

Probability and Computing

Chapter 5: Balls, Bins, and Random Graphs

Zeng Yuxiang

The Hong Kong University of Science and Technology

July 2, 2018

Table of Contents

- 1 Example: The Birthday Paradox
- 2 Balls into Bins
- 3 The Poisson Distribution
- 4 The Poisson Approximation
- 5 Application: Hashing
- 6 Random Graphs

Table of Contents

1 Example: The Birthday Paradox

2 Balls into Bins

3 The Poisson Distribution

4 The Poisson Approximation

5 Application: Hashing

6 Random Graphs

Example: The Birthday Paradox

Example

You notice that there are 30 people in the room. What is the probability that no two people in the room share the same birthday?

Analysis

Example: The Birthday Paradox

Example

You notice that there are 30 people in the room. What is the probability that no two people in the room share the same birthday?

We assume the birthday of each person is a random day from a 365-day year, chosen independently and uniformly at random for each person.

Analysis

Example: The Birthday Paradox

Example

You notice that there are 30 people in the room. What is the probability that no two people in the room share the same birthday?

We assume the birthday of each person is a random day from a 365-day year, chosen independently and uniformly at random for each person.

Analysis

First, there are 365^{30} possible configurations in total.

Example: The Birthday Paradox

Example

You notice that there are 30 people in the room. What is the probability that no two people in the room share the same birthday?

We assume the birthday of each person is a random day from a 365-day year, chosen independently and uniformly at random for each person.

Analysis

First, there are 365^{30} possible configurations in total.

Second, there are $\binom{365}{30}$ possible configurations to choose 30 different days from 365 days. These 30 days can be assigned to the people in any of the $30!$ possible orders.

Example: The Birthday Paradox

Example

You notice that there are 30 people in the room. What is the probability that no two people in the room share the same birthday?

We assume the birthday of each person is a random day from a 365-day year, chosen independently and uniformly at random for each person.

Analysis

First, there are 365^{30} possible configurations in total.

Second, there are $\binom{365}{30}$ possible configurations to choose 30 different days from 365 days. These 30 days can be assigned to the people in any of the $30!$ possible orders.

Finally, there are $\binom{365}{30} \cdot 30!$ configurations where no two people share the same birthday, out of the 365^{30} ways the birthdays could occur. Thus, the probability is $\frac{\binom{365}{30} \cdot 30!}{365^{30}} = \frac{\prod_{i=0}^{29} (365-i)}{365^{30}} \cdot 30! = \prod_{i=0}^{29} (1 - \frac{i}{365}) = \prod_{i=1}^{30} (1 - \frac{i-1}{365})$.

Example: The Birthday Paradox

Example

You notice that there are 30 people in the room. What is the probability that no two people in the room share the same birthday?

Analysis (Another Version)

Consider the probability by one person at a time. Specifically, the first person can choose any day.

Example: The Birthday Paradox

Example

You notice that there are 30 people in the room. What is the probability that no two people in the room share the same birthday?

Analysis (Another Version)

Consider the probability by one person at a time.

Specifically, the first person can choose any day.

The probability for the second person to choose a different day is

$(1 - \frac{1}{365})$.

Example: The Birthday Paradox

Example

You notice that there are 30 people in the room. What is the probability that no two people in the room share the same birthday?

Analysis (Another Version)

Consider the probability by one person at a time.

Specifically, the first person can choose any day.

The probability for the second person to choose a different day is

$$(1 - \frac{1}{365}).$$

The probability for the third person to choose a different day is $(1 - \frac{2}{365})$.

Similarly, the probability for the k -th person to choose a different day is

$$(1 - \frac{k-1}{365}).$$

Example: The Birthday Paradox

Example

You notice that there are 30 people in the room. **What is the probability that no two people in the room share the same birthday?**

Analysis (Another Version)

Consider the probability by one person at a time.

Specifically, the first person can choose any day.

The probability for the second person to choose a different day is

$$(1 - \frac{1}{365}).$$

The probability for the third person to choose a different day is $(1 - \frac{2}{365})$.

Similarly, the probability for the k -th person to choose a different day is

$$(1 - \frac{k-1}{365}).$$

Thus, the probability that 30 people all have different birthday is

$$(1 - \frac{1}{365})(1 - \frac{2}{365}) \cdots (1 - \frac{29}{365}) = \prod_{i=1}^{29} (1 - \frac{i}{365}) \approx 0.2937$$

Example: The Birthday Paradox

Example

You notice that there are 30 people in the room. What is the probability that no two people in the room share the same birthday?

A similar calculation shows that only 23 people need to be in the room before it is more likely (with probability higher than 50%) than not that two people share a birthday.

Example: The Birthday Paradox

Example

More generally, if there are m people and n possible birthdays, then what is the probability that all m have different birthdays?

Analysis

Example: The Birthday Paradox

Example

More generally, if there are m people and n possible birthdays, then what is the probability that all m have different birthdays?

Analysis

Inductively,

$$\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{m-1}{n}\right) = \prod_{i=1}^{m-1} \left(1 - \frac{i}{n}\right)$$

Example: The Birthday Paradox

Example

More generally, if there are m people and n possible birthdays, then what is the probability that all m have different birthdays?

Analysis

Inductively,

$$(1 - \frac{1}{n})(1 - \frac{2}{n}) \cdots (1 - \frac{m-1}{n}) = \prod_{i=1}^{m-1} (1 - \frac{i}{n})$$

Using that $1 - i/n \approx e^{-i/n}$ when i is small compared to n , we see that if m is small compared to n then

$$\begin{aligned} \prod_{i=1}^{m-1} (1 - \frac{i}{n}) &\approx \prod_{i=1}^{m-1} e^{-i/n} = \exp\{-\prod_{i=1}^{m-1} \frac{i}{n}\} \\ &= e^{-m(m-1)/2n} = e^{-m^2/2n} \end{aligned}$$

Example: The Birthday Paradox

Example

More generally, if there are m people and n possible birthdays, then what is the probability that all m have different birthdays?

Analysis

The probability that all m have different birthdays is $e^{-m^2/2n}$. Hence, the value for m at which the probability that m people all have different birthdays is $1/2$, which is approximately given by the equation

$$\frac{m^2}{2n} = \ln 2,$$

or $m = \sqrt{2n \ln 2}$.

Example: The Birthday Paradox

Example

More generally, if there are m people and n possible birthdays, then what is the probability that all m have different birthdays?

Analysis

The probability that all m have different birthdays is $e^{-m^2/2n}$. Hence, the value for m at which the probability that m people all have different birthdays is $1/2$, which is approximately given by the equation

$$\frac{m^2}{2n} = \ln 2,$$

or $m = \sqrt{2n \ln 2}$.

For the case $n = 365$, this approximation gives $m = 22.49$ to two decimal places.

Example: The Birthday Paradox

Example

More generally, if there are m people and n possible birthdays, then what is the probability that all m have different birthdays?

Loose Bound Analysis

let E_k be the event that the k th person's birthday does not match any of the birthdays of the first $k - 1$ people. Then the probability that the first k people fail to have distinct birthdays is

$$\Pr(\bar{E}_1 \cup \bar{E}_2 \cup \dots \cup \bar{E}_k) \leq \sum_{i=1}^k \bar{E}_i \leq \sum_{i=1}^k \frac{i-1}{n} = \frac{k(k-1)}{2n}$$

Example: The Birthday Paradox

Example

More generally, if there are m people and n possible birthdays, then what is the probability that all m have different birthdays?

Loose Bound Analysis

let E_k be the event that the k th person's birthday does not match any of the birthdays of the first $k - 1$ people. Then the probability that the first k people fail to have distinct birthdays is

$$\Pr(\bar{E}_1 \cup \bar{E}_2 \cup \dots \cup \bar{E}_k) \leq \sum_{i=1}^k \bar{E}_i \leq \sum_{i=1}^k \frac{i-1}{n} = \frac{k(k-1)}{2n}$$

If $k \leq \sqrt{n}$ for this probability is less than $1/2$, so with $\lfloor \sqrt{n} \rfloor$ people the probability is at least $1/2$ that all birthdays will be distinct.

Example: The Birthday Paradox

Example

More generally, if there are m people and n possible birthdays, then what is the probability that all m have different birthdays?

Loose Bound Analysis (Cont.)

Assume that the first $\lceil\sqrt{n}\rceil$ people all have distinct birthdays. Each person after has probability at least $\sqrt{n}/n = 1/\sqrt{n}$ of having the same birthday as one of these \sqrt{n} people. Thus, the probability that the next $\lceil\sqrt{n}\rceil$ people all have different birthdays from the first $\lceil\sqrt{n}\rceil$ is at most

$$\left(1 - \frac{1}{\sqrt{n}}\right)^{\lceil\sqrt{n}\rceil} < \frac{1}{e} < \frac{1}{2}.$$

Example: The Birthday Paradox

Example

More generally, if there are m people and n possible birthdays, then what is the probability that all m have different birthdays?

Loose Bound Analysis (Cont.)

Assume that the first $\lceil\sqrt{n}\rceil$ people all have distinct birthdays. Each person after has probability at least $\sqrt{n}/n = 1/\sqrt{n}$ of having the same birthday as one of these \sqrt{n} people. Thus, the probability that the next $\lceil\sqrt{n}\rceil$ people all have different birthdays from the first $\lceil\sqrt{n}\rceil$ is at most

$$\left(1 - \frac{1}{\sqrt{n}}\right)^{\lceil\sqrt{n}\rceil} < \frac{1}{e} < \frac{1}{2}.$$

Hence, once there are \sqrt{n} people, the probability is at most $1/e$ that all birthdays will be distinct.

Table of Contents

- 1 Example: The Birthday Paradox
- 2 Balls into Bins
- 3 The Poisson Distribution
- 4 The Poisson Approximation
- 5 Application: Hashing
- 6 Random Graphs

Definition: The Balls-and-Bins Model

The balls-and-bins model

We have m balls that are thrown into n bins, with the location of each ball chosen independently and uniformly at random from the n possibilities. What does the distribution of the balls in the bins look like?

Definition: The Balls-and-Bins Model

The balls-and-bins model

We have m balls that are thrown into n bins, with the location of each ball chosen independently and uniformly at random from the n possibilities.

What does the distribution of the balls in the bins look like?

- How many of the bins are empty?

Definition: The Balls-and-Bins Model

The balls-and-bins model

We have m balls that are thrown into n bins, with the location of each ball chosen independently and uniformly at random from the n possibilities.

What does the distribution of the balls in the bins look like?

- How many of the bins are empty?
- How many balls are in the fullest bin?

Definition: The Balls-and-Bins Model

The balls-and-bins model

We have m balls that are thrown into n bins, with the location of each ball chosen independently and uniformly at random from the n possibilities.

What does the distribution of the balls in the bins look like?

- How many of the bins are empty?
- How many balls are in the fullest bin?
- Whether there is a bin containing at least two balls? (Birthday Paradox)

Definition: The Balls-and-Bins Model

The balls-and-bins model

We have m balls that are thrown into n bins, with the location of each ball chosen independently and uniformly at random from the n possibilities. What does the distribution of the balls in the bins look like?

- How many of the bins are empty?
- How many balls are in the fullest bin?
- Whether there is a bin containing at least two balls? (**Birthday Paradox**)

Birthday Paradox

If m balls are randomly placed into n bins then, for some $m = \Omega(\sqrt{n})$, at least one of the bins is likely to have more than one ball in it.

Example: The Balls-and-Bins Model

Example

What is the maximum number of balls in a bin, or the maximum *load*.

Analysis

Example: The Balls-and-Bins Model

Example

What is the maximum number of balls in a bin, or the maximum *load*.

Analysis

Let us first consider the case $m = n$, so that the average *load* is 1. Of course the maximum *load* is m , but it is very unlikely that all m balls land in the same bin.

Example: The Balls-and-Bins Model

Example

What is the maximum number of balls in a bin, or the maximum *load*.

Analysis

Let us first consider the case $m = n$, so that the average *load* is 1. Of course the maximum *load* is m , but it is very unlikely that all m balls land in the same bin.

We seek an upper bound that holds with probability tending to 1 as n grows large.

Example: The Balls-and-Bins Model

Example

What is the maximum number of balls in a bin, or the maximum *load*.

Analysis

Let us first consider the case $m = n$, so that the average *load* is 1. Of course the maximum *load* is m , but it is very unlikely that all m balls land in the same bin.

We seek an upper bound that holds with probability tending to 1 as n grows large.

We can show that the maximum *load* is not more than $3 \ln n / \ln \ln n$ with probability at most $1/n$ for sufficiently large n via a direct calculation and a union bound.

Example: The Balls-and-Bins Model

Maximum Load

When n balls are thrown independently and uniformly at random into n bins, the probability that the maximum load is more than $3 \ln n / \ln \ln n$ is at most $1/n$ for n sufficiently large.

Example: The Balls-and-Bins Model

Maximum Load

When n balls are thrown independently and uniformly at random into n bins, the probability that the maximum load is more than $3 \ln n / \ln \ln n$ is at most $1/n$ for n sufficiently large.

Proof:

The probability that bin 1 receives at least M balls is at most

$$\binom{n}{M} \left(\frac{1}{n}\right)^M = \frac{\prod_{i=0}^{M-1} (n-i)}{M!} \cdot \frac{1}{n^M} \leq \frac{1}{M!} \leq \left(\frac{e}{M}\right)^M$$

The last inequality is because

$$\frac{k^k}{k!} < \sum_{i=0}^{\infty} \frac{k^i}{i!} = e^k \Rightarrow k! > (k/e)^k$$

Example: The Balls-and-Bins Model

Maximum Load

When n balls are thrown independently and uniformly at random into n bins, the probability that the maximum load is more than $3 \ln n / \ln \ln n$ is at most $1/n$ for n sufficiently large.

Proof:

The probability that bin 1 receives at least M balls is at most $(\frac{e}{M})^M$. Applying a **union bound** again allows us to find that, for $M \geq 3 \ln n / \ln \ln n$, the probability that any bin receives at least M balls is bounded above by

$$\begin{aligned} n \cdot \left(\frac{e}{M}\right)^M &\leq n \left(\frac{e \ln \ln n}{3 \ln n}\right)^{3 \ln n / \ln \ln n} \\ &< n \left(\frac{\ln \ln n}{\ln n}\right)^{3 \ln n / \ln \ln n} \\ &= e^{\ln n} (e^{\ln \ln \ln n - \ln \ln n})^{3 \ln n / \ln \ln n} \\ &= e^{-2 \ln n + 3(\ln n)(\ln \ln \ln n) / \ln \ln n} \leq \frac{1}{n} \end{aligned}$$

Application: Bucket Sort

Definition: Bucket Sort

Suppose that we have a set of $n = 2^m$ elements to be sorted and that each element is an integer chosen independently and uniformly at random from the range $[0, 2^k)$, where $k \geq m$. Using Bucket sort, we can sort the numbers in expected time $O(n)$. Here, expectation is over the choice of the random input, since Bucket sort is a completely deterministic algorithm.

Basic Idea: Bucket Sort

Bucket sort works in two stages. In the first stage, we place the elements into n buckets. The j -th bucket holds all elements whose first m binary digits correspond to the number j .

Application: Bucket Sort

Definition: Bucket Sort

Suppose that we have a set of $n = 2^m$ elements to be sorted and that each element is an integer chosen independently and uniformly at random from the range $[0, 2^k)$, where $k \geq m$. Using Bucket sort, we can sort the numbers in expected time $O(n)$. Here, expectation is over the choice of the random input, since Bucket sort is a completely deterministic algorithm.

Basic Idea: Bucket Sort

Bucket sort works in two stages. In the first stage, we place the elements into n buckets. The j -th bucket holds all elements whose first m binary digits correspond to the number j . For example, if $n = 2^{10}$, bucket 3 contains all elements whose first 10 binary digits are 0000000011. When $i < \mathcal{L}$ the elements of the i -th bucket all come before the elements in the \mathcal{L} -th bucket in the sorted order.

Application: Bucket Sort

Definition: Bucket Sort

Suppose that we have a set of $n = 2^m$ elements to be sorted and that each element is an integer chosen independently and uniformly at random from the range $[0, 2^k)$, where $k \geq m$. Using Bucket sort, we can sort the numbers in expected time $O(n)$. Here, expectation is over the choice of the random input, since Bucket sort is a completely deterministic algorithm.

Basic Idea: Bucket Sort

Bucket sort works in two stages. In the first stage, we place the elements into n buckets. The j -th bucket holds all elements whose first m binary digits correspond to the number j . Assuming that each element can be placed in the appropriate bucket in $O(1)$ time, this stage requires only $O(n)$ time. Because of the assumption that the elements to be sorted are chosen uniformly, the number of elements that land in a specific bucket follows a binomial distribution $B(n, 1/n)$.

Application: Bucket Sort

Basic Idea: Bucket Sort (Cont.)

In the second stage, each bucket is sorted using any standard quadratic time algorithm (such as Bubble sort or Insertion sort). Concatenating the sorted lists from each bucket in order gives us the sorted order for the elements. It remains to show that the expected time spent in the second stage is only $O(n)$.

Application: Bucket Sort

Basic Idea: Bucket Sort (Cont.)

In the second stage, each bucket is sorted using any standard quadratic time algorithm (such as Bubble sort or Insertion sort). Concatenating the sorted lists from each bucket in order gives us the sorted order for the elements. It remains to show that the expected time spent in the second stage is only $O(n)$.

The result relies on our assumption regarding the input distribution. Under the uniform distribution, Bucket sort falls naturally into the balls and bins model: the elements are balls, buckets are bins, and each ball falls uniformly at random into a bin.

Application: Bucket Sort

Basic Idea: Bucket Sort (Cont.)

Let X_j be the number of elements that land in the j -th bucket. The time to sort the j -th bucket is then at most $c(X_j)^2$ for some constant c . The expected time spent sorting in the second stage is at most

$$E\left[\sum_{j=1}^n c(X_j)^2\right] = c \sum_{j=1}^n E[X_j^2] = cnE[X_1^2].$$

Application: Bucket Sort

Basic Idea: Bucket Sort (Cont.)

Let X_j be the number of elements that land in the j -th bucket. The time to sort the j -th bucket is then at most $c(X_j)^2$ for some constant c . The expected time spent sorting in the second stage is at most

$$E\left[\sum_{j=1}^n c(X_j)^2\right] = c \sum_{j=1}^n E[X_j^2] = cnE[X_1^2].$$

Since X_1 is a binomial random variable $B(n, 1/n)$, we have

$$E[X_1^2] = n\frac{1}{n}\left(1 - \frac{1}{n}\right) + \left(n\frac{1}{n}\right)^2 = 2 - \frac{1}{n} < 2.$$

Hence the total expected time spent in the second stage is at most $2cn$, so Bucket sort runs in expected linear time.

Table of Contents

- 1 Example: The Birthday Paradox
- 2 Balls into Bins
- 3 The Poisson Distribution**
- 4 The Poisson Approximation
- 5 Application: Hashing
- 6 Random Graphs

Background: The Poisson Distribution

Example

We now consider the probability that **a given bin is empty** in the balls and bins model. For the first bin to be empty, it must be missed by all m balls. Since each ball hits the first bin with probability $1/n$, the probability the first bin remains empty is $(1 - 1/n)^m \approx e^{-m/n}$.

Background: The Poisson Distribution

Example

We now consider the probability that **a given bin is empty** in the balls and bins model. For the first bin to be empty, it must be missed by all m balls. Since each ball hits the first bin with probability $1/n$, the probability the first bin remains empty is $(1 - 1/n)^m \approx e^{-m/n}$. If X_j is a random variable that is 1 when the j -th bin is empty and 0 otherwise, then $E[X_j] = (1 - 1/n)^m$. Let X be a random variable that represents the number of empty bins. Then, by the linearity of expectations

$$E[X] = E\left[\sum_{j=1}^n X_j\right] = \sum_{j=1}^n E[X_j] = n\left(1 - \frac{1}{n}\right)^m = ne^{-m/n}$$

Background: The Poisson Distribution

Example

We now consider the probability that **a given bin is empty** in the balls and bins model. For the first bin to be empty, it must be missed by all m balls. Since each ball hits the first bin with probability $1/n$, the probability the first bin remains empty is $(1 - 1/n)^m \approx e^{-m/n}$. If X_j is a random variable that is 1 when the j -th bin is empty and 0 otherwise, then $E[X_j] = (1 - 1/n)^m$. Let X be a random variable that represents the number of empty bins. Then, by the linearity of expectations

$$E[X] = E\left[\sum_{j=1}^n X_j\right] = \sum_{j=1}^n E[X_j] = n\left(1 - \frac{1}{n}\right)^m = ne^{-m/n}$$

Thus, the expected fraction of empty bins is approximately $e^{-m/n}$.

Background: The Poisson Distribution

Example

We can generalize the preceding argument to find the expected fraction of bins with r balls for any constant r . The probability that a given bin has r balls is

$$\binom{m}{r} \left(\frac{1}{n}\right)^r \left(1 - \frac{1}{n}\right)^{m-r} = \frac{1}{r!} \frac{m(m-1) \cdots (m-r+1)}{n^r} \left(1 - \frac{1}{n}\right)^{m-r}$$

Background: The Poisson Distribution

Example

We can generalize the preceding argument to find the expected fraction of bins with r balls for any constant r . The probability that a given bin has r balls is

$$\binom{m}{r} \left(\frac{1}{n}\right)^r \left(1 - \frac{1}{n}\right)^{m-r} = \frac{1}{r!} \frac{m(m-1) \cdots (m-r+1)}{n^r} \left(1 - \frac{1}{n}\right)^{m-r}$$

When m and n are large compared to r , the second factor on the right-hand side is approximately $(m/n)^r$, and the third factor is approximately $e^{-m/n}$.

Background: The Poisson Distribution

Example

We can generalize the preceding argument to find the expected fraction of bins with r balls for any constant r . The probability that a given bin has r balls is

$$\binom{m}{r} \left(\frac{1}{n}\right)^r \left(1 - \frac{1}{n}\right)^{m-r} = \frac{1}{r!} \frac{m(m-1) \cdots (m-r+1)}{n^r} \left(1 - \frac{1}{n}\right)^{m-r}$$

When m and n are large compared to r , the second factor on the right-hand side is approximately $(m/n)^r$, and the third factor is approximately $e^{-m/n}$.

Hence, the probability p_r that a given bin has r balls is approximately $p_r = \frac{e^{-m/n} (m/n)^r}{r!}$ and the expected number of bins with exactly r balls is approximately np_r .

Definition: Poisson Random Variable

Discrete Poisson Random Variable

A discrete **Poisson** random variable X with parameter μ is given by the following probability distribution on $j = 0, 1, 2, \dots$ $\Pr(X = j) = \frac{e^{-\mu} \mu^j}{j!}$.

Verification

Let us verify that the definition gives a proper distribution in that the probabilities sum to 1:

$$\begin{aligned}\sum_{j=0}^{\infty} \Pr(X = j) &= \sum_{j=0}^{\infty} \frac{e^{-\mu} \mu^j}{j!} \\ &= e^{-\mu} \sum_{j=0}^{\infty} \frac{\mu^j}{j!} \\ &= 1\end{aligned}$$

where we use the Taylor expansion $e^x = \sum_{j=0}^{\infty} (x^j/j!)$.

Definition: Poisson Random Variable

Discrete Poisson Random Variable

A discrete **Poisson** random variable X with parameter μ is given by the following probability distribution on $j = 0, 1, 2, \dots$ $\Pr(X = j) = \frac{e^{-\mu} \mu^j}{j!}$.

Expectation

Next we show that the expectation of this random variable is μ :

$$\begin{aligned} E[X] &= \sum_{j=0}^{\infty} j \Pr(X = j) = \sum_{j=1}^{\infty} j \frac{e^{-\mu} \mu^j}{j!} \\ &= \mu \sum_{j=1}^{\infty} \frac{e^{-\mu} \mu^{j-1}}{(j-1)!} = \mu \sum_{j=0}^{\infty} \frac{e^{-\mu} \mu^j}{j!} \\ &= \mu \end{aligned}$$

Definition: Poisson Random Variable

Discrete Poisson Random Variable

A discrete **Poisson** random variable X with parameter μ is given by the following probability distribution on $j = 0, 1, 2, \dots$ $\Pr(X = j) = \frac{e^{-\mu} \mu^j}{j!}$.

Expectation

Next we show that the expectation of this random variable is μ :

$$\begin{aligned} E[X] &= \sum_{j=0}^{\infty} j \Pr(X = j) = \sum_{j=1}^{\infty} j \frac{e^{-\mu} \mu^j}{j!} \\ &= \mu \sum_{j=1}^{\infty} \frac{e^{-\mu} \mu^{j-1}}{(j-1)!} = \mu \sum_{j=0}^{\infty} \frac{e^{-\mu} \mu^j}{j!} \\ &= \mu \end{aligned}$$

where the distribution of the number of balls in a bin is approximately **Poisson** with $\mu = m/n$. It is exactly the average number of balls per bin.

Multiple Poisson Random Variable

Multiple Poisson Random Variable

The sum of a finite number of independent Poisson random variables is a Poisson random variable.

Proof:

Here, we consider independent two Poisson random variables X and Y with means μ_1 and μ_2 .

$$\begin{aligned}\Pr(X + Y = j) &= \sum_{k=0}^j \Pr((X = k) \cup (Y = j - k)) \\&= \sum_{k=0}^j \frac{e^{-\mu_1} \mu_1^k}{k!} \frac{e^{-\mu_2} \mu_2^{(j-k)}}{(j-k)!} = \frac{e^{-(\mu_1 + \mu_2)}}{j!} \sum_{k=0}^j \frac{j!}{k!(j-k)!} \mu_1^k \mu_2^{(j-k)} \\&= \frac{e^{-(\mu_1 + \mu_2)}}{j!} \sum_{k=0}^j \binom{j}{k} \mu_1^k \mu_2^{(j-k)} = \frac{e^{-(\mu_1 + \mu_2)} (\mu_1 + \mu_2)^j}{j!}\end{aligned}$$

Table of Contents

- 1 Example: The Birthday Paradox
- 2 Balls into Bins
- 3 The Poisson Distribution
- 4 The Poisson Approximation**
- 5 Application: Hashing
- 6 Random Graphs

Table of Contents

- 1 Example: The Birthday Paradox
- 2 Balls into Bins
- 3 The Poisson Distribution
- 4 The Poisson Approximation
- 5 Application: Hashing**
- 6 Random Graphs

Table of Contents

- 1 Example: The Birthday Paradox
- 2 Balls into Bins
- 3 The Poisson Distribution
- 4 The Poisson Approximation
- 5 Application: Hashing
- 6 Random Graphs**