Homework 7 Solutions

Question 1

1. A coin has a p chance of turning up heads and $1-p$ chance of tails. Each throw is independent. I toss the coin n times. What is the chance of exactly k heads? Do not just state the formula for this. You need to provide a proof showing why the formula is true.

2. In part (1), what is the expected value of the total number of heads?

3. Now consider the following process, there are two coins: one has heads on both sides and the other has tails on both sides. With probability p, I select the first coin and with probability $1-p$, I select the other coin. I now take only my selected coin and toss it n times. Let E_i be the event that the i-th roll was heads. What is $\Pr(E_i)$? Are the events E_i and E_j $(i \neq j)$ independent?

4. In part (3), what is the expected value of the total number of heads?

Solution:

1. For **any** fixed configuration of exactly k heads, the probability is $p^k(1-p)^{n-k}$. Since there are $\binom{n}{k}$ number of ways of getting exactly k heads, the total probability is $\binom{n}{k}p^k(1-p)^{n-k}$.

2. Let X be the total number of heads and X_i be the result of the i^{th} trial (e.g. 1 for head and 0 for tail). By linearity of expectation,

$$\mathbb{E}[X] = \mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i] = np.$$

3. Let A be the event that the one with heads on both sides is selected. By the law of total probability,

$$\begin{aligned}\mathbb{P}(E_i) &= \mathbb{P}(E_i \cap A) + \mathbb{P}(E_i \cap A^c) \\ &= \mathbb{P}(E_i|A)\mathbb{P}(A) + \mathbb{P}(E_i|A^c)\mathbb{P}(A^c) \\ &= 1 \cdot p + 0 \cdot (1-p) = p.\end{aligned}$$

Besides,

$$\begin{aligned}\mathbb{P}(E_i \cap E_j) &= \mathbb{P}(E_i \cap E_j|A)\mathbb{P}(A) + \mathbb{P}(E_i \cap E_j|A^c)\mathbb{P}(A^c) \\ &= 1 \cdot p + 0 \cdot (1-p) = p.\end{aligned}$$

We have $p = \mathbb{P}(E_i \cap E_j) \neq \mathbb{P}(E_i)\mathbb{P}(E_j) = p^2$. Therefore, the events E_i and E_j $(i \neq j)$ are not independent.

4. If the coin with heads on both sides is selected, which happens with probability p, we will get n heads out of n times; if the coin with tails on both sides is selected, which happens with probability $1-p$, we will get 0 head out of n times. Therefore, the expected value is $pn + (1-p) \cdot 0 = np$.

Question 2

There are n cities. Between every pair of cities there either is a road or there is no road. Prove that for any such road network, you can always divide the cities into two disjoint sets A and B such that at least half the roads in the network are between cities in two different sets. Every city in the network must be put into exactly one of the two sets A and B.

(**Hint:** Do the division of the cities into the two sets randomly. Also, use the linearity of expectation)

Solution: Let $a_{i,j} = 1$ if there is a road between city i and j, and $a_{i,j} = 0$ otherwise (there is no randomness here since the road network is predetermined). Suppose $\frac{1}{2}\sum_{i\neq j} a_{i,j} = m$, which means there are m roads totally. For each city independently, we assign it to set A with probability $1/2$ and set B with probability $1/2$. Let the random variable $X_{i,j} = \mathbf{1}\{\text{city i and j are not in the same set}\}$. It is easy to see $\mathbb{P}(X_{i,j} = 1) = 1/2$ (either i in A and j in B or j in A and i in B, and there are 4 ways of dividing them in total). The expected number of roads between cities is

$$\begin{aligned}
\mathbb{E}[\frac{1}{2}\sum_{i\neq j} a_{i,j}X_{i,j}] &= \frac{1}{2}\sum_{i\neq j} a_{i,j}\mathbb{E}[X_{i,j}] \\
&= \frac{1}{2}\sum_{i\neq j} a_{i,j}\mathbb{P}[X_{i,j}] \\
&= \frac{1}{2}\sum_{i\neq j} \frac{a_{i,j}}{2} \\
&= \frac{1}{4}\sum_{i\neq j} a_{i,j} \\
&= \frac{m}{2}.
\end{aligned}$$

This means the average number of roads between cities from all possible divisions of the cities is $\frac{m}{2}$. Since the maximum element in a set must be at least as large as its average, we can conclude out of the 2^n possible ways of dividing the cities, at least one of them should have at least half the roads in the network are between cities in two different sets.

Question 3

A fair coin is flipped n times independently. A “run" is a maximal sequence of consecutive flips that are all the same. For example, the sequence $HTHHHTTH$ with $n = 8$ has five runs, namely H, T, HHH, TT,H. What is the expected number of runs in a sequence of n throws?

(**Hint:** Define appropriate indicator random variables and then use the linearity of expectation.)

Solution: Let $(\Omega, \Pr)$ be the probability space of all possible results from n coin tosses. Let X be a random variable over Ω where $X(\omega)$ is the number of runs in the result $\omega \in \Omega$. Let X_i, $1 \leq i \leq n$, be the indicator random variable where $X_i = 1$ if toss i starts a new sequence, and $X_i = 0$ if toss i

continues the current sequence. Note that $X_1 = 1$ since the first throw will always start a new sequence. $\forall i > 1$, $X_i = \frac{1}{2}$. Then,

$$X = \sum_{i=1}^{n} X_i$$

Using linearity of expectation we can derive $E(X)$:

$$\begin{aligned} E(X) &= E\left(\sum_{i=1}^{n} X_i\right) \\ &= \sum_{i=1}^{n} E(X_i) \\ &= 1 + \sum_{i=2}^{n} \frac{1}{2} \\ &= \frac{n+1}{2} \end{aligned}$$

Question 4

A coin has a p chance of coming up heads. I toss it repeatedly until a heads shows up, and each coin toss is independent. What is the expected number of times that I will have to toss the coin? You may use the following formula without proof:

$$\sum_{i=0}^{\infty} r^i = \frac{1}{1-r}$$

The above formula holds for all real numbers $r \in (0, 1)$.

Solution: We assume that we cannot simply cite Theorem 5 from Lecture Notes 7 (The expected number of trials before the first success in a Bernoulli trial with probability of success p), since otherwise the provided formula would not be necessary.

Let $(\Omega, \Pr)$ be the probability space where Ω is all coin tosses of the form $T^{i-1}H$ for $i \in \mathbb{N}$. Let X be a random variable s.t $\forall \omega \in \Omega, \omega = T^{i-1}H$, we have $X(\omega) = i$. We seek to find $E(X)$:

$$E(X) = \sum_{\omega \in \Omega} X(\omega) \cdot \Pr(\omega) = \sum_{i=1}^{\infty} i \cdot \Pr(X = i) \qquad \text{(Theorem 2)}$$

$\forall i, \Pr(X = i) = (1-p)^{i-1}p$. This is because we must have a sequence of tails thrown of length $i - 1$

before an eventual success on throw i. Then,

$$\begin{aligned}
E(X) &= \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1} p \\
&= \frac{p}{1-p} \sum_{i=1}^{\infty} i \cdot (1-p)^{i} \\
&= \frac{p}{1-p} \left[\sum_{i=1}^{\infty} (1-p)^{i} + \sum_{i=1}^{\infty} (i-1)(1-p)^{i} \right] \\
&= \frac{p}{1-p} \left[\left(\sum_{i=0}^{\infty} (1-p)^{i} \right) - 1 + \sum_{i=0}^{\infty} i \cdot (1-p)^{i+1} \right] \\
&= \frac{p}{1-p} \left[\frac{1}{1-(1-p)} - 1 + \sum_{i=1}^{\infty} i \cdot (1-p)^{i+1} \right] && \text{(given)} \\
&= \frac{p}{1-p} \left[\frac{1-p}{p} + \frac{(1-p)^2}{p} \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1} p \right] \\
E(X) &= 1 + (1-p) \cdot E(X) \\
p \cdot E(X) &= 1 \\
E(X) &= \frac{1}{p}
\end{aligned}$$

□

(**Note:** The calculation above in order to evaluate the sum of the series $\sum_{i=1}^{\infty} i \cdot (1-p)^{i-1} p$, is equivalent to setting $S = \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1} p$, considering $(1-p)S$ and taking the difference to derive

$$S - (1-p)S = pS = \sum_{i=1}^{\infty} \cdot (1-p)^{i-1} p.$$

You can then use the provided formula in the question to derive that

$$pS = \sum_{i=1}^{\infty} \cdot (1-p)^{i-1} p = p \times \frac{1}{1-(1-p)} = 1.$$

where we are using $r = (1-p)$ in the provided formula (in the question).

This yields $pS = 1$ or $S = \frac{1}{p}$.)

Question 5

Every time we purchase a kid's meal at Taco Bell, we are graciously presented with a miniature "Racing Rocket" car together with a launching device which enables us to project our new vehicle across any tabletop or smooth floor at high velocity. Truly, our delight knows no bounds.

Racing Rocket cars come in n different colors. The color of the car awarded to us by the kind server at the Taco Bell register is selected uniformly and independently at random among these n colors. What is the expected number of kid's meals that we must purchase in order to acquire at least one Racing Rocket car of each color? You may leave your answer in terms of the sum of a simple series. In solving the problem, you may use the result of Q.4 if you wish (even if you haven't solved Q.4 yet.)

(**Hint:** Consider how long it takes to get a car with a new color that you haven't already received. Then use the linearity of expectation.)

Solution:
Let T be the total number of trials needed to collect all n colors. The expected value of T, $\mathbb{E}[T]$, can be found by summing the expected number of trials needed to collect each successive new car. Let t_i be the time to collect the $i-$th color after $i-1$ colors have been collected. So $T = t_1 + t_2 + ... + t_n$, and the probability of getting a new car is $p_i = \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$.
Therefore, the expected number of trials to collect all n colors is:

$$\begin{aligned}
\mathbb{E}[T] &= \mathbb{E}[t_1 + t_2 + ... + t_n] \\
&= \mathbb{E}[t_1] + \mathbb{E}[t_2] + ... + \mathbb{E}[t_n] \\
&= \frac{1}{p_1} + \frac{1}{p_2} + ... + \frac{1}{p_n} \\
&= \frac{n}{n} + \frac{n}{n-1} + ... + \frac{n}{1} \\
&= n(\frac{1}{1} + \frac{1}{2} + ... + \frac{1}{n})
\end{aligned}$$

Question 6

(1) Prove Markov's theorem (**Hint:** divide the terms in the expectation into two groups: one for when the random variable has a large value and the other for when it has a small value).

Why does the theorem require that the random variable be non-negative? Show an example where if we have a random variable that could be negative, then the bound given by the theorem does not hold.

Solution: The expectation of X, $\mathbb{E}[X]$, can be written as:

$$\mathbb{E}[X] = \sum_x x \cdot \Pr(X = x)$$

We divide the terms into two groups: when X has a value less than a and when X has a value of at least a:

$$\mathbb{E}[X] = \sum_{x<a} x \cdot \Pr(X = x) + \sum_{x \geq a} x \cdot \Pr(X = x)$$

Since X is non-negative and $x < a$ for the first group, we can ignore the first sum:

$$\mathbb{E}[X] \geq \sum_{x \geq a} x \cdot \Pr(X = x)$$

Now, since $x \geq a$ for all terms in the second sum, we can replace x with a to obtain a lower bound:

$$\mathbb{E}[X] \geq \sum_{x \geq a} a \cdot \Pr(X = x) = a \cdot \Pr(X \geq a)$$

Rearranging gives Markov's inequality:

$$\Pr(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

Why the Non-Negative Condition is Necessary:

The theorem requires X to be non-negative to ensure that the expectation $\mathbb{E}[X]$ does not underestimate the sum $\sum_{x \geq a} x \cdot \Pr(X = x)$. If X can take negative values, this sum might be much larger than the expectation of X, violating the inequality.

Example with a Negative Random Variable:

Consider a random variable Y which takes values -10 with probability 0.9 and 10 with probability 0.1. Then,

$$\mathbb{E}[Y] = -10 \cdot 0.9 + 10 \cdot 0.1 = -8$$

Applying Markov's theorem to check $\Pr(Y \geq 5)$:

$$\Pr(Y \geq 5) \leq \frac{\mathbb{E}[Y]}{5}$$

But this yields:

$$\Pr(Y \geq 5) \leq \frac{-8}{5} = -1.6$$

This result is impossible as probabilities cannot be negative and indeed there is a 10% chance of Y being more than 5.

(2) Prove that Var(X) $= E(X^2) - E(X)^2$.

Solution:

$$\begin{aligned}
\text{Var}[X] &= \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \\
&= \mathbb{E}\left[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2\right] \\
&= \mathbb{E}\left[X^2\right] - 2\mathbb{E}[X]\mathbb{E}\left[X\right] + \mathbb{E}[X]^2 \\
&= \mathbb{E}\left[X^2\right] - 2\mathbb{E}[X]^2 + \mathbb{E}[X]^2 \\
&= \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2
\end{aligned}$$

$\square$

(**Note:** in this calculation we have used the linearity of expectation, and also the fact that $E[X]$ is a constant (fixed number for a given random variable X). Thus, for instance, $E[E[X]] = E[X]$.)