CS 70 Discrete Mathematics and Probability Theory
Fall 2021 DIS 9B

1 Balls and Bins

Throw n balls into n labeled bins one at a time.

(a) What is the probability that the first bin is empty?

(b) What is the probability that the first k bins are empty?

(c) Let A be the event that at least k bins are empty. Notice that there are $m = \binom{n}{k}$ sets of k bins out of the total n bins. If we assume A_i is the event that the i^{th} set of k bins is empty. Then we can write A as the union of A_i's.

$$A = \bigcup_{i=1}^{m} A_i.$$

Write the union bound for the probability A.

(d) Use the union bound to give an upper bound on the probability A from part (c).

(e) What is the probability that the second bin is empty given that the first one is empty?

(f) Are the events that "the first bin is empty" and "the first two bins are empty" independent?

(g) Are the events that "the first bin is empty" and "the second bin is empty" independent?

Solution: Since the balls are thrown one at a time, there is an ordering, and so we are sampling with replacement where order matters rather than where it doesn't (which would correspond to each configuration in the stars and bars setup being equally likely).

(a) The probability that ball i does not land in the first bin is $\frac{n-1}{n}$. The probability that all of the balls do not land in the first bin is $\left(\frac{n-1}{n}\right)^n$.

(b) The probability that ball i does not land in the first k bins is $\frac{n-k}{n}$. The probability that all of the balls do not land in the first k bins is $\left(\frac{n-k}{n}\right)^n$.

(c) We use the union bound. Then

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{i=1}^{m} A_i\right) \leq \sum_{i=1}^{m} \mathbb{P}(A_i)$$

(d) We know the probability of the first k bins being empty from part (b), and this is true for any set of k bins, so

$$\mathbb{P}(A_i) = \left(\frac{n-k}{n}\right)^n.$$

Then,

$$\mathbb{P}(A) \leq m \cdot \left(\frac{n-k}{n}\right)^n = \binom{n}{k}\left(\frac{n-k}{n}\right)^n.$$

(e) Using Bayes' Rule:

$$\begin{aligned}\mathbb{P}[\text{2nd bin empty} \mid \text{1st bin empty}] &= \frac{\mathbb{P}[\text{2nd bin empty} \cap \text{1st bin empty}]}{\mathbb{P}[\text{1st bin empty}]} \\ &= \frac{(n-2)^n/n^n}{(n-1)^n/n^n} \\ &= \left(\frac{n-2}{n-1}\right)^n\end{aligned} \quad (1)$$

Alternate solution:
We know bin 1 is empty, so each ball that we throw can land in one of the remaining $n-1$ bins. We want the probability that bin 2 is empty, which means that each ball cannot land in bin 2 either, leaving $n-2$ bins. Thus for each ball, the probability that bin 2 is empty given that bin 1 is empty is $(n-2)/(n-1)$. For n total balls, this probability is $[(n-2)/(n-1)]^n$.

(f) They are dependent. Knowing the latter means the former happens with probability 1.

(g) In part (c) we calculated the probability that the second bin is empty given that the first bin is empty: $[(n-2)/(n-1)]^n$. The probability that the second bin is empty (without any prior information) is $[(n-1)/n]^n$. Since these probabilities are not equal, the events are dependent.

2 Head Count

Consider a coin with $\mathbb{P}(\text{Heads}) = 2/5$. Suppose you flip the coin 20 times, and define X to be the number of heads.

(a) Name the distribution of X and what its parameters are.

(b) What is $\mathbb{P}(X = 7)$?

(c) What is $\mathbb{P}(X \geq 1)$? Hint: You should be able to do this without a summation.

(d) What is $\mathbb{P}(12 \leq X \leq 14)$?

Solution:

(a) Since we have 20 independent trials, with each trial having a probability $2/5$ of success, $X \sim \text{Binomial}(20, 2/5)$.

(b)

$$\mathbb{P}(X = 7) = \binom{20}{7}\left(\frac{2}{5}\right)^7\left(\frac{3}{5}\right)^{13}.$$

(c)

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - \left(\frac{3}{5}\right)^{20}.$$

(d)

$$\begin{aligned}\mathbb{P}(12 \leq X \leq 14) &= \mathbb{P}(X = 12) + \mathbb{P}(X = 13) + \mathbb{P}(X = 14) \\ &= \binom{20}{12}\left(\frac{2}{5}\right)^{12}\left(\frac{3}{5}\right)^{8} + \binom{20}{13}\left(\frac{2}{5}\right)^{13}\left(\frac{3}{5}\right)^{7} + \binom{20}{14}\left(\frac{2}{5}\right)^{14}\left(\frac{3}{5}\right)^{6}.\end{aligned}$$

3 Exploring the Geometric Distribution

Suppose $X \sim \text{Geometric}(p)$ and $Y \sim \text{Geometric}(q)$ are independent. Find the distribution of $\min\{X, Y\}$ and justify your answer.

Solution:

X is the number of coins we flip until we see a heads from flipping a coin with bias p, and Y is the same as flipping a coin with bias q. Imagine we flip the bias p coin and the bias q coin at the same time. The min of the two random variables represents how many simultaneous flips occur before at least one head is seen.

The probability of not seeing a head at all on any given simultaenous flip is $(1-p)(1-q)$, so the probability that there will be a success on any particular trial is $p+q-pq$. Therefore, $\min\{X, Y\} \sim \text{Geometric}(p+q-pq)$.

We can also solve it algebraically. The probability that $\min\{X, Y\} = k$ for some positive integer k is the probability that the first $k-1$ coin flips for both X and Y were tails, then times the probability that we get heads on the k-th toss. Specifically,

$$((1-p)(1-q))^{k-1} \cdot (p+q-pq)$$

We recognize this as the formula for a geometric random variable with parameter $p+q-pq$.