# Massachusetts Institute of Technology

Department of Electrical Engineering & Computer Science

**6.041/6.431: Probabilistic Systems Analysis**

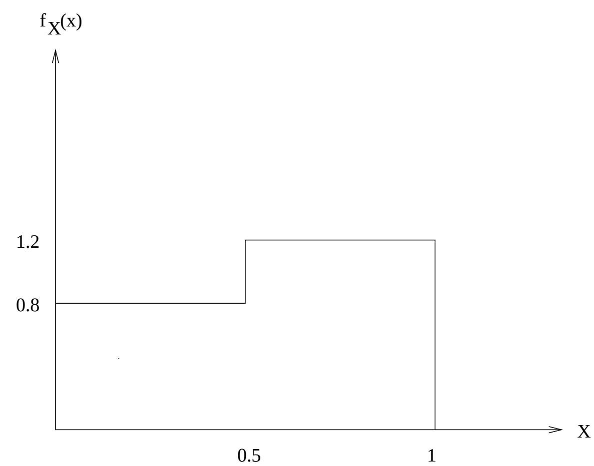
(Spring 2009)

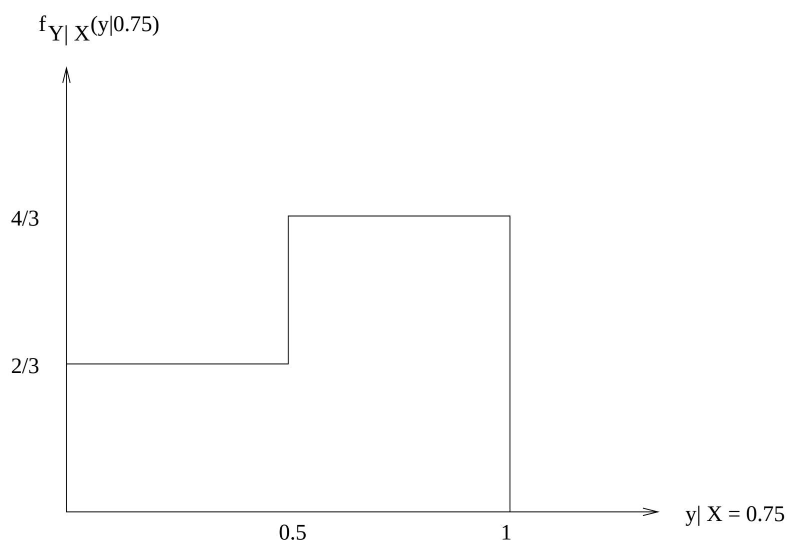
**Problem Set 5 Solutions**
**Due: March 18, 2009**

- **Text Sections: 3.1-3.5**
- **Continuous random variables and PDFs, CDFs, Normal random variables, Joint PDFs, Conditioning.**

1. (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.




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

(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 $0 < w \leq 1/2$ 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 $1/2 < w \leq 1$ 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

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

2. (a) An error will occur in the following cases:

- If the system concludes 0 or 1 was sent when -1 was actually sent.
- If the system concludes -1 or 1 was sent when 0 was actually sent.
- If the system concludes -1 or 0 was sent when 1 was actually sent.

Therefore, we can compute the probability of error, $\mathbf{P}(\varepsilon)$, by using the Total Probability Theorem:

$$\begin{aligned}
\mathbf{P}(\varepsilon) &= \mathbf{P}(\varepsilon \mid X=-1)\mathbf{P}(X=-1) + \mathbf{P}(\varepsilon \mid X=0)\mathbf{P}(X=0) + \mathbf{P}(\varepsilon \mid X=1)\mathbf{P}(X=1) \\
&= \mathbf{P}(Y > -0.5 \mid X=-1)\mathbf{P}(X=-1) + \mathbf{P}(Y < -0.5 \cup Y > 0.5 \mid X=0)\mathbf{P}(X=0) \\
&\qquad + \mathbf{P}(Y < 0.5 \mid X=1)\mathbf{P}(X=1) \\
&= \frac{1}{3}\left(\mathbf{P}(Y > -0.5 \mid X=-1) + \mathbf{P}(Y < -0.5 \mid X=0)\right. \\
&\qquad \left. + \mathbf{P}(Y > 0.5 \mid X=0) + \mathbf{P}(Y < 0.5 \mid X=1)\right) \\
&= \frac{1}{3}\left(\mathbf{P}(N > 0.5) + \mathbf{P}(N < -0.5) + \mathbf{P}(N > 0.5) + \mathbf{P}(N < -0.5)\right)
\end{aligned}$$

Since the normal distribution is symmetric, $\mathbf{P}(N > 0.5) = \mathbf{P}(N < -0.5)$. So, $\mathbf{P}(\varepsilon) = \frac{4}{3}\mathbf{P}(N < -0.5)$.
$\mathbf{P}(N < 0.5)$ can be computed from the CDF:

$$\mathbf{P}(N \geq n) = \Phi\left(\frac{n-\mu}{\sigma}\right)$$

Therefore

$$\begin{aligned}
\mathbf{P}(\varepsilon) &= \frac{4}{3}\mathbf{P}(N < -0.5) \\
&= \frac{4}{3}\Phi\left(\left(-\frac{1}{2}\right)\frac{1}{\sqrt{4}}\right) \\
&\approx \boxed{0.535}
\end{aligned}$$

(b) Similarly

$$\begin{aligned}
\mathbf{P}(\varepsilon) &= \mathbf{P}(\varepsilon \mid X=-2)\mathbf{P}(X=-2) + \mathbf{P}(\varepsilon \mid X=0)\mathbf{P}(X=0) + \mathbf{P}(\varepsilon \mid X=2)\mathbf{P}(X=2) \\
&= \mathbf{P}(Y > -1 \mid X=-2)\mathbf{P}(X=-2) + \mathbf{P}(Y < -1 \cup Y > 1 \mid X=0)\mathbf{P}(X=0) \\
&\qquad + \mathbf{P}(Y < 1 \mid X=2)\mathbf{P}(X=2) \\
&= \frac{1}{3}\left(\mathbf{P}(Y > -1 \mid X=-2) + \mathbf{P}(Y < -1 \mid X=0) + \mathbf{P}(Y > 1 \mid X=0) + \mathbf{P}(Y < 1 \mid X=2)\right) \\
&= \frac{1}{3}\left(\mathbf{P}(N > 1) + \mathbf{P}(N < -1) + \mathbf{P}(N > 1) + \mathbf{P}(N < -1)\right)
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}(\varepsilon) &= \frac{4}{3}P(N \geq 1) \\
&= \frac{4}{3}\Phi\left(-\frac{1}{\sqrt{4}}\right) \\
&\approx \boxed{0.41}
\end{aligned}$$

Therefore the probability of error $\mathbf{P}(\varepsilon)$ is smaller when transmitting signals with values $\{-2, 0, 2\}$ than when transmitting $\{-1, 0, 1\}$.

(c) Using transmission symbols $\{-m, 0, m\}$, one can make error as small as desired by choosing $m$ large enough.

3. 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

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

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

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

$X$ 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, $f_X(x) \cdot f_Y(y) \neq f_{X,Y}(x,y)$.

4. a) $P(S = S_n) = 0$ because $S, S_n$ are continuous random variables.
b) $P(S > S_n)$ is the probability that the host's roll is the biggest. By symmetry, this must be $\frac{1}{n+1}$

c) We are asked to find $P(S > S_1)$.

$$\begin{aligned} P(S > S_1) &= 1 - P(S \leq S_1) \\ &= 1 - \frac{1}{n+1} \\ &= \frac{n}{n+1}. \end{aligned}$$

d) We need to find: $P(S_j < S < S_k)$ for $1 \leq j < k \leq n$. Before the host spins the wheel, his roll has an equal a priori probability of being anywhere in the line of contestant rolls. Note that this approach yields the same answer for part (c). Thus the probability that the host will be inbetween $S_j$ and $S_k$ is $\frac{k-j}{n+1}$.

5. Let $X_1$, $X_2$ and $X_3$ denote the lifetime (in months) of bulbs A, B and C respectively. $X_1$, $X_2$ and $X_3$ are therefore independent, and exponentially distributed with parameters 1, $\frac{1}{2}$ and $\frac{1}{3}$ respectively.

(a)

$$\begin{aligned} \mathbf{P}(T \geq t) &= \mathbf{P}(X_1 \geq t, X_2 \geq t, X_3 \geq t) \\ &= \mathbf{P}(X_1 \geq t)\mathbf{P}(X_2 \geq t)\mathbf{P}(X_3 \geq t) \\ &= e^{-1 \cdot t} \cdot e^{-\frac{1}{2} \cdot t} \cdot e^{-\frac{1}{3} \cdot t} \\ &= e^{-t(1+\frac{1}{2}+\frac{1}{3})} \\ &= e^{-t \cdot \frac{11}{6}}. \end{aligned}$$

Thus $T$ is exponentially distributed with parameter $\lambda = 1 + \frac{1}{2} + \frac{1}{3} = \frac{11}{6}$.

(b)

$$\begin{aligned} \mathbf{P}(\text{First bulb to fail is bulb C}) &= \mathbf{P}(X_3 \leq X_2, X_3 \leq X_1) \\ &= \int_0^\infty \mathbf{P}(t \leq X_2, t \leq X_1 | X_3 = t) f_{X_3}(t)\, dt \\ &= \int_0^\infty \mathbf{P}(t \leq X_2)\mathbf{P}(t \leq X_1) \frac{1}{3} e^{-\frac{t}{3}}\, dt \\ &= \frac{1}{3} \int_0^\infty e^{-\frac{t}{2}} e^{-t} e^{-\frac{t}{3}}\, dt \\ &= \frac{1}{3} \frac{1}{1+\frac{1}{2}+\frac{1}{3}} \int_0^\infty (1 + \frac{1}{2} + \frac{1}{3}) e^{-\frac{t}{2}} e^{-t} e^{-\frac{t}{3}}\, dt \\ &= \frac{\frac{1}{3}}{1+\frac{1}{2}+\frac{1}{3}} = \frac{2}{11} \end{aligned}$$

# Massachusetts Institute of Technology

Department of Electrical Engineering & Computer Science

**6.041/6.431: Probabilistic Systems Analysis**

(Spring 2009)

---

G1†. (a) Since $0 \le X \le c$,

$$\begin{aligned} \mathbf{E}[X^2] &= \mathbf{E}[XX] \\ &\le \mathbf{E}[cX] \\ &= c\mathbf{E}[X]. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathrm{var}(X) &= \mathbf{E}[X^2] - (\mathbf{E}[X])^2 \\ &\le c\mathbf{E}[X] - (\mathbf{E}[X])^2 \\ &= c^2\left(\frac{\mathbf{E}[X]}{c}\right) - c^2\left(\frac{\mathbf{E}[X]}{c}\right)^2 \\ &= c^2\left(\frac{\mathbf{E}[X]}{c}\left(1 - \frac{\mathbf{E}[X]}{c}\right)\right) \\ &= c^2[\alpha(1-\alpha)], \end{aligned}$$

where $\alpha = \mathbf{E}[X]/c$. The value $\alpha$ is not something we choose, rather it is fixed by the specification of the random variable $X$ (among those satisfying $0 \le X \le c$). Thus we must take the *weakest* of the bounds obtained by variation of $\alpha$ to have a statement that is always true. The maximum of $\alpha(1-\alpha)$ for $\alpha \in [0,1]$ occurs at $\alpha = \frac{1}{2}$ and yields the bound

$$\mathrm{var}(X) \le \frac{c^2}{4}.$$

(b) Let $\tilde{X}$ be a discrete random variable that takes only two values 0 and $c$, each with probability $\frac{1}{2}$. The PMF of $\tilde{X}$ is therefore given by

$$p_{\tilde{X}}(x) = \begin{cases} 0.5 & x = 0, c \\ 0 & \text{otherwise.} \end{cases}$$

We calculate $\mathbf{E}(\tilde{X}) = c/2$ and $\mathbf{E}(\tilde{X}^2) = c^2/2$. Therefore,

$$\begin{aligned} \mathrm{var}(\tilde{X}) &= \mathbf{E}(\tilde{X}^2) - (\mathbf{E}(\tilde{X}))^2 \\ &= \frac{c^2}{2} - \frac{c^2}{4} \\ &= \frac{c^2}{4} \end{aligned}$$

---

†