CS 70 Fall 2018	Discrete Mathematics and Probability Theory Alistair Sinclair and Yun Song	DIS 11B

1 Rolling Dice

(a) If we roll a fair 6-sided die, what is the expected number of times we have to roll before we roll a 6? What is the variance?

(b) Suppose we have two independent, fair n-sided dice labeled Die 1 and Die 2. If we roll the two dice until the value on Die 1 is smaller than the value on Die 2, what is the expected number of times that we roll? What is the variance?

(c) Let $n = 6$, so we are back to fair 6-sided die. Suppose we roll Die 1 until a 6 comes up, and we roll Die 2 until a 6 comes up. Let X be a random variable representing the number of times Die 1 is rolled before getting a 6, and let Y be the corresponding random variable for Die 2. Compute $\mathbb{P}[\min(X,Y) = n]$ and $\mathbb{P}[X+Y = n]$, where n is an integer.

Solution:

(a) Since rolling a die until obtaining a 6 involves independent rolls with a constant probability of success per roll, the expected number of times we roll follows a geometric distribution.

This question seeks to review basic formulas for the geometric distribution. The probability of rolling a 6 is $1/6$. Recall that the expectation is the inverse of the probability, and that the variance is $(1-p)/p^2$. Thus the expectation is $1/(1/6) = 6$ rolls. Thus the variance is $(1-p)/p^2 = (1-1/6)/(1/6)^2 = 30$ rolls.

(b) If we roll the two dice, three outcomes are possible: the two dice show the same number, Die 1 is greater than Die 2, or Die 2 is greater than Die 1. The last two events occur with the same likelihood and the first event occurs with chance $n/n^2 = 1/n$, since there are n^2 possible rolls and n different numbers for which there could to be duplicates. Thus the number of ways that Die 1 is smaller than Die 2 on a given roll is $(n^2-n)/2$, so the probability that this occurs on a given roll is $(n^2-n)/(2n^2) = 1/2 - 1/(2n)$.

The expected number of times we roll is therefore geometrically distributed with

$$p = \frac{1}{2} - \frac{1}{2n}.$$

Plugging this into the formulas for expectation and variance yields the answer.

(c) To compute $\mathbb{P}[\min(X,Y) = n]$, we notice that $\min(X,Y) = n$ corresponds to $\{X = n \text{ and } Y \geq n\}$

or $\{X \geq n \text{ and } Y = n\}$. Using the inclusion-exclusion principle, we have:

$$\begin{aligned}
\mathbb{P}[\min(X,Y) = n] &= \mathbb{P}[\{X = n\} \cap \{Y \geq n\}] + \mathbb{P}[\{X \geq n\} \cap \{Y = n\}] - \mathbb{P}[\{X = n\} \cap \{Y = n\}] \\
&= \mathbb{P}[\{X = n\}] \cdot \mathbb{P}[\{Y \geq n\}] + \mathbb{P}[\{X \geq n\}] \cdot \mathbb{P}[\{Y = n\}] \\
&\quad - \mathbb{P}[\{X = n\}] \cdot \mathbb{P}[\{Y = n\}] \quad \text{(independence)}
\end{aligned}$$

To compute $\mathbb{P}[X \geq n]$ (which is also $\mathbb{P}[Y = n]$, by symmetry), we have:

$$\begin{aligned}
\mathbb{P}[X \geq n] &= \sum_{i \geq n} \mathbb{P}[X = i] \\
&= \sum_{i \geq n} \left[\left(\frac{5}{6}\right)^{i-1} \cdot \left(\frac{1}{6}\right) \right] \\
&= \frac{\left[\left(\frac{5}{6}\right)^{n-1} \cdot \left(\frac{1}{6}\right) \right]}{\left(1 - \frac{5}{6}\right)} \quad \text{(sum of infinite geometric series)} \\
&= \left(\frac{5}{6}\right)^{n-1}
\end{aligned}$$

We can now plug this value into our probability above.

$$\begin{aligned}
\mathbb{P}[\min(X,Y) = n] &= \mathbb{P}[\{X = n\}] \cdot \mathbb{P}[\{Y \geq n\}] + \mathbb{P}[\{X \geq n\}] \cdot \mathbb{P}[\{Y = n\}] \\
&\quad - \mathbb{P}[\{X = n\}] \cdot \mathbb{P}[\{Y = n\}] \\
&= \left(\frac{5}{6}\right)^{n-1} \cdot \left(\frac{1}{6}\right) \cdot \left(\frac{5}{6}\right)^{n-1} + \left(\frac{5}{6}\right)^{n-1} \cdot \left(\frac{5}{6}\right)^{n-1} \cdot \left(\frac{1}{6}\right) \\
&\quad - \left(\frac{5}{6}\right)^{n-1} \cdot \left(\frac{1}{6}\right) \cdot \left(\frac{5}{6}\right)^{n-1} \cdot \left(\frac{1}{6}\right) \\
&= 2 \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{2(n-1)} - \frac{1}{36} \cdot \left(\frac{5}{6}\right)^{2(n-1)} \\
&= \left(\frac{5}{6}\right)^{2(n-1)} \cdot \left(\frac{2}{6} - \frac{1}{36}\right) \\
&= \left(\frac{25}{36}\right)^{(n-1)} \cdot \frac{11}{36}
\end{aligned}$$

This looks suspiciously like a probability from a geometric distribution...surprising or not?

This brings us to an alternative approach for computing $\mathbb{P}[\min(X,Y) = n]$, we can define a new random variable $Z = \min(X,Y)$. $Z = n$ corresponds to the event that either Die 1 or Die 2 is a 6 at their n-th rolls, *and* neither Die 1 nor Die 2 is a 6 in any previous roll. Note that the events $\{\text{Die 1 or Die 2 are 6}\}$ and $\{\text{Die 1 and Die 2 are not 6}\}$ are complements, so we can correspond them to "success" and "failure" events. Thus, Z is also a geometric random variable, whose success probability is $\mathbb{P}[\text{Die 1 is 6, or Die 2 is 6}] = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}$. Using the distribution of a geometric random variable, we have $\mathbb{P}[Z = n] = (\frac{25}{36})^{n-1}(\frac{11}{36})$.

To compute $\mathbb{P}[X+Y]=n$, we notice that there are $(n-1)$ combinations of $X=i$ and $Y=n-i$ such that $X+Y=n$. Each combination of $X=i$ and $Y=n-i$ correspond to disjoint events. Therefore:

$$\begin{aligned}
\mathbb{P}[X+Y=n] &= \sum_{i=1}^{n-1} \mathbb{P}[\{X=i\}\cap\{Y=n-i\}] \\
&= \sum_{i=1}^{n-1} \mathbb{P}[\{X=i\}]\cdot\mathbb{P}[\{Y=n-i\}] \quad \text{(independence)} \\
&= \sum_{i=1}^{n-1} \left[\left(\frac{5}{6}\right)^{i-1}\cdot\left(\frac{1}{6}\right)\right]\left[\left(\frac{5}{6}\right)^{n-i-1}\cdot\left(\frac{1}{6}\right)\right] \\
&= \sum_{i=1}^{n-1} \left[\left(\frac{5}{6}\right)^{n-2}\cdot\left(\frac{1}{6}\right)^{2}\right] \\
&= \left(\frac{5}{6}\right)^{n-2}\cdot\left(\frac{1}{6}\right)^{2}\cdot\sum_{i=1}^{n-1} 1 \\
&= \left(\frac{5}{6}\right)^{n-2}\cdot\left(\frac{1}{6}\right)^{2}\cdot(n-1)
\end{aligned}$$

2 Trick or Treat

Shreyas and Jerry are trick or treating together, and are each trying to collect all n flavors of Laffy Taffy. At each house, they each receive a Laffy Taffy, chosen uniformly at random from all flavors. However, Shreyas will throw a tantrum if Jerry gets a flavor he doesn't, so they agree that if they receive different flavors, they'll both politely return the candy they got, and move on to the next house. What is the expected number of houses they need to visit until they get a full set of Laffy Taffy?

Solution:

Similar to the original coupon collector problem, we start by considering the expected number of houses that Shreyas and Jerry need to visit to get a new flavor, once they already have i unique ones. At each house, there is a $\frac{n-i}{n}$ chance that Jerry gets a new flavor. However, in order for her to keep the candy, she needs Shreyas to get the same flavor, which happens with probability $\frac{1}{n}$. Thus, the probability that Shreyas and Jerry add a new flavor to their collection once they already have i distinct ones is $\frac{n-i}{n^2}$. This tells us that the expected number of houses they need to visit to get the $(i+1)$st flavor is $\frac{n^2}{n-i}$. Hence, in order to get all n flavors, Shreyas and Jerry will have to visit on average

$$\begin{aligned}
\sum_{i=0}^{n-1} \frac{n^2}{n-i} &= n^2\cdot\sum_{i=0}^{n-1}\frac{1}{n-i} \\
&= n^2\cdot\sum_{j=1}^{n}\frac{1}{j}
\end{aligned}$$

where for the second equality, we did a variable substitution $j = n - i$ to simplify. As in the standard coupon collector problem, we know that the summation can be approximated by $\ln(n)$, so we have that in expectation, Shreyas and Jerry have to visit approximately $n^2 \ln(n)$ houses until they have all n flavors.

3 Unreliable Servers

In a single cluster of a Google competitor, there are a huge number of servers n, each with a uniform and independent probability of going down in a given day. On average, 4 servers go down in the cluster per day. Recall that as $n \to \infty$, a $\text{Binom}(n, \lambda/n)$ distribution will tend towards a $\text{Poisson}(\lambda)$ distribution.

(a) What is an appropriate distribution by which the number of servers that crash can be modeled?

(b) Compute the expected value and variance of the number of crashed servers for a certain cluster.

(c) Compute the probability that fewer than 3 servers crashed.

(d) Compute the probability at least 3 servers crashed.

Solution:

(a) Because each server goes down independently of the other servers, and with the same probability, the number of servers that crash on a given day follows a binomial distribution $\text{Binom}(n, p)$. Let n be the number of servers. Since on average, 4 servers crash per day, we have $p = \frac{4}{n}$. We are given that the number of servers in the cluster is large, so $n \gg p$ and we can model the number of servers that crash as a Poisson distribution with $\lambda = 4$.

(b) Recall that the expectation and variance are both equal to $\lambda = 4$.

(c) To compute the probability that fewer than 3 servers went down, we must add the probabilities that 0 servers go down, 1 server goes down, and the probability that 2 servers go down. The PMF of the Poisson distribution is

$$\mathbb{P}[X = i] = \frac{\lambda^i}{i!} e^{-\lambda}.$$

Thus

$$\mathbb{P}[X = 0 \text{ or } X = 1 \text{ or } X = 2] = e^{-4} + 4e^{-4} + \frac{4^2}{2} e^{-4} = e^{-4} + 4e^{-4} + 8e^{-4} = 13e^{-4}.$$

(d) $1 - \mathbb{P}[\text{fewer than 3 servers crashed}] = 1 - 13e^{-4}$.