Massachusetts Institute of Technology
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Spring 2011)

Problem Set 6 Solutions

1. Symmetry is a powerful argument, and most importantly, gets us the answer to (a) and (b).

   (a) 1/2

   (b) 1/3

   (c) $P(n = n_0) = P(x_1$ is the smallest of 1st $n_0 - 1$, *and* x_{n_0} is the smallest of first $n_0)$. So using the extension of the argument in (a) and (b), we have $P(n = n_0) = \frac{1}{n_0-1} \times \frac{1}{n_0}$.
   Now,
   $$P(n > n_0) = 1 - P(n \leq n_0) = 1 - \sum_{k=2}^{n_0} \frac{1}{k(k-1)}$$
   Alternatively, $P(n > n_0) = P(x_1$ is the smallest of 1st $n_0) = \frac{1}{n_0}$.

   (d) Using part (c) we can see that this expectation is ∞ since the infinite sum of $\frac{1}{n_i}$ goes to ∞. Note that here we are using the result from Problem Set 4 which states that $\sum_{i=1}^{\infty} P(n \geq n_i) = E(n)$.

   (e) We could use the symmetry between the maximum and the mininum to obtain the results. Alternatively, we could multiply the x_i's by -1 and convert it into a problem about minimums. Then we could apply the results from (c) and (d) to see that the answers do not change.

2. (a) The total area under the PDF must equal to 1. Hence if we integrate the PDF of X for all values of X and set the integration equal to 1, we can determine the value of c.
   $$\int_{-\infty}^{\infty} f_X(x)dx = \int_1^4 cx^{-2}dx = -cx^{-1}\Big|_1^4 = \frac{3c}{4} = 1 \quad \Rightarrow \quad c = \frac{4}{3}$$

   (b) $P(A)$ is simply the area of the region where $2 < X \leq 4$.
   $$P(A) = \int_2^4 \frac{4}{3}x^{-2}dx = \frac{1}{3}$$
   Using the PDF of X and $P(A)$, $f_{X|A}(x)$ can now be calculated easily.
   $$f_{X|A}(x) = \frac{f_X(x)}{P(A)} = \begin{cases} 4x^{-2}, & 2 < x \leq 4, \\ 0, & \text{otherwise.} \end{cases}$$

   (c) $E[Y \mid A] = E[X^2 \mid A]$. Since we already obtained the conditional PDF of X given A in part (b), we can compute $E[X^2 \mid A]$ as follows:
   $$E[Y \mid A] = E[X^2 \mid A] = \int_2^4 x^2 f_{X|A}(x)dx = \int_2^4 x^2 4x^{-2}dx = \int_2^4 4dx = 8$$
   To obtain the conditional variance of Y given A, we need $E[Y^2 \mid A]$.
   $$E[Y^2 \mid A] = E[X^4 \mid A] = \int_2^4 x^4 4x^{-2}dx = \int_2^4 4x^2 dx = \frac{224}{3}$$
   Therefore
   $$\text{var}(Y \mid A) = E[Y^2 \mid A] - E[Y \mid A]^2 = \frac{32}{3}$$

---

3. (a) Since X, Y and Z are independent, V and W are independent.
Therefore $f_{V,W}(v,w) = f_V(v) * f_W(w)$. Find the seperate CDFs of V and W and then differentiate.

$$\begin{aligned}
F_V(v) &= P(V \le v) = P(XY \le v) = P(XY \le v, Y \le v) + P(XY \le v, Y > v) \\
&= P(Y \le v) + P(X \le v/Y, Y > v) \\
&= \int_0^v f_Y(y)dy + \int_v^1 \int_0^{v/y} f_X(x) f_Y(y) dx dy \\
&= v + \int_v^1 \frac{v}{y} dy = v(1 - \log v), \quad 0 \le v \le 1 \\
F_W(w) &= P(W \le w) = P(Z^2 \le w) = P(Z \le \sqrt{w}) = \sqrt{w}, \quad 0 \le w \le 1 \\
f_V(v) &= \frac{dF_V(v)}{dv} = (1 - \log v) + v(-\frac{1}{v}) = \log(1/v), \quad 0 \le v \le 1 \\
f_W(w) &= \frac{dF_W(w)}{dw} = \frac{1}{2\sqrt{w}}, \quad 0 \le w \le 1 \\
f_{V,W}(v,w) &= \frac{\log(1/v)}{2\sqrt{w}}, \quad 0 \le v, w \le 1
\end{aligned}$$

(b)

$$\begin{aligned}
P(XY \le Z^2) &= P(V \le W) = \int_0^1 \int_0^w \frac{\log(1/v)}{2\sqrt{w}} dv dw = \int_0^1 \frac{v(1-\log v)}{2\sqrt{w}}\Big|_{v=0}^{w} dw \\
&= \int_0^1 \frac{\sqrt{w}(1 - \log w)}{2} dw = [\frac{w^{3/2}}{3}(\frac{5}{3} - \log w)]_{w=0}^1 = \frac{5}{9}
\end{aligned}$$

4. (a) Since $E[X] = 0$, We have $E[E[X|Y]] = E[X] = 0$. Hence

$$\mathrm{cov}\big(X, E[X \mid Y]\big) = E\big[XE[X \mid Y]\big] = E\big[E\big[XE[X \mid Y] \mid Y\big]\big] = E\big[\big(E[X \mid Y]\big)^2\big] \ge 0.$$

(b) We can actually prove a stronger statement than what is asked for in the problem, namely that both pairs of random variables have the same covariance (from which it immediately follows that their correlation coefficients have the same sign. We have

$$\mathrm{cov}(Y, E[X \mid Y]\big) = E\big[YE[X \mid Y]\big] = E\big[E[XY \mid Y]\big] = E[XY] = \mathrm{cov}(X, Y).$$

5. The law of total variance states that

$$\mathrm{var}(X) = \mathrm{var}(E[X \mid Y]) + E[\mathrm{var}(X \mid Y)].$$

The first term on the right is the same as $\mathrm{var}(\hat{X})$. It remains to show that $\mathrm{var}(\tilde{X})$ equals $E[\mathrm{var}(X \mid Y)]$. We have

$$E[\tilde{X}] = E\big[X - E[X \mid Y]\big] = E[X] - E[X] = 0,$$

so that

$$\text{var}(\tilde{X}) = E[\tilde{X^2}] = E\big[(X - E[X \mid Y])^2\big].$$

Using the law of iterated expectations, this is the same as

$$E\Big[E\big[(X - E[X \mid Y])^2 \mid Y\big]\Big] = E[\text{var}(X \mid Y)],$$

which concludes the argument.

G1†. We first find $E[X_n|X_{n-1} = k]$. Using the total expectation theorem,

$$\begin{aligned} E[X_n|X_{n-1} = k] = \quad & E[X_n|X_{n-1} = k, (k+1)^{st} \text{ toss is a H}] \cdot P((k+1)^{st} \text{ toss is a H}) \\ & + E[X_n|X_{n-1} = k, (k+1)^{st} \text{ toss is a T}] \cdot P((k+1)^{st} \text{ toss is a T}) \end{aligned}$$

Now, if we are given that $X_{n-1} = k$, then this means that the first time $(n-1)$ heads occurred in succession was on the k^{th} toss.

If in addition we are given that the $(k+1)^{st}$ toss is a H, then this means that the first time n heads occur in succes sion is on the $(k+1)^{st}$ toss, *i.e.* $X_n = k+1$. Hence,

$$E[X_n|X_{n-1} = k, (k+1)^{st} \text{ toss is a H}] = k + 1.$$

However, if the $(k+1)^{st}$ toss is given to be a T, then the first time n heads occur in succession in the part of the sequence starting from the $(k+2)^{nd}$ toss is also the first time that n heads occur in succession in the entire sequence. Since the toss es are independent, the additional number of tosses after the $(k+1)^{st}$ toss for this to happen, has the same distribution as X_n without any conditioning.

This gives:

$$E[X_n|X_{n-1} = k, (k+1)^{st} \text{ toss is a T}] = k + 1 + E[X_n].$$

Substituting in the above,

$$\begin{aligned} E[X_n|X_{n-1} = k] &= p \cdot (k+1) + (1-p) \cdot (k + 1 + E[X_n]) \\ &= k + 1 + (1-p) \cdot E[X_n] \\ \text{Hence,} \qquad E[X_n|X_{n-1}] &= X_{n-1} + 1 + (1-p) \cdot E[X_n] \end{aligned}$$

Taking expectation throughout,

$$\begin{aligned} E\Big[E[X_n|X_{n-1}]\Big] = E[X_n] &= E[X_{n-1}] + 1 + (1-p) \cdot E[X_n] \\ \Rightarrow E[X_n] &= \frac{1}{p} + \frac{1}{p} E[X_{n-1}] \end{aligned}$$

Now, X_1 is the number of tosses till the first head. Hence, X_1 is a geometric random variable with parameter p, and its mean is $E[X_1] = \frac{1}{p}$. Using this as the basis step, we can prove by induction that for all $n \geq 1$,

$$E[X_n] = \sum_{k=1}^{n} \frac{1}{p^k}$$