Problem Set 4: Solutions
Due: March 2, 2011

1. (a) Use the total probability theorem by conditioning on the number of questions that Professor Right has to answer. Let A be the event that she gives all wrong answers in a given lecture, let B_1 be the event that she gets one question in a given lecture, and let B_2 be the event that she gets two questions in a given lecture. Then

$$\mathbf{P}(A) = \mathbf{P}(A|B_1)\mathbf{P}(B_1) + \mathbf{P}(A|B_2)\mathbf{P}(B_2).$$

From the problem statement, she is equally likely to get one or two questions in a given lecture, so $\mathbf{P}(B_1) = \mathbf{P}(B_2) = \frac{1}{2}$. Also, from the problem statement, $\mathbf{P}(A|B_1) = \frac{1}{4}$, and, because of independence, $\mathbf{P}(A|B_2) = (\frac{1}{4})^2 = \frac{1}{16}$. Thus we have

$$\mathbf{P}(A) = \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{16} \cdot \frac{1}{2} = \frac{5}{32}.$$

(b) Let events A and B_2 be defined as in the previous part. Using Bayes's Rule:

$$\mathbf{P}(B_2|A) = \frac{\mathbf{P}(A|B_2)\mathbf{P}(B_2)}{\mathbf{P}(A)}.$$

From the previous part, we said $\mathbf{P}(B_2) = \frac{1}{2}$, $\mathbf{P}(A|B_2) = \frac{1}{16}$, and $\mathbf{P}(A) = \frac{5}{32}$. Thus

$$\mathbf{P}(B_2|A) = \frac{\frac{1}{16} \cdot \frac{1}{2}}{\frac{5}{32}} = \frac{1}{5}.$$

As one would expect, given that Professor Right answers all the questions in a given lecture, it's more likely that she got only one question rather than two.

(c) We start by finding the PMFs for X and Y. The PMF $p_X(x)$ is given from the problem statement:

$$p_X(x) = \begin{cases} \frac{1}{2}, & if x \in \{1, 2\} \\ 0, & otherwise. \end{cases}$$

The PMF for Y can be found by conditioning on X for each value that Y can take on. Because Professor Right can be asked at most two questions in any lecture, the range of Y is from 0 to 2. Looking at each possible value of Y, we find

$$p_Y(0) = \mathbf{P}(Y = 0|X = 0)\mathbf{P}(X = 1) + \mathbf{P}(Y = 0|X = 2)\mathbf{P}(X = 2) = \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{16} \cdot \frac{1}{2} = \frac{5}{32},$$

$$p_Y(1) = \mathbf{P}(Y = 1|X = 1)\mathbf{P}(X = 1) + \mathbf{P}(Y = 1|X = 2)\mathbf{P}(X = 2) = \frac{3}{4} \cdot \frac{1}{2} + 2 \cdot \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{1}{2} = \frac{9}{16},$$

$$p_Y(2) = \mathbf{P}(Y = 2|X = 1)\mathbf{P}(X = 1) + \mathbf{P}(Y = 2|X = 2)\mathbf{P}(X = 2) = 0 \cdot \frac{1}{2} + (\frac{3}{4})^2 \cdot \frac{1}{2} = \frac{9}{32}.$$

Note that when calculating $\mathbf{P}(Y = 1|X = 2)$, we got $2 \cdot \frac{3}{4} \cdot \frac{1}{4}$ because there are two ways for Professor Right to answer one question right when she's asked two questions: either she answers the first question correctly or she answers the second question correctly. Thus, overall

$$p_Y(y) = \begin{cases} 5/32, & \text{if } y = 0; \\ 9/16, & \text{if } y = 1; \\ 9/32, & \text{if } y = 2; \\ 0, & \text{otherwise.} \end{cases}$$

Now the mean and variance can be calculated explicitly from the PMFs:

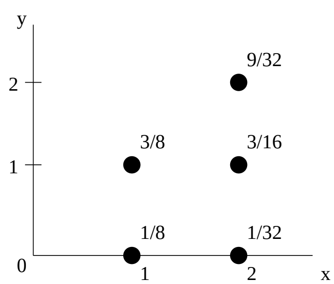
$$\mathbf{E}[X] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2},$$

$$\text{var}(X) = \left(1 - \frac{3}{2}\right)^2 \frac{1}{2} + \left(2 - \frac{3}{2}\right)^2 \frac{1}{2} = \frac{1}{4},$$

$$\mathbf{E}[Y] = 0 \cdot \frac{5}{32} + 1 \cdot \frac{9}{16} + 2 \cdot \frac{9}{32} = \frac{9}{8},$$

$$\text{var}(Y) = \left(0 - \frac{9}{8}\right)^2 \frac{5}{32} + \left(1 - \frac{9}{8}\right)^2 \frac{9}{16} + \left(2 - \frac{9}{8}\right)^2 \frac{9}{32} = \frac{27}{64}.$$

(d) The joint PMF $p_{X,Y}(x,y)$ is plotted below. There are only five possible (x,y) pairs. For each point, $p_{X,Y}(x,y)$ was calculated by $p_{X,Y}(x,y) = p_X(x)p_{Y|X}(y|x)$.



(e) By linearity of expectations,

$$\mathbf{E}[Z] = \mathbf{E}[X + 2Y] = \mathbf{E}[X] + 2\mathbf{E}[Y] = \frac{3}{2} + 2 \cdot \frac{9}{8} = \frac{15}{4}.$$

Calculating var(Z) is a little bit more tricky because X and Y are not independent; therefore we *cannot* add the variance of X to the variance of $2Y$ to obtain the variance of Z. (X and Y are clearly not independent because if we are told, for example, that $X = 1$, then we know that Y cannot equal 2, although normally without any information about X, Y could equal 2.)

To calculate var(Z), first calculate the PMF for Z from the joint PMF for X and Y. For each (x,y) pair, we assign a value of Z. Then for each value z of Z, we calculate $p_Z(z)$ by summing over the probabilities of all (x,y) pairs that map to z. Thus we get

$$p_Z(z) = \begin{cases} 1/8, & \text{if } z = 1; \\ 1/32, & \text{if } z = 2; \\ 3/8, & \text{if } z = 3; \\ 3/16, & \text{if } z = 4; \\ 9/32, & \text{if } z = 6; \\ 0, & \text{otherwise.} \end{cases}$$

In this example, each (x, y) mapped to exactly one value of Z, but this does not have to be the case in general. Now the variance can be calculated as:

$$\text{var}(Z) = \frac{1}{8}\left(1 - \frac{15}{4}\right)^2 + \frac{1}{32}\left(2 - \frac{15}{4}\right)^2 + \frac{3}{8}\left(3 - \frac{15}{4}\right)^2 + \frac{3}{16}\left(4 - \frac{15}{4}\right) + \frac{9}{32}\left(6 - \frac{15}{4}\right)^2 = \frac{43}{16}.$$

(f) For each lecture i, let Z_i be the random variable associated with the number of questions Professor Right gets asked plus two times the number she gets right. Also, for each lecture i, let D_i be the random variable $1000 + 40Z_i$. Let S be her semesterly salary. Because she teaches a total of 20 lectures, we have

$$S = \sum_{i=1}^{20} D_i = \sum_{i=1}^{20} 1000 + 40Z_i = 20000 + 40\sum_{i=1}^{20} Z_i.$$

By linearity of expectations,

$$\mathbf{E}[S] = 20000 + 40\mathbf{E}[\sum_{i=1}^{20} Z_i] = 20000 + 40(20)\mathbf{E}[Z_i] = 23000.$$

Since each of the D_i are independent, we have

$$\text{var}(S) = \sum_{i=1}^{20} \text{var}(D_i) = 20\text{var}(D_i) = 20\text{var}(1000 + 40Z_i) = 20(40^2\text{var}(Z_i)) = 86000.$$

(g) Let Y be the number of questions she will answer wrong in a randomly chosen lecture. We can find $\mathbf{E}[Y]$ by conditioning on whether the lecture is in math or in science. Let M be the event that the lecture is in math, and let S be the event that the lecture is in science. Then

$$\mathbf{E}[Y] = \mathbf{E}[Y|M]\mathbf{P}(M) + \mathbf{E}[Y|S]\mathbf{P}(S).$$

Since there are an equal number of math and science lectures and we are choosing randomly among them, $\mathbf{P}(M) = \mathbf{P}(S) = \frac{1}{2}$. Now we need to calculate $\mathbf{E}[Y|M]$ and $\mathbf{E}[Y|S]$ by finding the respective conditional PMFs first. The PMFs can be determined in an manner analagous to how we calculated the PMF for the number of correct answers in part (c).

$$p_{Y|S}(y) = \begin{cases} \frac{1}{2}\cdot\frac{3}{4} + \frac{1}{2}(\frac{3}{4})^2 = 21/32, & \text{if } y = 0; \\ \frac{1}{2}\cdot\frac{1}{4} + \frac{1}{2}\cdot 2\cdot\frac{1}{4}\cdot\frac{3}{4} = 5/16, & \text{if } y = 1; \\ \frac{1}{2}\cdot 0 + \frac{1}{2}(\frac{1}{4})^2 = 1/32, & \text{if } y = 2; \\ 0, & \text{otherwise.} \end{cases}$$

$$p_{Y|M}(y) = \begin{cases} \frac{1}{2}\cdot\frac{9}{10} + \frac{1}{2}(\frac{9}{10})^2 = 171/200, & \text{if } y = 0; \\ \frac{1}{2}\cdot\frac{1}{10} + \frac{1}{2}\cdot 2\cdot\frac{1}{10}\cdot\frac{9}{10} = 7/50, & \text{if } y = 1; \\ \frac{1}{2}\cdot 0 + \frac{1}{2}(\frac{1}{10})^2 = 1/200, & \text{if } y = 2; \\ 0, & \text{otherwise.} \end{cases}$$

Therefore

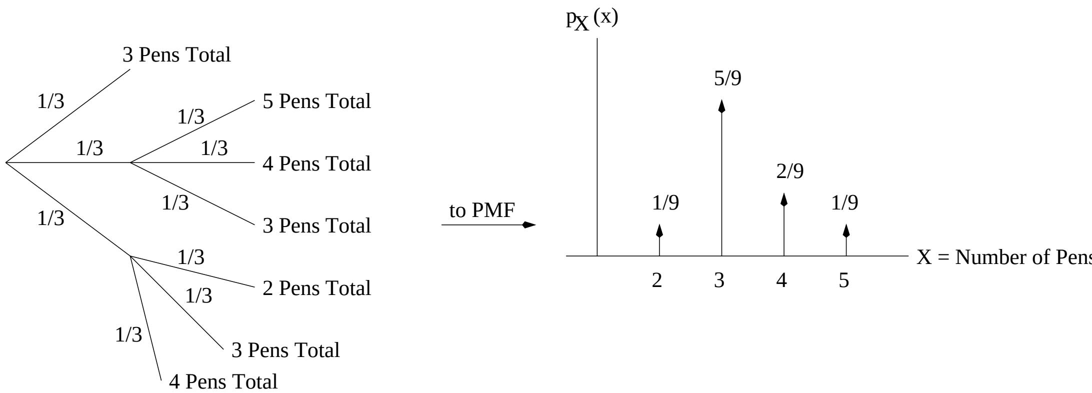
$$\mathbf{E}[Y|S] = 0\cdot\frac{21}{32} + 1\cdot\frac{5}{16} + 2\cdot\frac{1}{32} = \frac{3}{8},$$

$$\mathbf{E}[Y|M] = 0 \cdot \frac{171}{200} + 1 \cdot \frac{7}{50} + 2 \cdot \frac{1}{200} = \frac{3}{20}.$$

This implies that

$$\mathbf{E}[Y] = \frac{3}{20} \cdot \frac{1}{2} + \frac{3}{8} \cdot \frac{1}{2} = \frac{21}{80}.$$

2. (a) We begin by first drawing a tree diagram and from there derive the PMF of X, where X is the total number of pens that he gets on a particular day. The tree diagram and the resulting PMF are shown in the figures below.



From the PMF, we find $\mathbf{P}(A) = \frac{5}{9}$.

(b)

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(A)} = \frac{2/9}{5/9} = \frac{2}{5}$$

(c) Notice that N is the same as the X that we defined in (a) as a useful intermediary. Since we already have the PMF of N, computing $\mathbf{E}[N]$ is immediate:

$$\mathbf{E}[N] = \sum_n n p_N(n) = 2 \cdot \frac{1}{9} + 3 \cdot \frac{5}{9} + 4 \cdot \frac{2}{9} + 5 \cdot \frac{1}{9} = \frac{10}{3}$$

The conditional PMF $p_{N|C}(n)$ is given by

$$p_{N|C}(n) = \begin{cases} 2/3, & n = 4; \\ 1/3, & n = 5; \\ 0, & \text{otherwise} \end{cases}$$

so

$$\mathbf{E}[N|C] = \sum_n n p_{N|C}(n) = 4 \cdot \frac{2}{3} + 5 \cdot \frac{1}{3} = \frac{13}{3}.$$

(d) Using the conditional PMF from the previous part, one can compute

$$\mathbf{E}[N^2|C] = \sum_n n^2 p_{N|C}(n) = 4^2 \cdot \frac{2}{3} + 5^2 \cdot \frac{1}{3} = \frac{57}{3}.$$

Then,

$$\sigma^2_{N|C} = \mathbf{E}[N^2|C] - (\mathbf{E}[N|C])^2 = \frac{57}{3} - \frac{169}{9} = \frac{2}{9}$$

so $\sigma_{N|C} = \frac{\sqrt{2}}{3}$.

(e) The probability that Oscar receives more than three pens on any one day is

$$p_N(4) + p_N(5) = \frac{1}{3}.$$

Since all actions on separate days are independent, $\mathbf{P}(D) = \left(\frac{1}{3}\right)^{16}$.

3. Suppose the president decides to investigate A first. Then her expected costs will be

$$\mathbf{E}[\text{costs}] = D_A + pR_A + (1-p)\cdot(D_B + R_B),$$

whereas if she investigates B first, then

$$\mathbf{E}[\text{costs}] = D_B + (1-p)\cdot R_B + p\cdot(D_A + R_A).$$

In order that the first be smaller than the second, we need

$$pD_B > (1-p)D_A.$$

4. The problem statement says that a topping i is added to a pizza independently of all other toppings and the toppings on all other pizzas. Therefore, the number of pizzas with a certain topping i is independent of the number of pizzas with another topping j, and we may say that:

$$\mathbf{P}_{N_1,N_2,N_3,N_4}(n_1,n_2,n_3,n_4) = \mathbf{P}_{N_1}(n_1)\mathbf{P}_{N_2}(n_2)\mathbf{P}_{N_3}(n_3)\mathbf{P}_{N_4}(n_4)$$

Now we have to find the distribution $\mathbf{P}_{N_i}(n_i)$ for each topping i. The distribution will be binomial where we equate each of the n pizzas with an independent trial and p_i with the probability of success for each trial. Because $p_i = 2^{-i}$, we get:

$$\mathbf{P}_{N_i}(n_i) = \binom{n}{n_i}2^{-in_i}(1-2^{-i})^{n-n_i}, 0 \le n_i \le n$$

and

$$\mathbf{P}_{N_1,N_2,N_3,N_4}(n_1,n_2,n_3,n_4) = \prod_{i=1}^{i=4}\binom{n}{n_i}2^{-in_i}(1-2^{-i})^{n-n_i}, 0 \le n_1,n_2,n_3,n_4 \le n$$

G1†. (a) **Cauchy Schwartz Inequality**

Let us define a random variable $Z = X - tY$, where t is some constant.

$$E[Z^2] = E[(X-tY)(X-tY)] \ge 0$$

$$E[X^2 - 2tXY + t^2Y^2] \ge 0$$

$$E[X^2] - 2tE[XY] + t^2E[Y^2] \ge 0$$

Now, we can look at the left side of the last inequality as a function of t. Since the function is always greater than or equal to 0, the minimum value of the function is also greater than or equal to 0. To minimize the function we take the derivate of the function with respect to t and set it to 0. So,

$$\tfrac{d}{dt}(E[X^2] - 2tE[XY] + t^2E[Y^2]) = 0$$

$$2tE[Y^2] - 2E[XY]) = 0$$

$$t = \tfrac{E[XY]}{E[Y^2]}$$

If $E[Y^2] = 0$, of course dividing by $E[Y^2]$, as above, would be illegal. However, we don't need to worry about this case. If $E[Y^2] = 0$, that would imply that $Y = 0$, in which case the Cauchy Schwartz inequality still holds true : $E[XY]^2 = E[X * 0]^2 = 0 \leq E[X^2]E[0] = 0$. So, we can assume that $E[Y^2] \geq 0$.
Substituting the value of t into the function, we obtain:

$$E[X^2] - 2\tfrac{E[XY]^2}{E[Y^2]} + \tfrac{E[XY]^2}{E[Y^2]^2}E[Y^2] \geq 0$$

$$E[Y^2]E[X^2] - E[XY]^2 \geq 0$$

$$E[XY]^2 \leq E[X^2]E[Y^2]$$

(b) The result can be readily obtained by applying Jensen's inequality: The function $f(x) = 1/x$ is convex and hence $E[1/X] \geq 1/E[X]$.

$$\begin{aligned} E\left[\frac{X}{Y}\right] &= E[X]E\left[\frac{1}{Y}\right] \\ &\geq E[X]\frac{1}{E[Y]} \\ &= 1 \end{aligned}$$

where the first equality in the chain is true by the fact that X and Y are independent (and so is X and $1/Y$,) and the last equality is true because they are identically distributed.

Finally, note that unless Y (and X) are deterministic we strictly have $E[X/Y] > 1$ (as well as $E[Y/X] > 1$.)

(a) Suppose $X = I_A$, then $E[X] = E[I_A]$. $I_A = 1$ for every element of A and 0 for every element of A^c. Therefore, $E[X] = (1)P(\text{outcome is an element of A}) + (0)P(\text{outcome is an element of } A^c) = P(A)$.

(b) We wish to show that the infimum (or the greatest lower bound) of $E[(X-a)^2]$ is attained at $a = E[X]$. $E[(X-a)^2] = E[X^2-2aX+a^2] = E[X^2]-2aE[X]+E[a^2] = E[X^2]-2aE[X]+a^2$. Let $\frac{d}{da}E[(X-a)^2] = 0$, then $-2E[X] + 2a = 0$, or $a = E[X]$. Since $E[(X-a)^2]$ is a convex parabolic function of a, $a = E[X]$ is the value which yields the minimum of $E[(X-a)^2]$. Therefore,

$$var(X) = \inf_{-\infty<a<\infty} E[(X-a)^2] = E[(X - E[X])^2]$$