MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2009)

Problem Set 5: Solutions
Due October 19, 2009

1. In order to solve this problem, it is useful to begin by deriving the marginal densities $f_X(x)$ and $f_Y(y)$.

   The marginal densities $f_X(x)$ and $f_Y(y)$ are obtained by integrating the joint density over all y and over all x respectively:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy \tag{1}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx \tag{2}$$

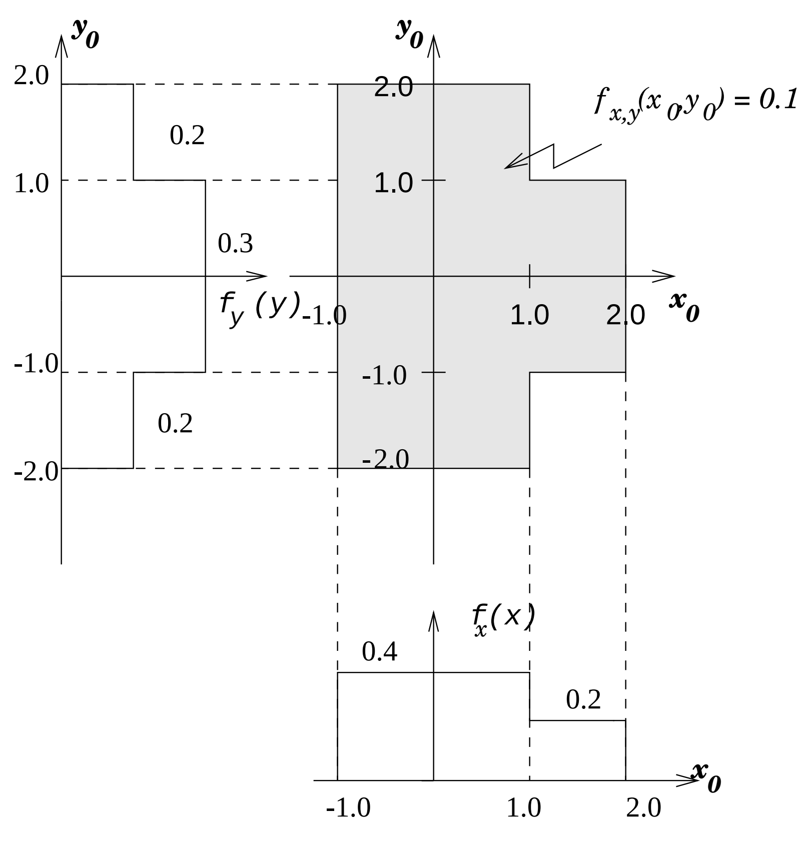


Figure 1: Marginal densities $f_X(x)$ and $f_Y(y)$.

(a) The conditional PDF $f_{Y|X}(y|x)$ is given by

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

A good way to visualize the conditional PDF $f_{Y|X}(y|x)$ is to imagine a vertical slice of the joint PDF at $X = x$. Essentially, the conditional PDF has the same shape as the joint PDF except for a scaling factor $f_X(x)$ which ensures that,

$$\int f_{Y|X}(y|x)dy = 1$$

Similarly the conditional PDF $f_{X|Y}(x|y)$ is obtained using,

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

To visualize $f_{X|Y}(x|y)$, imagine a horizontal slice through the joint PDF at $Y = y$. Again, the conditional PDF has the same shape except for the scaling factor of $f_Y(y)$. The conditional PDFs are as shown in the figure below.

(b) X and Y are **NOT** independent since $f_{X,Y}(x,y) \neq f_X(x)f_Y(y)$. Also, from the figures we have $f_{X|Y}(x|y) \neq f_X(x)$ and $f_{Y|X}(y|x) \neq f_Y(y)$.

(c)

$$\begin{aligned} f_{X,Y|A}(x,y) &= \begin{cases} \frac{f_{X,Y}(x,y)}{\mathbf{P}(A)} & (x,y) \in A \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} \frac{0.1}{\pi 0.1} & (x,y) \in A \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

(d) It can be seen that $f_{X|Y}(x|y)$ is a uniform density for each value of y. Noting the fact that the expected value of a random variable distributed uniformly between a and b is given by $\frac{(a+b)}{2}$, the conditional expected values can be written as

$$\mathbf{E}[X|Y = y] = \left\{ \begin{array}{ll} 0 & -2.0 \leq y \leq -1.0 \\ \frac{1}{2} & -1.0 \leq y \leq 1.0 \\ 0 & 1.0 \leq y \leq 2.0 \end{array} \right\}$$

Again, noting the fact that the variance of a random variable distributed uniformly between a and b is given by $\frac{(b-a)^2}{12}$, the conditional variance $\text{var}(X|Y = y)$ is given by

$$\text{var}(X|Y = y) = \left\{ \begin{array}{ll} \frac{4}{12} & -2.0 \leq y \leq -1.0 \\ \frac{9}{12} & -1.0 \leq y \leq 1.0 \\ \frac{4}{12} & 1.0 \leq y \leq 2.0 \end{array} \right\}$$

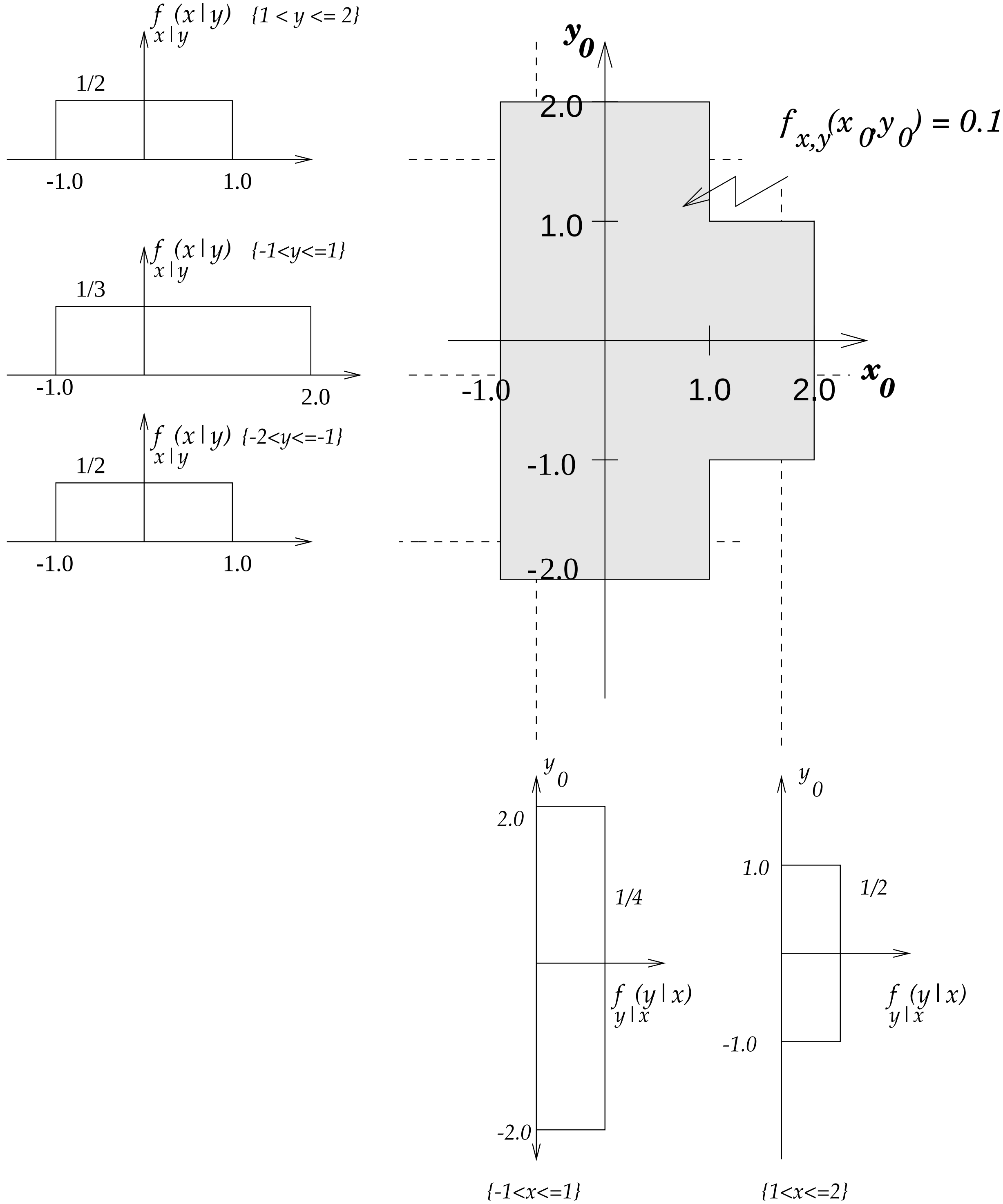


Figure 2: Conditional densities $f_{X|Y}(x|y)$ and $f_{Y|X}(y|x)$

2. (a) First, write down the general PDF of a normal random variable Z with mean μ_Z and variance σ_Z^2:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}\sigma_Z} e^{-\frac{(z-\mu_Z)^2}{2\sigma_Z^2}}$$

In this question X is a normal random variable with mean 1 and variance 4. Therefore $E[X] = 1$ and $\sigma_x = \sqrt{\text{var}(X)} = \sqrt{4} = 2$. Substitute these values into the general PDF:

$$f_X(x) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{(x-1)^2}{8}}$$

Note that Y is a linear function of X. Thus Y is also a normal random variable, and it can be completely characterized by its expectation and variance. These quantities are simple to find:

$$\begin{aligned} E[Y] &= E[3X-1] = 3 \cdot E[X] - 1 = 3 \cdot 1 - 1 = 2 \\ \text{var}(Y) &= \text{var}(3\text{X}-1) = \text{var}(3X) = 3^2 \cdot \text{var}(X) = 9 \cdot 4 = 36 \end{aligned}$$

Then we can write down the PDF of Y by substituting its expectation and variance into the generalized PDF above:

$$f_Y(y) = \frac{1}{6\sqrt{2\pi}} e^{-\frac{(y-2)^2}{72}}$$

(b) We are given $W = Y^2$. To compute the $\mathbf{E}[W]$, we use the expression for the variance of Y where $\text{var}(Y) = \mathbf{E}[Y^2] - (\mathbf{E}[Y])^2$.

$$\begin{aligned} \mathbf{E}[W] &= \mathbf{E}[Y^2] \\ &= \text{var}(Y) + (\mathbf{E}[Y])^2 \\ &= 36 + 2^2 = 40. \end{aligned}$$

3. (a) Since X and Y are independent, their joint PDF is

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \begin{cases} 1 & 0 < x, y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Define the event R_1 as $X \le 0.25$ and the event R_2 as $Y \le 0.25$. A message is *received* 15 minutes after A sent both messages whenever at least one of R_1 and R_2 occurs. The probability we wish to compute is thus

$$\mathbf{P}(R_1 \cup R_2) = \mathbf{P}(R_1) + \mathbf{P}(R_2) - \mathbf{P}(R_1 \cap R_2).$$

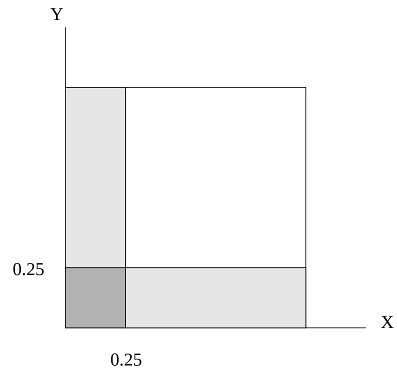
We compute the individual terms as

$$\begin{aligned} \mathbf{P}(R_1) &= \mathbf{P}\left(X \le \frac{1}{4}\right) = \frac{1}{4}, \\ \mathbf{P}(R_2) &= \mathbf{P}\left(Y \le \frac{1}{4}\right) = \frac{1}{4}, \\ \mathbf{P}(R_1 \cap R_2) &= \mathbf{P}(R_1) \cdot \mathbf{P}(R_2) = \left(\frac{1}{4}\right)^2 = \frac{1}{16}, \quad \text{since } X \text{ and } Y \text{ are independent.} \end{aligned}$$

Thus the desired probability is

$$\frac{1}{4}+\frac{1}{4}-\frac{1}{16} = \frac{7}{16}.$$

Note also that the probability is the total area of the shaded regions in the following sketch.



(b) Let B be the event that the message is received but not verified within 15 minutes. Then

$$B = R_1 \cap R_2^c \cup R_1^c \cap R_2$$

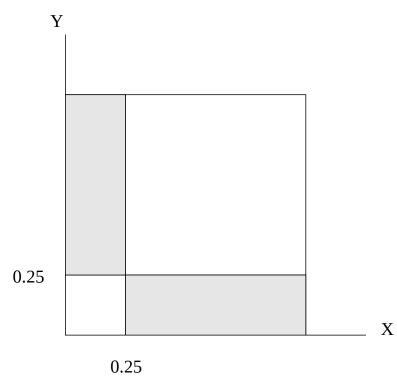
Note that this is a union of disjoint events, so we have

$$\mathbf{P}(B) = \mathbf{P}(R_1 \cap R_2^c) + \mathbf{P}(R_1^c \cap R_2),$$

and the independence of R_1 and R_2 allows the simplification

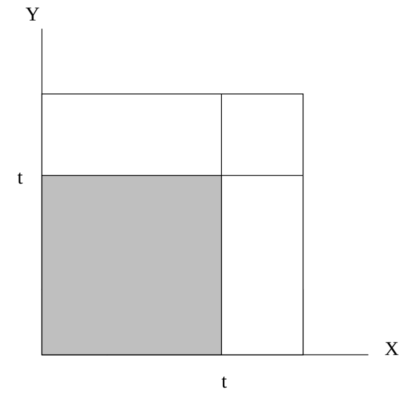
$$\begin{aligned}\mathbf{P}(B) &= \mathbf{P}(R_1)\cdot\mathbf{P}(R_2^c) + \mathbf{P}(R_1^c)\cdot\mathbf{P}(R_2) \\ &= \frac{1}{4}\cdot\frac{3}{4}+\frac{3}{4}\cdot\frac{1}{4} = \frac{3}{8}.\end{aligned}$$

Note also that the probability is the total area of the shaded regions in the following sketch.



(c) Verification occurs when the second of the messages arrives, so for $t \in [0,1]$

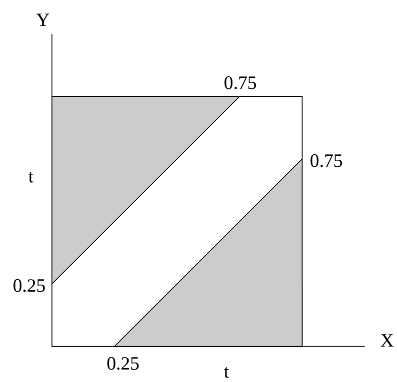
$$\begin{aligned}F_T(t) &= \mathbf{P}(T \le t) = \mathbf{P}(X \le t \cap Y \le t) \\ &= \mathbf{P}(X \le t)\cdot\mathbf{P}(Y \le t) \quad \text{by the independence of } X \text{ and } Y \\ &= t \cdot t = t^2.\end{aligned}$$



From this we can deduce the full CDF and differentiate to determine the PDF:

$$F_T(t) = \begin{cases} 0 & \text{if } -\infty < t \leq 0, \\ t^2 & \text{if } 0 < t \leq 1, \\ 1 & \text{if } 1 < t < \infty \end{cases} \qquad \Rightarrow \qquad f_T(t) = \begin{cases} 2t & \text{if } 0 \leq t \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

(d) The event that the clerk will be there to receive the message is $\{|X - Y| > \frac{1}{4}\}$. We can deduce the probability of this event easily from a sketch:



$$\mathbf{P}(|X - Y| > \frac{1}{4}) = 2 \cdot \left(\frac{1}{2} \cdot \frac{3}{4} \cdot \frac{3}{4} \right) = \frac{9}{16}.$$

(e) We know the strategy from (d) has $\frac{9}{16}$ probability of verification. The other strategy, of sending the employee home after 45 minutes, has probability of verification $\mathbf{P}(T \leq \frac{3}{4}) = \frac{9}{16}$ by evaluating the expression from part (c). Therefore, the two strategies are equally effective.

4. The purported solution is not correct. This problem illustrates the danger in writing down expressions without keeping track of the ranges on which they hold.

   We use the convenient "indicator" notation

$$1_C = \begin{cases} 1, & \text{when the condition } C \text{ holds;} \\ 0, & \text{otherwise.} \end{cases}$$

   to write

$$f_{X,Y}(x, y) = 1_{x \in [0,1]} \, 1_{y \in [x, x+1]}.$$

Correct calculations give

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_{-\infty}^{\infty} 1_{x\in[0,1]}\,1_{y\in[x,x+1]}\,dy \\ &= 1_{x\in[0,1]}\int_x^{x+1} 1\cdot\,dy = 1_{x\in[0,1]} \end{aligned}$$

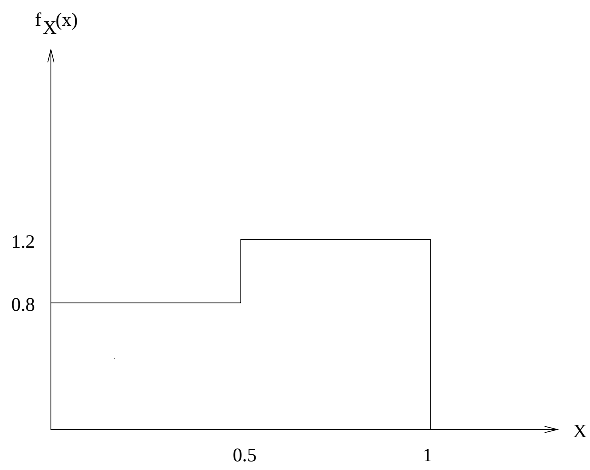
$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx \;=\; \int_0^1 1_{y\in[x,x+1]}\cdot\,dx \\ &= \begin{cases} 0, & y<0; \\ \int_0^y 1\,dx, & 0\le y<1; \\ \int_{y-1}^1 1\,dx, & 1\le y<2; \\ 0, & y\ge 2; \end{cases} \;=\; \begin{cases} 0, & y<0; \\ y, & 0\le y<1; \\ 2-y, & 1\le y<2; \\ 0, & y\ge 2; \end{cases} \end{aligned}$$

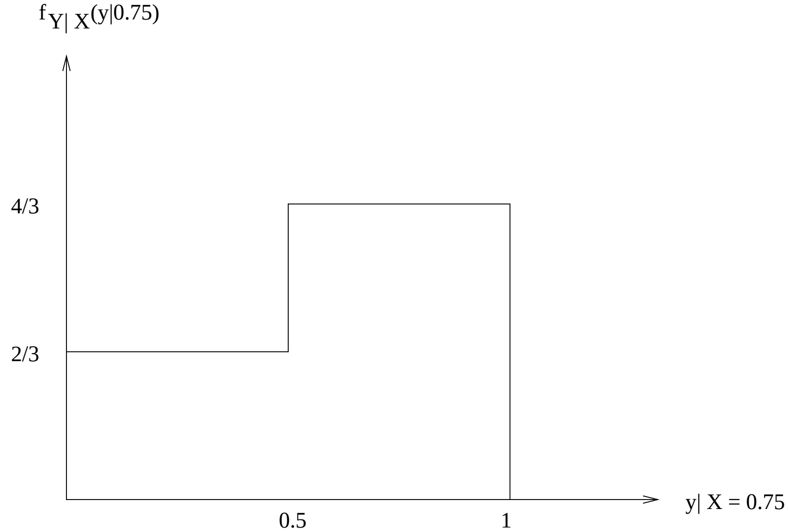
$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = 1_{y\in[x,x+1]} \quad \text{for } x\in[0,1] \text{ (and } f_{Y|X}(y|x) \text{ is undefined otherwise)}$$

X and Y are *not* independent because $f_{Y|X}(y|x)$ depends on x (the range for which the expression "1" holds is a dependence on x!). Alternatively, $f_X(x)\cdot f_Y(y) \neq f_{X,Y}(x,y)$.

5. (a) X and Y are not independent because there exist x and y such that $f_{X,Y}(x,y) \neq f_X(x)f_Y(y)$. For instance, $f_{X,Y}(\frac{2}{3},\frac{1}{3}) = 0.8$, $f_X(\frac{2}{3}) = \int_0^1 f_{X,Y}(\frac{2}{3},y)dy = 1.2$, $f_Y(\frac{1}{3}) = \int_0^1 f_{X,Y}(x,\frac{1}{3})dx = 0.8$, but $f_{X,Y}(\frac{2}{3},\frac{1}{3}) \neq f_X(\frac{1}{3})f_Y(\frac{1}{3})$.
   We can see this intuitively in the graph: For example, if X is larger than 0.5, then y is more likely to be large.

   (b) The plots are shown below.





$$f_X(x) = \begin{cases} 0.8, & 0 < x \le 1/2 \\ 1.2, & 1/2 < x \le 1 \\ 0, & \text{otherwise} \end{cases} \qquad f_{Y|X}(y \mid 0.75) = \frac{f_{X,Y}(0.75,y)}{f_X(0.75)} = \begin{cases} 2/3, & 0 < y \le 1/2 \\ 4/3, & 1/2 < y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

(c) Conditioned on event A, X and Y are independent. Thus

$$\mathbf{E}[R \mid A] = \mathbf{E}[XY \mid A] = \mathbf{E}[X \mid A]\mathbf{E}[Y \mid A] = (1/4)(1/2) = 1/8.$$

(d) It is easiest to see the CDF of W in this case as the integral of the PDF over an L-shaped area. For $0 < w \le 1/2$ the CDF would be the integral over the PDF of the L-shaped area given by $((1)(w) + (w)(1-w))(0.8)$. Similarly, for $1/2 < w \le 1$ the CDF would take on the values $(0.8)(3/4) + ((w-0.5)(0.5) + (1-w)(w-0.5))(1.6)$. Thus the entire CDF is given by

$$F_W(w) = \begin{cases} 0, & w \le 0 \\ (2w - w^2)(0.8), & 0 < w \le 1/2 \\ 1 - (1-w)^2(1.6), & 1/2 < w \le 1 \\ 1, & w > 1 \end{cases}$$

6. (a) Random variable X is a mixed random variable which is equal to an exponential with parameter $\lambda = 1$ with probability 1/2 and equal to an exponential with parameter $\lambda = 3$ also with probability 1/2. To find the expected value of X, we can use the total expectation theorem:

$$\mathbf{E}[X] = \mathbf{E}[X|A]\mathbf{P}(A) + \mathbf{E}[X|B]\mathbf{P}(B) = 1 \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} = \frac{2}{3}$$

(b) Using the total probability theorem, and noting that for an exponential random variable with parameter λ that $P(X > x) = e^{-\lambda x}$,

$$\mathbf{P}(D) = \mathbf{P}(D|A)\mathbf{P}(A) + \mathbf{P}(D|B)\mathbf{P}(B) = \frac{1}{2}e^{-\tau} + \frac{1}{2}e^{-3\tau}$$

(c) Using the Bayes' theorem, and the results from part (b):

$$\begin{aligned}\mathbf{P}(A|D) &= \frac{\mathbf{P}(D|A)\mathbf{P}(A)}{\mathbf{P}(D)} \\ &= \frac{\frac{1}{2}e^{-\tau}}{\frac{1}{2}e^{-\tau} + \frac{1}{2}e^{-3\tau}} \\ &= \frac{1}{1+e^{-2\tau}}\end{aligned}$$

(d) We need to determine $E[X|D]$. Using the total expectation theorem:

$$E[X|D] = E[X|A,D]P(A|D) + E[X|B,D]P(B|D)$$

where the probability $P(A|D)$ is the conditional probability that the first bulb used is a type-A bulb given that there are no failures during the first τ hours. From part (c) $P(A|D) = \frac{1}{1+e^{-2\tau}}$, and so $P(B|D) = 1 - P(A|D) = \frac{e^{-2\tau}}{1+e^{-2\tau}}$. To find the conditional expectations for each type of bulb given that $X > \tau$, we make use of the memoryless property of the exponential. The memoryless property of the exponential says that for an exponential random variable Y conditioned on the event $Y > \tau$, the distribution on $Y - \tau$ is the same as the distribution on Y. The following steps show how to use this property to find $E[X|D]$:

$$\begin{aligned}E[X|D] &= E[X|A,D]P(A|D) + E[X|B,D]P(B|D) \\ &= E[X-\tau+\tau|A,D)P(A|D) + E[X-\tau+\tau|B,D)P(B|D) \\ &= \tau(P(A|D)+P(B|D)) + E[X-\tau|A,D]P(A|D) \\ &\quad +E[X-\tau|B,D]P(B|D) \\ &= \tau + E[X|A]P(A|D) + E[X|B]P(B|D) \\ &= \tau + 1\cdot\frac{1}{1+e^{-2\tau}} + \frac{1}{3}\cdot\frac{e^{-2\tau}}{1+e^{-2\tau}}\end{aligned}$$

where the steps $E[X-\tau|A,D] = E[X|A]$ and $E[X-\tau|B,D] = E[X|B]$ follow from the memoryless property of the exponential.

7. (a) Let A be the event that the machine is functional. Conditioned on the random variable Q taking on a particular value q, $\mathbf{P}(A|Q=q) = q$. Using the continuous form of the total probability theorem, the probability of event A is given by:

$$\begin{aligned}\mathbf{P}(A) &= \int_0^1 \mathbf{P}(A|Q=q)f_Q(q)dq \\ &= \int_0^1 q\,dq \\ &= 1/2\end{aligned}$$

(b) Let B be the event that the machine is functional on m out of the last n days. Conditioned on random variable Q taking on value q (a probability q of being functional) the probability of event B is binomial with n trials, m successes, and a probability q of success in each trial. Again using the total probability theorem, the probability of event B is given by:

$$\begin{aligned}\mathbf{P}(B) &= \int_0^1 \mathbf{P}(B|Q=q)f_Q(q)dq \\ &= \int_0^1 \binom{n}{m} q^m(1-q)^{n-m} f_Q(q)dq \\ &= \binom{n}{m}\frac{m!(n-m)!}{(n+1)!}\end{aligned}$$

We then find the distribution on Q conditioned on event B using Bayes rule:

$$\begin{aligned}f_{Q|B}(q) &= \frac{\mathbf{P}(B|Q=q)f_Q(q)}{\mathbf{P}(B)} \\ &= \frac{q^m(1-q)^{n-m}}{\frac{m!(n-m)!}{(n+1)!}} \quad 0 \le q \le 1, \quad n \ge m\end{aligned}$$

(c) Let event C be the probability that the machine is functional today. The probability $\mathbf{P}(C|B)$ is given by:

$$\begin{aligned}\mathbf{P}(C|B) &= \int_0^1 \mathbf{P}(C|Q=q,B)f_{Q|B}(q)dq \\ &= \int_0^1 \mathbf{P}(C|Q=q)f_{Q|B}(q)dq \\ &= \int_0^1 \frac{q\cdot q^m(1-q)^{n-m}}{\frac{m!(n-m)!}{(n+1)!}}dq \quad n \ge m \\ &= \int_0^1 \frac{q^{m+1}(1-q)^{(n+1)-(m+1)}}{\frac{m!(n-m)!}{(n+1)!}}dq \quad n \ge m \\ &= \frac{\frac{(m+1)!(n-m)!}{(n+2)!}}{\frac{m!(n-m)!}{(n+1)!}} \quad n \ge m \\ &= \frac{m+1}{n+2} \quad n \ge m\end{aligned}$$

where the second equality follows since events C and B are conditionally independent given Q.

G1†. Let $p = (\cos\theta, \sin\theta)$ and $q = (a, b)$. The other two vertices of R are $(\cos\theta, b)$ and $(a, \sin\theta)$. If $|a| \le |\cos\theta|$ and $|b| \le |\sin\theta|$, then each vertex (x, y) of R satisfies $x^2+y^2 \le \cos^2\theta + \sin^2\theta = 1$ and no points of R can lie outside of C. Conversely if no points of R lies outside of C, then applying this to the two vertices other than p and q, we find

†Required for 6.431; optional challenge problem for 6.041

$$\cos^2\theta + b^2 \le 1, \quad \text{and} \quad a^2 + \sin^2\theta \le 1. \tag{3}$$

or equivalently

$$|b| \le |\sin\theta|, \quad \text{and} \quad |a| \le |\cos\theta| \tag{4}$$

These conditions imply that (a, b) lies inside or on C, so for any given θ, the probability that the random point $q = (a, b)$ satisfies (4) is

$$\frac{2|\cos\theta| \cdot 2|\sin\theta|}{\pi} = \frac{2}{\pi}|\sin(2\theta)|$$

and the overall probability is

$$\frac{1}{2\pi}\int_0^{2\pi} \frac{2}{\pi}|\sin(2\theta)|d\theta = \frac{4}{\pi^2}\int_0^{\pi/2} \sin(2\theta)d\theta = \frac{4}{\pi^2}$$

G2†. (a) Let $X_1, X_2, \ldots X_n$ be independent, identically distributed (IID) random variables. We note that

$$\mathbf{E}[X_1 + \cdots + X_n \mid X_1 + \cdots + X_n = x_0] = x_0.$$

It follows from the linearity of expectations that

$$\begin{aligned} x_0 &= \mathbf{E}[X_1 + \cdots + X_n \mid X_1 + \cdots + X_n = x_0] \\ &= \mathbf{E}[X_1 \mid X_1 + \cdots + X_n = x_0] + \cdots + \mathbf{E}[X_n \mid X_1 + \cdots + X_n = x_0] \end{aligned}$$

Because the X_i's are identically distributed, we have the following relationship.

$$\mathbf{E}[X_i \mid X_1 + \cdots + X_n = x_0] = \mathbf{E}[X_j \mid X_1 + \cdots + X_n = x_0], \text{ for any } 1 \le i \le n, 1 \le j \le n.$$

Therefore,

$$\begin{aligned} n\mathbf{E}[X_1 \mid X_1 + \cdots + X_n = x_0] &= x_0 \\ \mathbf{E}[X_1 \mid X_1 + \cdots + X_n = x_0] &= \frac{x_0}{n}. \end{aligned}$$

(b) Note that we can rewrite $\mathbf{E}[X_1 \mid S_n = s_n, S_{n+1} = s_{n+1}, \ldots, S_{2n} = s_{2n}]$ as follows:

$$\begin{aligned} &\mathbf{E}[X_1 \mid S_n = s_n, S_{n+1} = s_{n+1}, \ldots, S_{2n} = s_{2n}] \\ = \; &\mathbf{E}[X_1 \mid S_n = s_n, X_{n+1} = s_{n+1} - s_n, X_{n+2} = s_{n+2} - s_{n+1}, \ldots, X_{2n} = s_{2n} - s_{2n-1}] \\ = \; &\mathbf{E}[X_1 \mid S_n = s_n], \end{aligned}$$

where the last equality holds due to the fact that the X_i's are independent. We also note that

$$\mathbf{E}[X_1 + \cdots + X_n \mid S_n = s_n] = \mathbf{E}[S_n \mid S_n = s_n] = s_n.$$

It follows from the linearity of expectations that

$$\mathbf{E}[X_1 + \cdots + X_n \mid S_n = s_n] = \mathbf{E}[X_1 \mid S_n = s_n] + \cdots + \mathbf{E}[X_n \mid S_n = s_n].$$

Because the X_i's are identically distributed, we have the following relationship:

$$\mathbf{E}[X_i \mid S_n = s_n] = \mathbf{E}[X_j \mid S_n = s_n], \text{ for any } 1 \le i \le n, 1 \le j \le n.$$

Therefore,

$$\mathbf{E}[X_1 + \cdots + X_n \mid S_n = s_n] = n\mathbf{E}[X_1 \mid S_n = s_n] = s_n \Rightarrow \mathbf{E}[X_1 \mid S_n = s_n] = \frac{s_n}{n}.$$

†Required for 6.431; optional challenge problem for 6.041