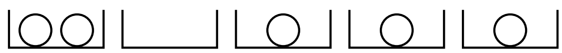
CS 70 Discrete Mathematics and Probability Theory
Summer 2026 Tate

# DIS 4D

## Balls and Bins Intro

Note 15

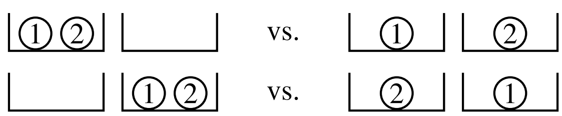
Recall the *balls and bins* (aka *stars and bars*) technique we used to count the number of ways we can sample with replacement, where order does not matter. Consider a process where we throw balls into bins, where each ball is independent and equally likely to land in any bin.



$$\star\star \mid \; \mid \star \mid \star \mid \star\star$$

Unfortunately, outcomes have different probabilities. For instance, with 2 balls and 2 bins, it's less likely for both balls to be in bin 1 than for one ball to be in each bin. We don't have a uniform probability space.


To work with a uniform probability space, we will *label* the balls. Order does matter now: throwing the first ball into bin 1 and the second ball into bin 2 (top right) is a different outcome from throwing the first ball into bin 2 and the second ball into bin 1 (bottom right).




## 1 Balls and Bins

Note 15

Suppose you throw $b$ balls into $n$ labeled bins one at a time, where $b \geq 1$ and $n \geq 2$.

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

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

(c) If $k$ bins were randomly selected, what is the probability that they are all empty?

(d) Let $A$ be the event that at least $k$ bins are empty. Let $m$ be the number of subsets of $k$ bins out of the total $n$ bins. If we assume $A_i$ is the event that the $i$th subset of $k$ bins is empty. Then we can write $A$ as the union of $A_i$'s:

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

Compute $m$ in terms of $n$ and $k$, and use the union bound to give an upper bound on $\mathbb{P}[A]$.

(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?

## 2 Monopoly Card Collector

Note 15

It's that time of the year again—Safeway is offering its Monopoly Card promotion. Each time you visit Safeway, you are given one of $n$ different Monopoly Cards with equal probability. You need to collect them all to redeem the grand prize.

(a) Suppose you have $k$ unique Monopoly Cards already, where $k \leq n$. What is the probability that you'll get a new unique Monopoly Card on your next visit to Safeway?

Now, suppose that you have visited Safeway $r$ times already.

(b) Let $X_i$ be the event that you are missing Monopoly Card $i$, where $i \in \{1, \ldots, n\}$. Find $\mathbb{P}[X_i]$.

(c) Let $T$ be the event that you are missing any Monopoly Card. Express $T$ in terms of the events $X_i$. *Hint*: Events are subsets of the probability space $\Omega$, so you should be thinking of set operations.

(d) Show that $ne^{-r/n}$ is an upper bound for $\mathbb{P}[T]$. That is, show that $\mathbb{P}[T] \leq ne^{-r/n}$. (One inequality you may find helpful is $1 - x \leq e^{-x}$. Is there another inequality that would be helpful?)

(e) Find the minimum number of Safeway visits $r$ such that part (d) guarantees that $\mathbb{P}[T]$ is at most $\frac{1}{n^2}$?

## 3 Planetary Party

Note 15

For this question, you may use a calculator.

(a) Suppose we are at party on a planet where every year is 2849 days. If 30 people attend this party, what is the exact probability that at least two people will share the same birthday? You may leave your answer as an unevaluated expression.

(b) From lecture, we know that given $n$ bins and $m$ balls, $\mathbb{P}[\text{no collision}] \approx \exp(-m^2/(2n))$. Using this, give an approximation for the probability in part (a).

(c) Using the approximation in part (b), what is the approximate minimum number of people that need to attend this party to ensure that the probability that any two people share a birthday is at least 0.5?

(d) Now suppose that 70 people attend this party. What the is probability that none of these 70 individuals have the same birthday? You can use the approximation you used in the previous parts.