Assignment 4 Solutions

COMP2804 Fall 2019

November 12, 2019

Name: Lenny Learning Combinatorics
Student ID: 100000000

1 Rolling two D20

For these questions we're working in a uniform sample space $S = \{(d_1, d_2) : d_1, d_2 \in \{1, \ldots, 20\}\}$ of size $20 \cdot 20 = 400$ and it helps to explicitly know what A, B, and C, are.

$$A = \{(13,1), (13,2), (13,3), \ldots, (13,19), (13,20)\}$$

$$B = \{(1,14), (2,13), (3,12), \ldots, (13,2), (14,1)\}$$

$$C = \{(1,20), (2,19), (3,18), \ldots, (19,2), (20,1)\}$$

We see that $|A| = 20$, $|B| = 14$, and $|C| = 20$.

$$\Pr(A) = \frac{|A|}{|S|} = \frac{20}{400} = \frac{1}{20}$$

$$\Pr(B) = \frac{|B|}{|S|} = \frac{14}{400} = \frac{7}{200}$$

$$\Pr(C) = \frac{|C|}{|S|} = \frac{20}{400} = \frac{1}{20}$$

1. $A \cap B = \{(13,2)\}$ so $|A \cap B| = 1$ and

   $$\Pr(A \cap B) = \frac{|A \cap B|}{|S|} = 1/400 \neq \Pr(A) \cdot \Pr(B) = \frac{1}{20}\frac{7}{200} = \frac{7}{4000} \ .$$

   Therefore A and B are not independent.

2. $A \cap C = \{(13,8)\}$ so $|A \cap C| = 1$ and

   $$\Pr(A \cap C) = \frac{|A \cap C|}{|S|} = 1/400 = \frac{1}{20} \cdot \frac{1}{20} = \Pr(A) \cdot \Pr(B) \ .$$

   Therefore A and C are independent.

2 Randomized Leader Election

1. Person x_i leaves the circle in the first round if they toss heads and their two neighbours x_{i-1} and x_{i+1} toss tails. Therefore, no two adjacent people leave the circle in the first round. Therefore, the maximum number of people who leave the circle in the first round is not more than $\lfloor n/2 \rfloor$. On the other hand, if the coin tosses alternate between tails and heads so that $c_0, \ldots, c_{n-1} = T, H, T, H, T, \ldots$ then persons $x_1, x_3, x_5, x_7, \ldots, x_{2\lfloor n/2 \rfloor - 1}$ will all leave the circle, so the maximum number of people who can leave the circle in the first round is not less than $\lfloor n/2 \rfloor$.

2. Since the coin tosses are independent

$$\Pr(\text{“}x_i \text{ survives”}) = 1 - \Pr(\text{“}c_{i-1} = T \text{ and } c_i = H \text{ and } c_{i+1} = T\text{”}) = 1 - \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = 7/8 \ .$$

3. Person x_i survives the first r rounds if they survive round 1 and they survive round 2 and then they survive round 3,..., and then they survive round r. From the first question, we know that the number of people who survive up to the beginning of round r' is at least $n/2^{r'-1} > 3$ for $r' - 1 < \log_2(n/3)$. Therefore, if x_i survives to the beginning of Round r' then round r' proceeds under the same assumptions we used for the previous question. Therefore,

$$\Pr(\text{“}x_i \text{ survives round } r'\text{”} \mid \text{“}x_i \text{ survives rounds } 1, \ldots, r' - 1\text{”}) = 7/8 \ .$$

Therefore,

$$\begin{aligned}
\Pr(\text{“}x_i \text{ survives rounds } 1, \ldots, r\text{”}) &= \prod_{r'=1}^{r} \Pr(\text{“}x_i \text{ survives round } r'\text{”} \mid \text{“}x_i \text{ survives rounds } 1, \ldots, r' - 1\text{”}) \\
&= \prod_{r'=1}^{r} 7/8 = (7/8)^r \ .
\end{aligned}$$

If we let I_i be the indicator variable

$$I_i = \begin{cases} 1 & \text{if } x_i \text{ survives rounds } 1, \ldots, r \\ 0 & \text{otherwise} \end{cases}$$

then $\mathbf{E}(I_i) = \Pr(I_i = 1) = (7/8)^r$. Then the expected number of people who survive the first r rounds is

$$\mathbf{E}\left(\sum_{i=1}^{n} I_i\right) = \sum_{i=1}^{n} \mathbf{E}(I_i) = n(7/8)^r \ .$$

3 Sampling With Replacement

For this one, recall the sum that we used to analyze geometric random variables: For any $0 < p < 1$,

$$\sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} \cdot p = \frac{1}{p}$$

1. For each $k \in \mathbb{N}$, $X = k$ if and only if $\pi_1 = \pi_2 = \cdots = \pi_{k-1} = T$ and $\pi_k = H$. Since the coin tosses are independent:

$$\begin{aligned}
\Pr(X = k) &= \Pr(\pi_1 = \pi_2 = \cdot = \pi_{k-1} = T \text{ and } \pi_k = H) \\
&= \Pr(\pi_1 = T) \cdot \Pr(\pi_2 = T) \cdot \cdots \cdot \Pr(\pi_{k-1} = T) \cdot \Pr(\pi_k = H) \\
&= (2/n)^{k-1} \cdot (1 - 2/n) \ .
\end{aligned}$$

Now, using the definition of expected value, we get

$$\mathbf{E}(X) = \sum_{k=1}^{\infty} k \Pr(X = k) = \sum_{k=1}^{\infty} k \cdot (2/n)^{k-1} \cdot (1 - 2/n) = \frac{1}{1 - 2/n} = \frac{n}{n-2} \ .$$

(Or we can observe that X is a gemetric$(1 - 2/n)$ random variable to get the same result.)

2. To compute $\mathbf{E}(Y)$, we proceed exactly as above except reversing the roles of $1 - 2/n$ and $2/n$ to finish with

$$\mathbf{E}(Y) = \sum_{k=1}^{\infty} k \cdot (1 - 2/n)^{k-1} \cdot (2/n) = \frac{n}{2} \ .$$

4 Sampling without Replacement

1. Computing $\mathbf{E}(X)$ is not too difficult because X has only three possible values:

   (a) $X = 1$ and this happens when π_1 is a beer bottle. There are $n - 2$ choices for π_1 and $(n-1)!$ choices for the permutation $\pi_2, \ldots, \pi_n$ of the remaining $n - 1$ bottles. So,

   $$\Pr(X = 1) = \frac{(n-2) \cdot (n-1)!}{n!} = \frac{n-2}{n}$$

   (b) $X = 2$ and this happens when π_1 is a cider bottle and π_2 is a beer bottle. There are 2 choices for the cider bottle π_1, there are $n - 2$ choices for the beer bottle π_2, and then there are $(n-2)!$ choices for the permutation $\pi_3, \ldots, \pi_n$ of the remaining $n - 2$ bottles. So

   $$\Pr(X = 2) = \frac{2 \cdot (n-2) \cdot (n-2)!}{n!} = \frac{2(n-2)}{n(n-1)}$$

   (c) $X = 3$ and this happens when π_1 and π_2 are cider bottles and π_3 is a beer bottle. There are $2! = 2$ choices for the ordering of the cider bottles (either $\pi_1\pi_2 = c_1c_2$ or $\pi_1\pi_2 = c_2c_1$) and then there are $(n-2)!$ choices for the ordering $\pi_3, \ldots, \pi_n$ of the remaining $n - 2$ beer bottles. So,

   $$\Pr(X = 3) = \frac{2 \cdot (n-2)!}{n!} = \frac{2}{n(n-1)}$$

   Applying the definition of expected value, we get

   $$\begin{aligned} \mathbf{E}(X) &= \sum_{k \in \{1,2,3\}} k \cdot \Pr(X = k) \\ &= \frac{n-2}{n} + \frac{4(n-2)}{n(n-1)} + \frac{6}{n(n-1)} \\ &= \frac{n+1}{n-1} \ . \end{aligned}$$

2. To compute $\mathbf{E}(Y)$ we should figure out $\Pr(Y = k)$ for each $k \in \{1, \ldots, n\}$. Now, $X = k$ exactly when $\pi_1, \ldots, \pi_{k-1}$ are beer bottles and π_k is a cider bottle. We can count the number of such permutations using the Product Rule:

   (a) Select the beer bottles $\pi_1, \ldots, \pi_{k-1}$. There are $n - 2$ choices for π_1 and $n - 3$ choices for π_2,..., and $n - 2 - (k - 2) = n - k$ choices for π_{k-1}, for a total of $(n-2)!/(n-k-1)!$ ways to execute this step.

   (b) Select a cider bottle π_k from c_1 or c_2. There are two ways to execute this step.

(c) Select a permutation $\pi_{k+1}, \ldots, \pi_n$ of the remaining $n-k$ bottles. There are $(n-k)!$ ways to perform this step.

Therefore, there are

$$\begin{aligned}\frac{(n-2)!}{(n-k-1)!} \cdot 2 \cdot (n-k)! &= (n-2)(n-3)\cdots(n-k+1)(n-k) \cdot 2 \cdot (n-k)(n-k-1)\cdots 1 \\ &= (n-2)! \cdot 2 \cdot (n-k)\end{aligned}$$

permutations $\pi_1, \ldots, \pi_n$ for which $Y = k$. Therefore,

$$\Pr(Y = k) = \frac{(n-2)! \cdot 2 \cdot (n-k)}{n!} = \frac{2(n-k)}{n(n-1)}$$

Finally, we finish by applying the definition of expected value

$$\begin{aligned}\mathbf{E}(Y) &= \sum_{k=1}^{n} k \cdot \Pr(X = k) \\ &= \sum_{k=1}^{n} k \cdot \frac{2(n-k)}{n(n-1)} \\ &= \frac{1}{n(n-1)} \cdot \sum_{k=1}^{n} k \cdot 2(n-k) \\ &= \frac{1}{n(n-1)} \cdot \left(\sum_{k=1}^{n} 2kn - \sum_{k=1}^{n} 2k^2 \right) \\ &= \frac{1}{n(n-1)} \cdot \left(2n \sum_{k=1}^{n} k - \sum_{k=1}^{n} 2k^2 \right) \\ &= \frac{1}{n(n-1)} \cdot \left(n^2(n+1) - \sum_{k=1}^{n} 2k^2 \right) \\ &= \frac{1}{n(n-1)} \cdot \left(n^2(n+1) - \sum_{k=1}^{n} 2k^2 \right) \\ &= \frac{1}{n(n-1)} \cdot \left(n^2(n+1) - 2\sum_{k=1}^{n} k^2 \right) \\ &= \frac{1}{n(n-1)} \cdot \left(n^2(n+1) - \frac{n(n+1)(2n+1)}{3} \right) \\ &= \frac{1}{n(n-1)} \cdot n(n+1) \left(n - \frac{(2n+1)}{3} \right) \\ &= \frac{1}{n(n-1)} \cdot n(n+1) \left(\frac{n-1}{3} \right) \\ &= \frac{n+1}{3}\end{aligned}$$

Notice that, although this random variable Y looks a lot like the one in Question 3.1, its expected value is quite a bit different.

5 Doing (much) Better by Taking the Minimum

1. We are told that $\Pr(X \geq i) \leq a/i$, so

$$\mathbf{E}(X) = \sum_{i=1}^{n} i \cdot \Pr(X = i) = \sum_{i=1}^{n} \Pr(X \geq i) \leq \sum_{i=1}^{n} a/i = aH_n \ ,$$

   where $H_n = \sum_{i=1}^{n} 1/i$ is the n-th harmonic number.

2. Since X_1 and X_2 are independent,

$$\Pr(Z \geq i) = \Pr(X_1 \geq i \text{ and } X_2 \geq i) = \Pr(X_1 \geq i) \cdot \Pr(X_2 \geq i) \leq (a/i)^2 \ .$$

3. Following the same procedure we used for $\mathbf{E}(X)$.

$$\mathbf{E}(Z) = \sum_{i=1}^{n} \Pr(Z \geq i) \leq \sum_{i=1}^{n} (a/i)^2 = a^2 \sum_{i=1}^{n} 1/i^2 \ .$$

   Now we're stuck until we can say something about $\sum_{i=1}^{n} 1/i^2$.

4. Following the link provided to the Basel Problem explains that $\sum_{i=1}^{\infty} 1/i^2 = \pi^2/6$. We can use this by continuing from the previous derivation:

$$\mathbf{E}(Z) \leq a^2 \sum_{i=1}^{n} 1/i^2 \leq a^2 \sum_{i=1}^{\infty} 1/i^2 = \frac{(a\pi)^2}{6} \ .$$

Notice that, by taking the minimum of two samples we went from a random variable whose expected value was $H_n \approx \ln n$ to a random variable whose expected value is at most $(a\pi)^2/6$—a constant that doesn't depend on n. This idea of taking the best of 2 (or more) samples has useful applications in randomized algorithms.