1 Balls and Bins

Suppose you 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 $m$ subsets 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:

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

Use the union bound to give an upper bound on the probability $\mathbb{P}[A]$.

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

(e) Are the events that “the first bin is empty” and “the first two bins are empty” independent?

(f) 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

$$\mathbb{P}[A] = \mathbb{P}\left[\bigcup_{i=1}^{m} A_i\right] \leq \sum_{i=1}^{m} \mathbb{P}[A_i].$$

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,

$$\mathbb{P}[A] \leq m \cdot \left(\frac{n-k}{n}\right)^n = \binom{n}{k} \left(\frac{n-k}{n}\right)^n.$$
(d) Using Bayes’ Rule:
\[
P[\text{2nd bin empty} \mid \text{1st bin empty}] = \frac{P[\text{2nd bin empty} \cap \text{1st bin empty}]}{P[\text{1st bin empty}]} = \frac{(n-2)^n/n^n}{(n-1)^n/n^n} = \left(\frac{n-2}{n-1}\right)^n
\]

**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 \(\left(\frac{n-2}{n-1}\right)^n\). For \(n\) total balls, this probability is \(\left(\frac{n-2}{n-1}\right)^n\).

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

(f) In part (c) we calculated the probability that the second bin is empty given that the first bin is empty: \(\left(\frac{n-2}{n-1}\right)^n\). The probability that the second bin is empty (without any prior information) is \(\left(\frac{n-1}{n}\right)^n\). Since these probabilities are not equal, the events are dependent.

### 2 Head Count

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

(a) What is \(P[X = k]\), for some \(0 \leq k \leq 20\)?

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

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

(d) What is \(P[12 \leq X \leq 14]\)?

**Solution:**

(a) There are a total of \(\binom{20}{k}\) ways to select \(k\) coins to be heads. The probability that the selected \(k\) coins to be heads is \(\left(\frac{2}{5}\right)^k\), and the probability that the rest are tails is \(\left(\frac{3}{5}\right)^{20-k}\). Putting this together, we have
\[
P[X = k] = \begin{pmatrix} 20 \\ k \end{pmatrix} \left(\frac{2}{5}\right)^k \left(\frac{3}{5}\right)^{20-k}.
\]

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

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

(d) \[
P[12 \leq X \leq 14] = P[X = 12] + P[X = 13] + P[X = 14]
\]
\[
= \begin{pmatrix} 20 \\ 12 \end{pmatrix} \left(\frac{2}{5}\right)^{12} \left(\frac{3}{5}\right)^8 + \begin{pmatrix} 20 \\ 13 \end{pmatrix} \left(\frac{2}{5}\right)^{13} \left(\frac{3}{5}\right)^7 + \begin{pmatrix} 20 \\ 14 \end{pmatrix} \left(\frac{2}{5}\right)^{14} \left(\frac{3}{5}\right)^6.
\]
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:** Let $X$ be the number of coins we flip until we see a heads from flipping a coin with bias $p$, and let $Y$ similarly be the number of coins we flip until we see a heads from flipping a coin with bias $q$.

Imagine we flip the bias $p$ coin and the bias $q$ coin at the same time. The minimum 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 simultaneous flip is $(1-p)(1-q)$; this corresponds to a failure. This means that the probability that there will be a success on any particular trial is $1 - (1-p)(1-q) = p + q - pq$. Therefore, $\min(X, Y) \sim \text{Geometric}(p + q - pq)$.

We can also solve this 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, and we get heads on the $k$th toss (this can come from either $X$ or $Y$). Specifically, this occurs with probability

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

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

CS 70, Fall 2022, DIS 9B 3