Massachusetts Institute of Technology
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Spring 2010)

Problem Set 4: Solutions
Due: March 3, 2010

1. (a) From the joint PMF, there are six (x, y) coordinate pairs with nonzero probabilities of occurring. These pairs are (1, 1), (1, 3), (2, 1), (2, 3), (4, 1), and (4, 3). The probability of a pair is proportional to the product of the x and y coordinate of the pair. Because the probability of the entire sample space must equal 1, we have:

$$(1 \cdot 1)c + (1 \cdot 3)c + (2 \cdot 1)c + (2 \cdot 3)c + (4 \cdot 1)c + (4 \cdot 3)c = 1.$$

Solving for c, we get $c = \boxed{\frac{1}{28}}$

(b) There are three sample points for which $Y < X$.

$$\mathbf{P}(Y < X) = \mathbf{P}(\{(2,1)\}) + \mathbf{P}(\{(4,1)\}) + \mathbf{P}(\{(4,3)\}) = \frac{2 \cdot 1}{28} + \frac{4 \cdot 1}{28} + \frac{4 \cdot 3}{28} = \boxed{\frac{18}{28}}$$

(c) There are two sample points for which $Y > X$.

$$\mathbf{P}(Y > X) = \mathbf{P}(\{(1,3)\}) + \mathbf{P}(\{(2,3)\}) = \frac{1 \cdot 3}{28} + \frac{2 \cdot 3}{28} = \boxed{\frac{9}{28}}$$

(d) There is only one sample point for which $Y = X$.

$$\mathbf{P}(Y = X) = \mathbf{P}(\{(1,1)\}) = \frac{1 \cdot 1}{28} = \boxed{\frac{1}{28}}$$

Notice that, using the above two parts:

$$\mathbf{P}(Y < X) + \mathbf{P}(Y > X) + \mathbf{P}(Y = X) = \frac{18}{28} + \frac{9}{28} + \frac{1}{28} = 1$$

as expected.

(e) There are three sample points for which $y = 3$.

$$\mathbf{P}(Y = 3) = \mathbf{P}(\{(1,3)\}) + \mathbf{P}(\{(2,3)\}) + \mathbf{P}(\{(4,3)\}) = \frac{3}{28} + \frac{6}{28} + \frac{12}{28} = \boxed{\frac{21}{28}}$$

(f) In general, for two discrete random variables X and Y for which a joint PMF is defined, we have

$$p_X(x) = \sum_{y=-\infty}^{\infty} p_{X,Y}(x,y) \qquad \text{and} \qquad p_Y(y) = \sum_{x=-\infty}^{\infty} p_{X,Y}(x,y).$$

In this problem the number of possible (X, Y) pairs is quite small, so we can determine the marginal PMFs by enumeration. For example,

$$p_X(2) = \mathbf{P}(\{(2,1)\}) + \mathbf{P}(\{(2,3)\}) = \frac{8}{28}.$$

Overall, we get:

$$p_X(x) = \begin{cases} 4/28, & x = 1; \\ 8/28, & x = 2; \\ 16/28, & x = 4; \\ 0, & \text{otherwise} \end{cases} = \begin{cases} 1/7, & x = 1; \\ 2/7, & x = 2; \\ 4/7, & x = 4; \\ 0, & \text{otherwise} \end{cases}$$

and

$$p_Y(y) = \begin{cases} 7/28, & y = 1; \\ 21/28, & y = 3; \\ 0, & \text{otherwise} \end{cases} = \begin{cases} 1/4, & y = 1; \\ 3/4, & y = 3; \\ 0, & \text{otherwise.} \end{cases}$$

(g) In general, the expected value of any discrete random variable X is given by

$$\mathbf{E}[X] = \sum_{x=-\infty}^{\infty} x p_X(x).$$

For this problem,

$$\mathbf{E}[X] = 1 \cdot \frac{1}{7} + 2 \cdot \frac{2}{7} + 4 \cdot \frac{4}{7} = \boxed{3}$$

and

$$\mathbf{E}[Y] = 1 \cdot \frac{1}{4} + 3 \cdot \frac{3}{4} = \boxed{\tfrac{5}{2}}$$

(h) The variance of a random variable X can be computed as $\mathbf{E}[X^2] - \mathbf{E}[X]^2$ or as $\mathbf{E}[(X - \mathbf{E}[X])^2]$. Here we use the second approach.

$$\text{var}(X) = (1-3)^2 \cdot \frac{1}{7} + (2-3)^2 \cdot \frac{2}{7} + (4-3)^2 \cdot \frac{4}{7} = \boxed{\tfrac{10}{7}}$$

$$\text{var}(Y) = \left(1 - \frac{5}{2}\right)^2 \frac{1}{4} + \left(3 - \frac{5}{2}\right)^2 \frac{3}{4} = \frac{9}{16} + \frac{3}{16} = \boxed{\tfrac{3}{4}}$$

2. a) Using the Total Probability Theorem:

$$\begin{aligned} P(\$1.00 \text{ win in a single game}) &= P(N=1|M=1)P(M=1) + P(N=1|M=2)P(M=2) \\ &\quad + P(N=1|M=3)P(M=3) \\ &= \frac{1}{3}\left(\binom{1}{1}\frac{1}{2} + \binom{2}{1}(\frac{1}{2})^2 + \binom{3}{1}(\frac{1}{2})^3\right) \\ &= \frac{11}{24} \end{aligned}$$

b) Using the Total Expectation Theorem:

$$\begin{aligned} \text{Expected winning} &= \mathbf{E}[N|M=1]P(M=1) + \mathbf{E}[N|M=2]P(M=2) \\ &\quad + \mathbf{E}[N|M=3]P(M=3) \\ &= \frac{1}{3}(1 \times p + 2 \times p + 3 \times p) \\ &= 2 \times p \\ &= \$1 \end{aligned}$$

where p is the probability of a head in the coin toss and since the coin is fair $p = 0.5$.

c)

$$P(M = m|N = 1) = \frac{P(M = m, N = 1)}{P(N = 1)} = \frac{P(N = 1|M = m)P(M = m)}{P(N = 1)}$$

where $P(N = 1)$ is given by part (a) and we denote it with p_a.

$$P(M = m|N = 1) = \begin{cases} \frac{1/2}{3p_a} = \frac{4}{11} & m = 1 \\ \frac{2(1/2)^2}{3p_a} = \frac{4}{11} & m = 2 \\ \frac{3(1/2)^3}{3p_a} = \frac{3}{11} & m = 3 \end{cases}$$

d)

$$\begin{aligned} P(\text{All tosses come head}) &= P(M=1, N=1) + P(M=2, N=2) + P(M=3, N=3) \\ &= P(N=1|M=1)P(M=1) + P(N=2|M=2)P(M=2) \\ &\quad +P(N=3|M=3)P(M=3) \\ &= \frac{1}{3}\left(\binom{1}{1}\frac{1}{2} + \binom{2}{2}(\frac{1}{2})^2 + \binom{3}{3}(\frac{1}{2})^3\right) \\ &= \frac{1}{3}(\frac{1}{2} + \frac{1}{4} + \frac{1}{8}) \\ &= \frac{7}{24} \end{aligned}$$

We denote this probability with p_d.

e) Since the drawn card is replaced in the box, probability of drawing a 1, a 2 or a 3 in any game remains the same, i.e., 1/3. Hence, the games are independent from each other. Now, we define "success" in each game as the event all-heads, which occurs with $p = 7/24$ as calculated in part (d), independent of the game number. This is exactly the case of a geometric distribution with parameter p_d from part (d). The expected value is equal to $1/p_d = 24/7$.

3. (a)

$$\begin{aligned} P(\text{bin}|\text{is empty}) &= P(\text{none of } n \text{ balls fall in bin 1}) && (1) \\ &= \prod_{i=1}^{n} P(\text{ball } i \text{ doesn't fall in bin 1}) && (2) \\ &= [P(\text{ball } i \text{ doesn't fall in bin 1})]^n && (3) \\ &= \left(1 - \frac{1}{n}\right)^n && (4) \end{aligned}$$

(b) Let the random variable X_i be defined as the follows:

$$X_i = \begin{cases} 1 & \text{if bin } i \text{ is empty} \\ 0 & \text{otherwise} \end{cases}$$

Let Y denote the total number of empty bin. Then, we have

$$Y = X_1 + X_2 + \cdots + X_n$$

and

$$E[Y] = E[X_1] + E[X_2] + \cdots + E[X_n].$$

By symmetry, we have

$$E[Y] = n \cdot E[X_1]$$

where

$$E[X_1] = 1 \cdot P(A) + 0 \cdot P(A^c) = \left(1 - \frac{1}{n}\right)^n$$

Thus, we have

$$E[Y] = n \cdot \left(1 - \frac{1}{n}\right)^n$$

4. (a) Determine the joint PMF of H and W, i.e., $p_{H,W}(h, w)$.

$$p_{H,W}(h,w) = p_{W|H}(w|h) \cdot p_H(h) = \begin{cases} \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}, & h = 0, w = 1 \\ \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}, & h = 0, w = 2 \\ \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}, & h = 1, w = 0 \\ \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}, & h = 1, w = 1 \\ 0, & \text{otherwise} \end{cases}$$

(b) Determine $p_{H|W}(0|1)$.

$p_{H|W}(0|1) = \frac{p_{H,W}(0,1)}{p_W(1)} = \frac{3}{4}$

(c) Are W and H independent? Explain why or why not.

No, W and H are not independent. The conditional PMFs of W conditioned on H have different support. i.e, $p_{W|H}(w|h)$ is non-zero when $w = 0$,1 if $h = 1$ and when $w = 1$,2 if $h = 0$.

It can also be quickly noted for example, that $p_{H|W}(0|1) = \frac{3}{4} \neq \frac{2}{3} = p_H(0)$.

5. (a) The answer is NO. Knowing something about V gives information about W. For instance, if we know that V is 6, then W must be 0.

(b)

$$p_V(v) = \begin{cases} \frac{1}{6} & \text{if } v = 2, 3, 5, 6 \\ \frac{2}{6} & \text{if } v = 4 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} E[V] &= 4 \\ Var(V) &= E[(V - E[V])^2] \\ &= E[(V-4)^2] \\ &= \frac{1}{6}(4) + \frac{1}{6}(1) + \frac{2}{6}(0) + \frac{1}{6}(1) + \frac{1}{6}(4) \\ &= \frac{5}{3}. \end{aligned}$$

(c)

$$p_{V,W}(v,w) = \begin{cases} \frac{1}{6} & \text{if } v = 2, 6 \text{ and } w = 0 \\ \frac{1}{6} & \text{if } v = 3 \text{ and } w = 1 \\ \frac{1}{6} & \text{if } v = 5 \text{ and } w = -1 \\ \frac{1}{6} & \text{if } v = 4 \text{ and } w = -2, 2 \\ 0 & \text{otherwise} \end{cases}$$

(d)

$$p_{V|W}(v \mid w > 0) = \begin{cases} \frac{1}{2} & \text{if } v = 3, 4 \\ 0 & \text{otherwise} \end{cases}$$

$$E[V \mid W > 0] = 3.5$$

.

(e)

$$p_{W|V}(w \mid 4) = \begin{cases} \frac{1}{2} & \text{if } w = -2, 2 \\ 0 & \text{otherwise} \end{cases}$$

$$E[W \mid V = 4] = 0$$

$$Var(W \mid V = 4] = \frac{1}{2}(4) + \frac{1}{2}(4) = 4$$

(f)

$$p_{X|V}(x \mid 2) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$p_{X|V}(x \mid 3) = \begin{cases} 1 & \text{if } x = 2 \\ 0 & \text{otherwise} \end{cases}$$

$$p_{X|V}(x \mid 4) = \begin{cases} \frac{1}{2} & \text{if } x = 1, 3 \\ 0 & \text{otherwise} \end{cases}$$

$$p_{X|V}(x \mid 5) = \begin{cases} 1 & \text{if } x = 2 \\ 0 & \text{otherwise} \end{cases}$$

$$p_{X|V}(x \mid 6) = \begin{cases} 1 & \text{if } x = 3 \\ 0 & \text{otherwise} \end{cases}$$

G1†. Notice that only the relative distance between the fly and the spider matters here, and not the absolute positions of the fly and the spider.
Denote:

- A_d the event that initially the spider and the fly are d units apart.
- B_d the event that after one second the spider and the fly are d units apart.

Our approach will be to first apply the (conditional version of the) total expectation theorem to compute $\mathbf{E}(T \mid A_1)$, then use the result to compute $\mathbf{E}(T \mid A_2)$, and similarly compute sequentially $\mathbf{E}(T \mid A_d)$ for all relevant values of d. We will then apply the (unconditional version of the) total expectation theorem to compute $\mathbf{E}(T)$, using the given PMF of d.

We have

$$A_d = A_d B_d + A_d B_{d-1} + A_d B_{d-2}, \qquad \text{if } d > 1.$$

This is because if the spider and the fly are at a distance $d > 1$ apart, then one second later their distance will be d (if the fly moved away from the spider) or $d - 1$ (if the fly did not move) or $d-2$ (if the fly moved towards the spider). We also have, for the case where the spider and the fly start one unit apart,

$$A_1 = A_1B_1 + A_1B_0.$$

Using the total expectation theorem, we obtain

$$\begin{aligned}
\mathbf{E}(T \mid A_d) &= \mathbf{P}(A_dB_d)\mathbf{E}(T \mid A_dB_d) \\
&\quad + \mathbf{P}(A_dB_{d-1})\mathbf{E}(T \mid A_dB_{d-1}) \qquad \text{if } d > 1, \\
&\quad + \mathbf{P}(A_dB_{d-2})\mathbf{E}(T \mid A_dB_{d-2})
\end{aligned}$$

while for the case $d = 1$,

$$\mathbf{E}(T \mid A_1) = \mathbf{P}(A_1B_1)\mathbf{E}(T \mid A_1B_1) + \mathbf{P}(A_1B_0)\mathbf{E}(T \mid A_1B_0).$$

It can be seen based on the problem data that

$$\mathbf{P}(A_1B_1) = 2p, \qquad \mathbf{P}(A_1B_0) = 1 - 2p,$$

$$\mathbf{E}(T \mid A_1B_1) = 1 + \mathbf{E}(T \mid A_1), \qquad \mathbf{E}(T \mid A_1B_0) = 1,$$

so by applying the theorem with $d = 1$, we obtain

$$\mathbf{E}(T \mid A_1) = 2p\,(1 + \mathbf{E}(T \mid A_1)) + (1 - 2p),$$

or

$$\mathbf{E}(T \mid A_1) = \frac{1}{1-2p}.$$

By applying the theorem with $d = 2$, we obtain

$$\mathbf{E}(T \mid A_2) = p\mathbf{E}(T \mid A_2B_2) + (1-2p)\mathbf{E}(T \mid A_2B_1) + p\mathbf{E}(T \mid A_2B_0).$$

We have

$$\begin{aligned}
\mathbf{E}(T \mid A_2B_0) &= 1, \\
\mathbf{E}(T \mid A_2B_1) &= 1 + \mathbf{E}(T \mid A_1), \\
\mathbf{E}(T \mid A_2B_2) &= 1 + \mathbf{E}(T \mid A_2),
\end{aligned}$$

so by substituting these relations in the expression for $\mathbf{E}(T \mid A_2)$, we obtain

$$\begin{aligned}
\mathbf{E}(T \mid A_2) &= p\,(1 + \mathbf{E}(T \mid A_2)) + (1-2p)\,(1 + \mathbf{E}(T \mid A_1)) + p \\
&= p\,(1 + \mathbf{E}(T \mid A_2)) + (1-2p) \cdot \left(1 + \frac{1}{1-2p}\right) + p.
\end{aligned}$$

This equation yields after some calculation

$$\mathbf{E}(T \mid A_2) = \frac{2}{1-p}.$$

Similarly, we obtain for $d > 2$,

$$\begin{aligned}\mathbf{E}(T \mid A_d) &= p\,(1+\mathbf{E}(T \mid A_d)) \\ &\quad +(1-2p)\,(1+\mathbf{E}(T \mid A_{d-1})) \\ &\quad +p\,(1+\mathbf{E}(T \mid A_{d-2}))\end{aligned}$$

Simplifying:

$$\mathbf{E}(T \mid A_d) \;=\; \frac{1}{1-p}\left[1+(1-2p)\mathbf{E}(T \mid A_{d-1})+p\mathbf{E}(T \mid A_{d-2})\right]$$

So $\mathbf{E}(T \mid A_d)$ can be generated recursively for any initial distance d, using the initial conditions $\mathbf{E}(T \mid A_1) = \dfrac{1}{1-2p}$ and $\mathbf{E}(T \mid A_2) = \dfrac{2}{1-p}$ as obtained earlier.

Since d is a random variable taking on values 1,2,3,4,5, we calculate $\mathbf{E}(T \mid A_3), \mathbf{E}(T \mid A_4)$ and $\mathbf{E}(T \mid A_5)$:

$$\begin{aligned}\mathbf{E}(T \mid A_3) &= \frac{1}{1-p}\left[1+\frac{2}{1-p}(1-2p)+\frac{p}{1-2p}\right] \\ \mathbf{E}(T \mid A_4) &= \frac{1}{1-p}\left[1+\frac{2p}{1-p}+\frac{1-2p}{1-p}\left(1+2\frac{1-2p}{1-p}+\frac{p}{1-2p}\right)\right] \\ &= \frac{2}{(1-p)^2}[1+(1-2p)^2] \\ \mathbf{E}(T \mid A_5) &= \frac{1}{1-p}\left[1+(1-2p)\frac{2}{(1-p)^2}[1+(1-2p)^2]+\frac{p}{1-p}\left(1+\frac{2}{1-p}(1-2p)+\frac{p}{1-2p}\right)\right]\end{aligned}$$

Finally, the expected value of T can be obtained using the uniform PMF for the initial distance d and the total expectation theorem:

$$\mathbf{E}(T) = \sum_{d_0} p_d(d_0)\mathbf{E}(T \mid A_{d_0})$$

Since d is a uniform random variable taking on values $1, 2, 3, 4, 5$:

$$\mathbf{E}(T) = \frac{1}{5}[\mathbf{E}(T \mid A_1)+\mathbf{E}(T \mid A_2)+\mathbf{E}(T \mid A_3)+\mathbf{E}(T \mid A_4)+\mathbf{E}(T \mid A_5)]$$

† Optional (not graded for anyone); recommended for 6.431 students.