This week on the problem set you will get practice calculating joint PMFs and covariance. Especially challenging questions, or questions that are not appropriate for an exam, are indicated with one or more asterisks.

1. From the textbook, chapter 2, problems 24, 26, 32, 37.

2. From the supplementary problems, chapter 2, problem 15, 18, 21.

3. We have a biased coin (probability of heads is equal to p). We toss it first for n times and let X denote the number of heads that came up. Then we toss it for another n times. Let Y denote the number of times heads came up in all $2n$ tosses. Find the range, joint PMF and covariance of (X, Y). Compute $p_{X|Y}(k|l)$ for any (k, l) in the range of (X, Y).

Solution: The range is the set $\{(k, l) | 1 \le k \le n, k \le l \le k + n\}$. Let Z be the number of heads in the second round of n tosses, so $Y = X + Z$. Clearly X and Z are independent. Also both X and Z are Binomial(n, p) and Y is Binomial$(2n, p)$. The joint PMF is

$$\begin{aligned} p_{X,Y}(k, l) &= \mathbb{P}(X = k, Y = l) = \mathbb{P}(X = k, Z = l - k) = \mathbb{P}(X = k)\mathbb{P}(Z = l - k) \\ &= \binom{n}{k} p^k (1-p)^{n-k} \binom{n}{l-k} p^{l-k} (1-p)^{n-l+k} = \binom{n}{k}\binom{n}{l-k} p^l (1-p)^{2n-l}. \end{aligned}$$

To compute covariance write $Y = X + Z$ and

$$\begin{aligned} \mathrm{cov}(X, Y) &= \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y) = \mathbb{E}(X(X + Z)) - (\mathbb{E}X)(\mathbb{E}(X + Z)) = \mathbb{E}(X^2 + XZ) \\ &\quad - (\mathbb{E}X)^2 - (\mathbb{E}X)(\mathbb{E}Z) = \mathrm{var}(X) - \mathrm{cov}(X, Z) = np(1-p) - 0 = np(1-p). \end{aligned}$$

Now

$$p_{X|Y}(k|l) = \frac{\mathbb{P}(X = k, Y = l)}{\mathbb{P}(Y = l)} = \frac{\binom{n}{k}\binom{n}{l-k} p^l (1-p)^{2n-l}}{\binom{2n}{l} p^l (1-p)^{2n-l}} = \frac{\binom{n}{k}\binom{n}{l-k}}{\binom{2n}{l}}$$

4. We have two fair coins, one is red and the other one is blue. We start tossing them simultaneously, and repeat this indefinitely. Let X denote the number tosses up to (and including) the time when at least one heads appeared. Let Y denote the number of tosses up to (and including) the time that heads appeared on the red coin. Find the range, joint PMF, and the covariance of (X, Y). Compute $p_{X|Y}(k|l)$ for any (k, l) in the range of (X, Y).

Solution: The range is $\{(k, l) | 1 \le k \le l\}$. To get both $X = k$ and $Y = l$ we first need to get all tails in the first $k - 1$ tosses. Then if $k = l$ we only need to get heads on the red coin in the k-th toss. If $k < l$ we need to get tails on the red coin and heads on the blue coin in k-th toss, then we need to have tails on the red coin in tosses $k + 1$ to $l - 1$ and then we need to get heads on the red coin in the l-th toss. This gives

$$p_{X,Y}(k, l) = \begin{cases} \frac{1}{4^{k-1}} \cdot \frac{1}{2}, & \text{if } k = l \\ \frac{1}{4^k} \frac{1}{2^{l-k}}, & \text{if } k < l. \end{cases}$$

To find the covariance compute

$$\mathbb{E}(XY) = \sum_{k=1}^{\infty} \left(k^2 \mathbb{P}(X = k, Y = k) + \sum_{l=k+1}^{\infty} kl\mathbb{P}(X = k, Y = l) \right).$$

The inner sum is

$$k\sum_{l=k+1}^{\infty} l\mathbb{P}(X=k,Y=l) = \frac{k}{4^k}\sum_{l=k+1}^{\infty}\frac{l}{2^{l-k}} = \frac{k}{4^k}\sum_{i=1}^{\infty}\frac{k+i}{2^i}$$
$$= \frac{k^2}{4^k}\sum_{i=1}^{\infty}2^{-i} + \frac{k}{4^k}\sum_{i=1}^{\infty}\frac{i}{2^i} = \frac{k^2}{4^k} + \frac{k}{4^k}\frac{1/2}{(1-1/2)^2} = \frac{k^2+2k}{4^k},$$

where the second sum was obtained by differentiating geometric series. Then

$$\mathbb{E}(XY) = \sum_{k=1}^{\infty}\left(k^2\frac{1}{2\cdot 4^{k-1}} + \frac{k^2+2k}{4^k}\right) = 3\sum_{k=1}^{\infty}\frac{k^2}{4^k} + 2\sum_{k=1}^{\infty}\frac{k}{4^k}$$
$$= \frac{3}{4}\sum_{k=1}^{\infty}\frac{(k+1)k}{4^{k-1}} - \sum_{k=1}^{\infty}\frac{k}{4^k} = \frac{3}{4}\frac{2}{(1-1/4)^3} - \frac{1/4}{(1-1/4)^2} = 28/9.$$

On the other hand X is Geometric(3/4) and Y is Binomial(1/2) so $\mathbb{E}(X) = 4/3$ and $\mathbb{E}(Y) = 2$ and

$$\mathrm{cov}(X,Y) = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y) = 28/9 - 8/3 = 4/9.$$

When $k = l$ we have

$$p_{X|Y}(k|k) = \frac{p_{X,Y}(k,k)}{p_Y(k)} = \frac{\frac{1}{4^{k-1}}\frac{1}{2}}{\frac{1}{2^{k-1}}} = \frac{1}{2^{k-1}}.$$

When $k < l$ we have

$$p_{X|Y}(k|l) = \frac{p_{X,Y}(k,l)}{p_Y(l)} = \frac{\frac{1}{4^k}\frac{1}{2^{l-k}}}{\frac{1}{2^l}} = \frac{1}{2^k}.$$

5. Recall Problem 7 from the last week homework (n people coming to your party). Compute the variance of the number of presents you receive.

Solution: If X_i is the indicator function that the i-th person gives you a present then for the total number of presents $X = \sum_{i=1}^n X_i$ we get

$$\mathrm{var}(X) = \sum_{i=1}^{n}\mathrm{var}(X_i) = np(1-p).$$

6. Let (X,Y) be a pair of random variables whose range is $\{(k,l)|1 \le k \le 10, 1 \le l \le 10\}$.
   (a) Express $p_{X|Y}(k|l)$ in terms of the values of the joint PMF $p_{X,Y}$.
   (b) Express $p_{X|Y}(k|l)$ in terms of the values of $p_{Y|X}(l|k)$ and the marginal PMFs p_X and p_Y.

Solution:

(a)

$$p_{X|Y}(k|l) = \frac{\mathbb{P}(X=k,Y=l)}{\mathbb{P}(Y=l)} = \frac{p_{X,Y}(k,l)}{\sum_{i=1}^{10}p_{X,Y}(i,l)}.$$

(b) Apply the Bayes formula to get

$$p_{X|Y}(k|l) = \mathbb{P}(X = k|Y = l) = \frac{\mathbb{P}(Y = l|X = k)\mathbb{P}(X = k)}{\sum_{i=1}^{10} \mathbb{P}(Y = l|X = i)\mathbb{P}(X = i)} = \frac{p_{Y|X}(l|k)p_X(k)}{\sum_{i=1}^{10} p_{Y|X}(l|i)p_X(i)}.$$

7. Recall question 7 from Problem Set 2. Assume that the loser in the game pays the winner the amount $\$x$, which is defined as follows: x is the difference between the highest card of the winner and the highest card of the loser. Your friend tells you that his higher card is k, and wants you to tell him if, given this information, the conditional expectation of his income in the game is positive or negative.
   (a) Find the expression for the conditional expectation of the your friends income.
   (b) Try to simplify the expression and determine for which values of k it is positive and for which values of k it is negative.

Solution: If X is the highest card of your friend and Y of his opponent then we have to find $\mathbb{E}(X - Y|X = k) = \mathbb{E}(X|X = k) - \mathbb{E}(Y|X = k) = k - \mathbb{E}(Y|X = k)$. Observe that

$$\mathbb{P}(Y = 20|X = k) = \frac{\mathbb{P}(X = k, Y = 20)}{\mathbb{P}(X = k)} = \frac{(k-1)\cdot 17}{(k-1)\binom{18}{2}} = \frac{17}{\binom{18}{2}},$$

since there are $(k-1)\binom{18}{2}$ ways to give cards so that your friend higher card is k and $(k-1)\cdot 17$ ways so that your friend's higher card is k and his opponent's 20. Using the same argument we have

$$\mathbb{P}(Y = 19|X = k) = \frac{16}{\binom{18}{2}},$$

all the way to

$$\mathbb{P}(Y = k+1|X = k) = \frac{k-2}{\binom{18}{2}}.$$

However $\mathbb{P}(Y = k|X = k) = 0$ and after this point we have to argue a bit more carefully, that the number of ways that your friend's higher card is k and his opponent's higher card is l (for $l < k$) is $(l-1)(k-3)$. This gives

$$\mathbb{P}(Y = k-1|X = k) = \frac{(k-2)(k-3)}{(k-1)\binom{18}{2}},$$

$$\mathbb{P}(Y = k-2|X = k) = \frac{(k-3)(k-3)}{(k-1)\binom{18}{2}},$$

all the way to

$$\mathbb{P}(Y = 2|X = k) = \frac{1\cdot(k-3)}{(k-1)\binom{18}{2}}.$$

Now just plug it in

$$\mathbb{E}(X - Y|X = k) = k - \sum_{i=k+1}^{20} i\frac{i-3}{\binom{18}{2}} - \sum_{i=2}^{k-1} i\frac{(i-1)(k-3)}{(k-1)\binom{18}{2}}$$

which can be simplified with enough patience.

First compute the term

$$\sum_{i=k+1}^{20} i(i-3) = \sum_{i=k+1}^{20} i^2 - 3\sum_{i=k+1}^{20} i = \sum_{i=1}^{20} i^2 - \sum_{i=1}^{k} i^2 - 3\sum_{i=1}^{20} i + 3\sum_{i=1}^{k} i$$
$$= \frac{20\cdot 21\cdot 41}{6} - \frac{k(k+1)(2k+1)}{6} - 3\cdot\frac{20\cdot 21}{2} + 3\cdot\frac{k(k+1)}{2} = 2240 - \frac{k(k+1)(k+5)}{3}.$$

Then compute the term

$$\sum_{i=2}^{k-1} i(i-1) = \sum_{i=1}^{k-1} i^2 - 1 - \sum_{i=1}^{k-1} i + 1 = \frac{(k-1)k(2k-1)}{6} - \frac{(k-1)k}{2} = \frac{(k-2)(k-1)k}{3}.$$

We get

$$\mathbb{E}(X-Y|X=k) = k - \sum_{i=k+1}^{20} i\frac{i-3}{\binom{18}{2}} - \sum_{i=2}^{k-1} i\frac{(i-1)(k-3)}{(k-1)\binom{18}{2}}$$
$$= k - \frac{2240}{\binom{18}{2}} + \frac{\frac{k(k+1)(k+5)}{3}}{\binom{18}{2}} - \frac{k-3}{(k-1)\binom{18}{2}}\cdot\frac{(k-2)(k-1)k}{3}$$
$$= \frac{k(11k + 3\cdot\binom{18}{2} - 1)}{3\cdot\binom{18}{2}} - \frac{2240}{\binom{18}{2}}$$

For this to be positive we would need

$$11k^2 + 458k - 6720 > 0.$$

By solving quadratic equation we find one root at around 11.497 so the answer is that it makes sense for your friend to stay in the game if $k \geq 12$.