Problem Set 4 Solutions

Note: Thanks to Kevin Lee for some of the solutions.

1. The winner must win the last game, as well as three others and the other person can win between zero and three inclusive. Thus, the probability mass function is

$$f(k) = \binom{k-1}{3}\left(p^4(1-p)^{k-4} + (1-p)^4 p^{k-4}\right)$$

The two terms are for the cases where Ann wins or Bob wins. For $p = 0.6$, this formula gives us:

k	4	5	6	7
$f(k)$	0.1552	0.2688	0.29952	0.27648

The expectation of this variable is therefore $4 \cdot f(4) + 5 \cdot f(5) + 6 \cdot f(6) + 7 \cdot f(7) = 5.69728$. The variance is $f(4) \cdot (4 - 5.69728)^2 + f(5) \cdot (5 - 5.69728)^2 + f(6) \cdot (6 - 5.69728)^2 + f(7) \cdot (7 - 5.69728)^2 = 1.07444$

2. First Solution: In order for k red balls to be drawn, k red balls must be drawn first and 1 blue ball must be drawn after. Thus, the mass function is

$$f(k) = \frac{r(r-1)\cdots(r-k+1)b}{(r+b)(r+b-1)\cdots(r+b-k)} = \frac{r!b}{(r+b)!}\frac{(r+b-k-1)!}{(r-k)!} = \frac{\binom{b+r-k-1}{b-1}}{\binom{b+r}{b}}.$$

Note that we can alternately derive the last expression directly by noting that if k red balls are drawn, the $k+1$st ball is blue and then $b-1$ blue balls are drawn out of the remaining $b+r-k-1$ balls. Since between 0 and r red balls are drawn, the expectation is given by the formula

$$\begin{aligned}
E(k) &= \sum_{k=0}^{r} kf(k) = \sum_{k=0}^{r} k\frac{\binom{b+r-k-1}{b-1}}{\binom{b+r}{b}} = \frac{1}{\binom{b+r}{b}}\sum_{k=0}^{r}[r-(r-k)]\binom{b+r-k-1}{r-k} \\
&= \frac{1}{\binom{b+r}{b}}\left[r\sum_{k=0}^{r}\binom{b+r-k-1}{r-k} - b\sum_{k=0}^{r}\binom{b+r-k-1}{r-k-1}\right] = \frac{1}{\binom{b+r}{b}}\left[r\sum_{t=0}^{r}\binom{b-1+t}{t} - b\sum_{t=0}^{r-1}\binom{b+t}{t}\right] \\
&= \frac{1}{\binom{b+r}{b}}\left[r\binom{b+r}{r} - b\binom{b+r}{r-1}\right] = \frac{b!r!}{(b+r)!}\left[\frac{(b+r)!}{b!(r-1)!} - \frac{(b+r)!}{(r-1)!(b-1)!(b+1)}\right] = r - \frac{br}{b+1} = \frac{r}{b+1}.
\end{aligned}$$

Second Solution: Let R be the set of red balls. For $i \in R$, let I_i be the indicator function for the event that ball i is drawn. Then, we see that $E(k) = \sum_{i=1}^{r} E(I_i)$. But $E(I_i) = P(I_i) = \frac{1}{b+1}$, since red ball i is drawn if it appears before all blue balls in some ordering. So we can sum to find that $E(k) = \frac{r}{b+1}$.

3. Prisoner $n+2$ must be given a card that is not his. Therefore, we must distribute the remaining $n+1$ cards to the remaining $n+1$ prisoners, but allowing for the possibility that the prisoner whose card prisoner $n+2$ has can get "his own card" (If we were to swap his card and with prisoner $n+2$'s card). Thus, this gives us a total of $H_{n+1} + H_n$ ways to distribute the remaining cards. Since there are $n+1$ cards that prisoner $n+2$ can get, the recursion is

$$H_{n+2} = (n+1)(H_{n+1} + H_n)$$

If we reindex, we can get that $H_{n+1} = n(H_n + H_{n-1})$. This means that $H_{n+1} - (n+1)H_n = -H_n + nH_{n-1}$. If we let $U_n = H_n - nH_{n-1}$, then our recurrence becomes $U_n = -U_{n-1}$. Given that $H_0 = 1$ and $H_1 = 0$, we know that $U_1 = -1$. Thus, $U_n = (-1)^n$. This implies that $H_n = (-1)^n + nH_{n-1} \implies H_n = n!\sum_{k=0}^{n}\frac{(-1)^k}{k!}$. Notice that $\lim_{n\to\infty} H_n/n! = 1/e$.

In order to solve the matching problem, note that the solution when we want r elements to be the same is $\binom{n}{r}\frac{H_{n-r}}{n!} = \frac{n!}{r!(n-r)!}\frac{(n-r)!}{n!}\sum_{k=0}^{n-r}\frac{(-1)^k}{k!} = \frac{1}{r!}\left(\frac{1}{2!} - \frac{1}{3!} + \cdots \pm \frac{1}{(n-r)!}\right)$. This is the desired formula.

4. (a) Since the case is symmetric whether there is a curse or not, we don't need to consider it. If the better team will ultimately win, this happens with probability $\frac{2^4}{3^7}$. Otherwise, the probability is $\frac{2^3}{3^7}$. This gives us a total probability of $\frac{2^3+2^4}{3^7} = \frac{2^3}{3^6} = \frac{8}{729} \approx 1.10\%$.
(b) If there is a curse, their probability of losing the first three games is $\frac{2^3}{3^3}$. If there isn't a curse, their probability of losing the first three games is $\frac{1}{3^3}$. Therefore the conditional probability would be $\frac{2^3}{2^3+1} = \frac{8}{9} = 88.9\%$.
(c) If the curse does exist, then there is a $\frac{1}{3^4}$ probability that they will win the next four games. Otherwise, there is a $\frac{2^4}{3^4}$ probability. Thus, using the earlier calculation about conditional probability, we have that the total probability is $\frac{1}{9}\frac{2^4}{3^4} + \frac{8}{9}\frac{1}{3^4} = \frac{24}{729} = \frac{8}{243} \approx 3.29\%$. This is lower than the probability that they would win the next four games if the curse were not an issue, which is $\frac{1}{2^4} = \frac{1}{16} = 6.25\%$.
(d) Using the previous calculations, we get that the probability is $\frac{\frac{8}{9}\frac{1}{3^4}}{\frac{8}{243}} = \frac{1}{3} \approx 66.7\%$

5. (a) $D = B \cup C \cup E$.
(b) By inclusion-exclusion on D we have

$$\mathbb{P}(D) = \mathbb{P}(B) + \mathbb{P}(C) + \mathbb{P}(E) - \mathbb{P}(B \cap C) - \mathbb{P}(C \cap E) - \mathbb{P}(B \cap E) + \mathbb{P}(B \cap C \cap E) = \mathbb{P}(B) + \mathbb{P}(C) + \mathbb{P}(E) - \mathbb{P}(A),$$

where all the other intersections are empty. Rearranging, we obtain $\mathbb{P}(E) = \mathbb{P}(D) - \mathbb{P}(B) - \mathbb{P}(C) + \mathbb{P}(A)$.
(c) Because X and Y are discrete, they each have a countable set of possible values $\{x_i\}$ and $\{y_j\}$. Then, we have

$$\begin{aligned}\mathbb{P}(X \leq x, Y \leq y) = \sum_{x_i \leq x}\sum_{y_j \leq y} \mathbb{P}(X = x_i, Y = y_j) &= \sum_{x_i \leq x}\sum_{y_j \leq y} \mathbb{P}(X = x_i)\mathbb{P}(Y = y_j) \\ &= \left(\sum_{x_i \leq x} \mathbb{P}(X = x_i)\right)\left(\sum_{y_j \leq y} \mathbb{P}(Y = y_j)\right) = \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y),\end{aligned}$$

so $\{X \leq x\}$ and $\{Y \leq y\}$ are independent.
(d) Note that we have by **(b)** that

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X \leq x, Y \leq y) - \mathbb{P}(X < x, Y \leq y) - \mathbb{P}(X \leq x, Y < y) + \mathbb{P}(X < x, Y < y).$$

Notice that

$$\mathbb{P}(X < x, Y \leq y) = \lim_{t \to x^-} \mathbb{P}(X \leq x, Y \leq y) = \lim_{t \to x^-} \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y) = \mathbb{P}(X < x)\mathbb{P}(Y \leq y).$$

Likewise, we have that $\mathbb{P}(X \leq x, Y < y) = \mathbb{P}(X \leq x)\mathbb{P}(Y < y)$ and $\mathbb{P}(X < x, Y < y) = \mathbb{P}(X < x)\mathbb{P}(Y < y)$. Substituting in, we see that

$$\begin{aligned}\mathbb{P}(X = x, Y = y) &= \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y) - \mathbb{P}(X < x)\mathbb{P}(Y \leq y) - \mathbb{P}(X \leq x)\mathbb{P}(Y < y) + \mathbb{P}(X < x)\mathbb{P}(Y < y) \\ &= [\mathbb{P}(X \leq x) - \mathbb{P}(X < x)][\mathbb{P}(Y \leq y) - \mathbb{P}(Y < y)] = \mathbb{P}(X = x)\mathbb{P}(Y = y),\end{aligned}$$

so $\{X = x\}$ and $\{Y = y\}$ are independent.