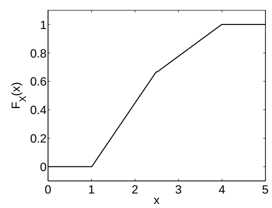
1.a. (2 pts) Which **one** of the following plots could correspond to a valid cumulative distribution function?



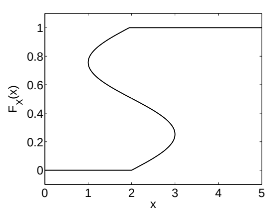
(i)



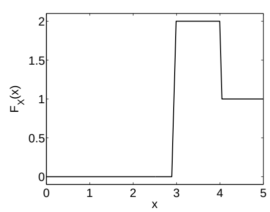
(ii)



(iii)



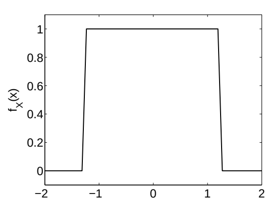
(iv)



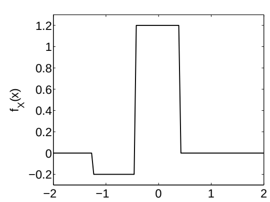
(v)

Answer: (iii). F_X **fails to be valid because it is not (i) nondecreasing; (ii) bounded above by 1; (iv) a function; (v) bounded above by 1 or nondecreasing.**

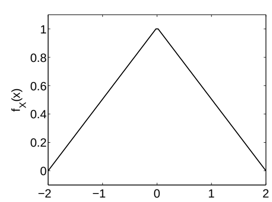
1.b. (2 pts) Which **one** of the following plots could correspond to a valid probability density function?



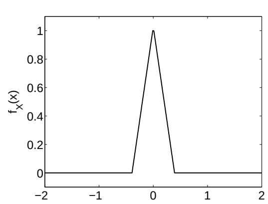
(i)



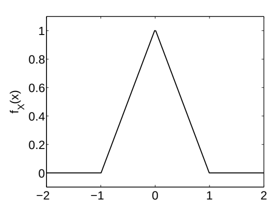
(ii)



(iii)



(iv)

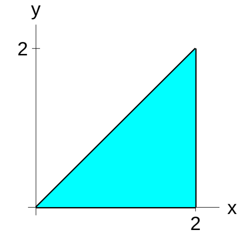


(v)

Answer: (v). The others are not nonnegative functions that integrate to 1.

Question 2

Let X and Y be jointly continuous random variables with joint probability density function



$$f_{X,Y}(x,y) = \begin{cases} \frac{3}{8}x, & \text{for } (x,y) \text{ in the shaded triangle shown to the right;} \\ 0, & \text{otherwise.} \end{cases}$$

2.a. (6 pts) Find $f_X(x)$. Give both an expression and a clearly-labeled sketch.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \begin{cases} \int_0^x \frac{3}{8}x\,dy, & \textbf{for } x \in [0,2]; \\ 0, & \textbf{otherwise} \end{cases} = \begin{cases} \frac{3}{8}x^2, & \textbf{for } x \in [0,2]; \\ 0, & \textbf{otherwise}. \end{cases}$$

(A student solution must also have a sketch.)

2.b. (6 pts) Find $\mathbf{E}[X]$.

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} x\,f_X(x)\,dx = \int_0^2 x \cdot \frac{3}{8}x^2\,dx = \frac{3}{32}x^4 \Big|_{x=0}^{x=2} = \frac{3}{2}$$

2.c. (6 pts) Find var(X).

$$\mathbf{E}[X^2] = \int_{-\infty}^{\infty} x^2\,f_X(x)\,dx = \int_0^2 x^2 \cdot \frac{3}{8}x^2\,dx = \frac{3}{40}x^5 \Big|_{x=0}^{x=2} = \frac{12}{5}$$

$$\mathbf{var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{12}{5} - \frac{9}{4} = \frac{3}{20}$$

2.d. (6 pts) Find $f_{Y|X}(y\,|\,x)$. Make sure that your answer covers all (x,y) pairs.

Since $f_{X,Y}(x,y)$ depends on y only through the shape of the shaded region, the desired conditional PDF is uniform on the appropriate interval. Specifically,

$$f_{Y|X}(y\,|\,x) = \begin{cases} 1/x, & \textbf{for } y \in [0,x]; \\ 0, & \textbf{otherwise} \end{cases}$$

for the $x \in (0,2)$ values that are valid for the conditioning. For $x \notin (0,2)$, the conditional PDF $f_{Y|X}(y\,|\,x)$ is undefined.

2.e. (6 pts) Find $\mathbf{E}[Y]$.

One could compute this directly, using $f_Y(y)$. The simple (uniform on $[0,X]$) conditional distribution of Y given X along with the computations of previous parts makes it more attractive to use the law of iterated expectations:

$$\mathbf{E}[Y] = \mathbf{E}[\,\mathbf{E}[Y\,|\,X]\,] = \mathbf{E}[X/2] = \mathbf{E}[X]/2 = 3/4.$$

2.f. (6 pts) Find var(Y).

One could compute this directly, using $f_Y(y)$. The simple (uniform on $[0, X]$) conditional distribution of Y given X along with the computations of the previous parts makes it more attractive to use the law of total variance:

$$\mathbf{var}(Y) \;=\; \mathbf{E}[\,\mathbf{var}(Y \mid X)\,] + \mathbf{var}(\,\mathbf{E}[Y \mid X]\,) \;=\; \mathbf{E}\left[\frac{1}{12}X^2\right] + \mathbf{var}(X/2) \;=\; \frac{1}{12}\cdot\frac{12}{5} + \frac{1}{4}\cdot\frac{3}{20} \;=\; \frac{19}{80}.$$

2.g. (6 pts) Find $\mathbf{P}(\mathbf{E}[Y \mid X] > 1/2)$.

Remember that $\mathbf{E}[Y \mid X]$ is a random variable. In this case, $\mathbf{E}[Y \mid X] = X/2$. The computation is straightforward:

$$\begin{aligned}
\mathbf{P}(\mathbf{E}[Y \mid X] > 1/2) &= \mathbf{P}(X/2 > 1/2) \;=\; \mathbf{P}(X > 1) \\
&= \int_1^\infty f_X(x)\,dx \;=\; \int_1^2 \frac{3}{8}x^2\,dx \;=\; \frac{3}{24}x^3\Big|_{x=1}^{x=2} \;=\; \frac{3}{24}(8-1) \;=\; \frac{7}{8}.
\end{aligned}$$

Question 3

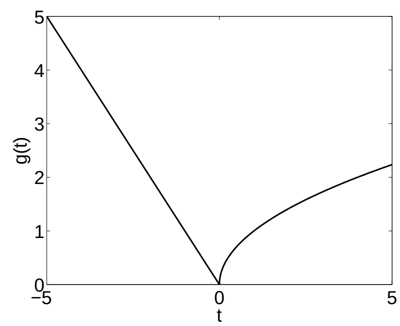
Let X have the normal distribution with mean 0 and variance 1, i.e.,

$$f_X(x) \;=\; \frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

Also, let $Y = g(X)$ where

$$g(t) \;=\; \begin{cases} -t, & \text{for } t \le 0; \\ \sqrt{t}, & \text{for } t > 0, \end{cases}$$

as shown to the right.



Find the probability density function of Y.

Because of the definition of g, the random variable Y takes on only nonnegative values. Thus $f_Y(y) = 0$ for any negative y. For $y > 0$,

$$\begin{aligned}
F_Y(y) &= \mathbf{P}(Y \le y) \\
&= \mathbf{P}\left(X \in [-y, 0]\right) + \mathbf{P}\left(X \in (0, y^2]\right) \\
&= \left(F_X(0) - F_X(-y)\right) + \left(F_X(y^2) - F_X(0)\right) \\
&= F_X(y^2) - F_X(-y).
\end{aligned}$$

Taking the derivative of $F_Y(y)$ (and using the chain rule),

$$\begin{aligned}
f_Y(y) &= 2y f_X(y^2) + f_X(-y) \\
&= \frac{1}{\sqrt{2\pi}}\left(2ye^{-y^4/2} + e^{-y^2/2}\right).
\end{aligned}$$

Question 4

Rajeev has an unlimited number of widgets to sell. Each day (numbered 1, 2, 3, ...) he receives an offer for one widget, and he must decide immediately whether to accept it. (Once he turns down an offer, it is withdrawn permanently.) Denote the offers by Z_1, Z_2, Z_3, ..., and assume that these offers are independent random variables following the exponential distribution with parameter 2.

Rajeev will use **two different strategies** to decide whether to accept offers: one in parts (a)–(d) and another in parts (e)–(h). **In parts (a)–(d), assume that Rajeev accepts an offer if and only if it is greater than 2.**

4.a. (6 pts) Find the probability density function of the purchase price of the first sold widget. **Also provide a clearly-labeled sketch of this PDF.**

Let W be the sale price of a widget. Since offers are only accepted when they exceed 2, we always have $W > 2$. By the memorylessness property of the exponential distribution, the excess above 2 follows an exponential distribution with parameter 2. Specifically, $W = V + 2$ where V is an exponential random variable with parameter 2. Thus,

$$f_W(w) \;=\; \begin{cases} 2e^{-2(w-2)}, & \textbf{for } w \geq 2; \\ 0, & \textbf{otherwise}. \end{cases}$$

(A student solution must also have a sketch.)

4.b. (6 pts) Define a random process taking values 0 and 1 based on whether Rajeev accepts the offer on a particular day. Specifically,

$$X_i \;=\; \begin{cases} 1, & \text{if Rajeev makes a sale at price } Z_i \text{ on day } i; \\ 0, & \text{otherwise.} \end{cases}$$

Is X_1, X_2, X_3, ... a Bernoulli process? (For full credit, you should clearly demonstrate your knowledge of the defining characteristics of a Bernoulli process.)

Yes, X_1, X_2, X_3, ... is a Bernoulli process. (i) Each X_i is a Bernoulli random variable. (ii) The offers are independent on different days, and Rajeev applies the same fixed rule every day in deciding whether to accept the offer, so the collection of random variables is independent. (iii) Since the offers have the same distribution every day, the probability of the offer exceeding 2 is the same every day.

4.c. (6 pts) What is the probability that Rajeev makes his third sale on Day 10?

The Bernoulli process has probability of success

$$p \;=\; \mathbf{P}(Z_i > 2) \;=\; \int_2^\infty 2e^{-2z}\,dz \;=\; -e^{-2z}\Big|_{z=2}^{z=\infty} \;=\; e^{-4}.$$

The PMF of the day on which the third sale is made is thus the Pascal PMF of order 3

$$\mathbf{P}(\textbf{third sale on day } i) \;=\; \binom{i-1}{2} p^3 (1-p)^{i-3} \;=\; \binom{i-1}{2} e^{-12}(1-e^{-4})^{i-3} \qquad \textbf{for } i \geq 3,$$

and

$$\mathbf{P}(\textbf{third sale on day 10}) \;=\; \binom{9}{2} e^{-12}(1-e^{-4})^7 \;=\; 36e^{-12}(1-e^{-4})^7.$$

4.d. (6 pts) Suppose Rajeev makes a sale on Day 1 (at price Z_1). If the offer the next day Z_2 is smaller, he is especially happy; if Z_2 is larger, he experiences *seller's regret* because he could have made more money on the first widget. Define $S = Z_2 - Z_1$ to quantify his seller's regret (which may be positive or negative). Find the probability density function of S, conditioned on there being a sale on Day 1.

S is the sum of two random variables that are conditionally independent given the event $\{Z_1 > 2\}$: $-Z_1$ and Z_2. Thus, $f_{S|\{Z_1>2\}}$ can be found by convolving

$$f_{-Z_1|\{Z_1>2\}}(z) = \begin{cases} 2e^{2(z+2)}, & \textbf{for } z \leq -2; \\ 0, & \textbf{otherwise}. \end{cases}$$

and

$$f_{Z_2|\{Z_1>2\}}(z) = \begin{cases} 2e^{-2z}, & \textbf{for } z \geq 0; \\ 0, & \textbf{otherwise}. \end{cases}$$

$$\begin{aligned}
f_{S|\{Z_1>2\}}(s) &= \int_{-\infty}^{\infty} f_{-Z_1|\{Z_1>2\}}(z)\, f_{Z_2|\{Z_1>2\}}(s-z)\, dz \\
&= \int_{-\infty}^{-2} 2e^{2(z+2)}\, f_{Z_2|\{Z_1>2\}}(s-z)\, dz \\
&= \int_{-\infty}^{\min\{-2,s\}} 2e^{2(z+2)}\, 2e^{-2(s-z)}\, dz \\
&= 4e^{-2(s-2)} \int_{-\infty}^{\min\{-2,s\}} e^{4z}\, dz \\
&= 4e^{-2(s-2)}\, \frac{1}{4} e^{4z} \Big|_{z=-\infty}^{z=\min\{-2,s\}} \\
&= 4e^{-2(s-2)} \cdot \frac{1}{4} e^{4\min\{-2,s\}} \\
&= \begin{cases} e^{-2(s-2)} e^{4s}, & \textbf{for } s < -2; \\ e^{-2(s-2)} e^{-8}, & \textbf{otherwise} \end{cases} \\
&= \begin{cases} e^{2(s+2)}, & \textbf{for } s < -2; \\ e^{-2(s+2)}, & \textbf{otherwise} \end{cases} \\
&= e^{-2|s+2|}
\end{aligned}$$

An offer is called a *record high* when it is higher than all previous offers. Denote the sequence of record-high offers $R_1, R_2, R_3, \ldots$. The first offer is always a record high, so $R_1 = Z_1$. The first offer among $Z_2, Z_3, \ldots$ that exceeds Z_1 is the second record high R_2. Similarly, R_3 is defined as the first offer that exceeds R_2, etc. **For the remaining parts, assume that Rajeev accepts an offer if and only if it is a record high.**

4.e. (6 pts) Find $\mathbf{E}[R_2]$, the expected value of the price of the second sale.

R_2 is the first offer that exceeds $R_1 = Z_1$. Because of the dependence on R_1, it makes sense to evaluate $\mathbf{E}[R_2]$ via $\mathbf{E}[\,\mathbf{E}[R_2 \,|\, R_1]\,]$. In analogy to 4.a, because of the memoryless property of the exponential distribution, $\mathbf{E}[R_2 \,|\, R_1 = r_1] = r_1 + 1/2$. Thus $\mathbf{E}[R_2 \,|\, R_1] = R_1 + 1/2$, and $\mathbf{E}[\,\mathbf{E}[R_2 \,|\, R_1]\,] = \mathbf{E}[R_1 + 1/2] = 1/2 + 1/2 = 1$.

4.f. (6 pts) Generalize the previous part by finding $\mathbf{E}[R_n]$ for any positive integer n.

R_n is the first offer that exceeds R_{n-1}. Using the law of iterated expectations and the memoryless property of the exponential distribution,

$$\mathbf{E}[R_n] \;=\; \mathbf{E}[\,\mathbf{E}[R_n \,|\, R_{n-1}]\,] \;=\; \mathbf{E}[R_{n-1} + 1/2] \;=\; \mathbf{E}[R_{n-1}] + 1/2.$$

This is the inductive step to go with the base case in 5.e to show that $\mathbf{E}[R_n] = n/2$.

4.g. (6 pts) Define a random process taking values 0 and 1 based on whether Rajeev accepts the offer on a particular day. Specifically,

$$Y_i \;=\; \begin{cases} 1, & \text{if the offer on day } i \text{ is a record high;} \\ 0, & \text{otherwise.} \end{cases}$$

Explain why $Y_1, Y_2, Y_3, \ldots$ is or is not a Bernoulli process.

This is not a Bernoulli process because $\mathbf{P}(Y_i = 1)$ is not the same for every i. Specifically, $\mathbf{P}(\{Y_1 = 1\}) = 1$ because the first offer is always a record high. Then, $\mathbf{P}(\{Y_2 = 1\}) = 1/2$ because $\{Z_1 > Z_2\}$ and $\{Z_2 > Z_1\}$ are equally likely. Similarly, there is an arrival at time 3 when Z_3 is the largest of $\{Z_1, Z_2, Z_3\}$, so $\mathbf{P}(\{Y_3 = 1\}) = 1/3$. In general, $\mathbf{P}(\{Y_n = 1\}) = 1/n$.

4.h. (6 pts) Find the probability mass function of the second arrival time in the process $Y_1, Y_2, Y_3, \ldots$.

The first arrival is always at time 1. The second arrival is at time k when Z_k is the largest of $\{Z_1, Z_2, \ldots, Z_k\}$ and Z_1 is the second largest. (If Z_1 is not second largest, then the second arrival is earlier than Day k.) Since the Z_is are identically distributed, all $k!$ rankings of them are equally likely. The rankings that result in $T_2 = k$ are $(k-2)!$ of the permutations. Thus,

$$p_{T_2}(k) \;=\; \frac{(k-2)!}{k!} \;=\; \frac{1}{k(k-1)}, \qquad \textbf{for } k = 2, 3, \ldots$$

and $p_{T_2}(k) = 0$ otherwise.