MATH 20: PROBABILITY

Important Distributions

Xingru Chen
xingru.chen.gr@dartmouth.edu
Important Distributions

Discrete Uniform Distribution
\[ m(\omega) = \frac{1}{n} \]

Binomial Distribution
\[ b(n, p, k) = \binom{n}{k} p^k q^{n-k} \]

Geometric Distribution
\[ P(T = n) = q^{n-1} p \]

Hypergeometric Distribution
\[ h(N, k, n, x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} \]

Negative Binomial Distribution
\[ u(x, k, p) = \binom{x-1}{k-1} p^k q^{x-k} \]

Poisson Distribution
\[ P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \]
Discrete Uniform Distribution

- In the sample space $S$, $m(\omega) = \frac{1}{n}$ for all $\omega \in S$. 

<table>
<thead>
<tr>
<th>Discrete Uniform Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m(\omega) = \frac{1}{n}$</td>
</tr>
</tbody>
</table>
Discrete Uniform Distribution

Toss a coin

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>head or tail</td>
<td></td>
</tr>
</tbody>
</table>

\[ m(\omega) = \frac{1}{2} \]
Discrete Uniform Distribution

Roll a dice

1, 2, 3, 4, 5, or 6

\[ m(\omega) = \frac{1}{6} \]
Draw a poker card

(10 + 3) × 4

\[ m(\omega) = \frac{1}{52} \]

Draw a poker card: heart

spade, diamond, club, heart

\[ m(\omega) = \frac{1}{4} \]

Draw a poker card: Q

1, 2, 3, ..., 10, J, Q, K

\[ m(\omega) = \frac{1}{13} \]
Pick a date
1, 2, ..., 365

\[ m(\omega) = \frac{1}{365} \]

Pick a month
1, 2, ..., 12

\[ m(\omega) = \frac{1}{12} \]

Pick a season
spring, summer, fall, winter

\[ m(\omega) = \frac{1}{4} \]
- Suppose you're on a game show, and you're given the choice of three doors: behind one door is a car, behind the other two, goats.

<table>
<thead>
<tr>
<th>Have a baby</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>boy or girl</td>
<td></td>
</tr>
</tbody>
</table>

\[ m(\omega) = \frac{1}{2} \]
Binomial Distribution

- Let $n$ be a positive integer and let $p$ be a real number between 0 and 1.
- Let $B$ be the random variable which counts the number of successes in a Bernoulli trials process with parameters $n$ and $p$.
- Then the distribution $b(n, p, k)$ of $B$ is called the binomial distribution.

\[
b(n, p, k) = \binom{n}{k} p^k q^{n-k}\]
Bernoulli Trials

A Bernoulli trials process is a sequence of \( n \) chance experiments such that

- Each experiment has two possible outcomes, which we may call **success** and **failure**.
- The probability \( p \) of success on each experiment is the same for each experiment, and this probability is not affected by any knowledge of previous outcomes. The probability \( q \) of failure is given by \( q = 1 - p \).
TOSS A COIN

Toss a coin 5 times. What is the probability that there are 2 flips that land heads?
Suppose you're on a game show, and you're given the choice of three doors: behind one door is a car; behind the others, goats.

Toss a coin 5 times. What is the probability that there are 2 flips that land heads?

<table>
<thead>
<tr>
<th>Toss a coin</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 5, p = \frac{1}{2}, k = 2$</td>
</tr>
</tbody>
</table>

$$b(n, p, k) = b\left(5, \frac{1}{2}, 2\right) = \binom{5}{2} \left(\frac{1}{2}\right)^5$$
ROLL A DICE

Roll a dice 10 times. What is the probability that 6 is obtained twice?
Suppose you're on a game show, and you're given the choice of three doors: behind one door is a car; behind the others, goats.

Roll a dice 10 times. What is the probability that 6 is obtained twice?

<table>
<thead>
<tr>
<th>Roll a dice</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n = 10, p = \frac{1}{6}, k = 2 )</td>
</tr>
</tbody>
</table>

\[
b(n, p, k) = b\left(10, \frac{1}{6}, 2\right) = \binom{10}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^8
\]
Suppose you're on a game show, and you're given the choice of three doors: behind one door is a car; behind the others, goats.

Monte Carlo Casino

In a game of roulette at the Monte Carlo Casino on August 18, 1913, the ball fell in black 26 times in a row. This was an extremely uncommon occurrence: the probability of a sequence of either red or black occurring 26 times in a row is around 1 in 66.6 million, assuming the mechanism is unbiased.

Gamblers lost millions of francs betting against black, reasoning incorrectly that the streak was causing an imbalance in the randomness of the wheel, and that it had to be followed by a long streak of red.
Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car; behind the others, goats.  

In a game of roulette at the Monte Carlo Casino on August 18, 1913, the ball fell in black 26 times in a row.

Monte Carlo Casino

\[
\begin{array}{|c|c|}
\hline
n & 26, \ p = \frac{1}{2}, \ k = 26 \\
\hline
b(n, p, k) = b\left(26, \frac{1}{2}, 26\right) = \binom{26}{26} \left(\frac{1}{2}\right)^{26} \\
& = \left(\frac{1}{2}\right)^{26} \approx 1.5 \times 10^{-8} \\
\hline
\end{array}
\]
Suppose you're on a game show, and you're given the choice of three doors: behind one door is a car; behind the others, goats.

**Binomial Distribution**

\[ b(n, p, k) = \binom{n}{k} p^k q^{n-k} \]
Suppose you’re on a game show, and you’re given the choice of three doors: behind one door is a car; behind the others, goats.

Binomial Distribution

\[ b(n, p, k) = \binom{n}{k} p^k q^{n-k} \]
Colorful Balls in a Jar

Experiment
Blindly pick up 3 balls.

Random variable
Number of brown balls.

Distribution
Binomial distribution or not?
TOSS A COIN

Toss a coin until first head shows up.
Infinite Discrete Sample Space

First Head

- The experiment is to repeatedly toss a coin until first head shows up.
- Possible outcomes are sequences of $T$ that, if finite, end with a single $H$, and an infinite sequence of $T$:
  \[ \Omega = \{H, TH, TTH, TTTTH, TTTTTTH, \ldots \} \]
- One random variable is defined most naturally as the length of an outcome.
- It draws values from the set of whole numbers augmented by the symbol of infinity:
  \[ \{1, 2, 3, 4, 5, \ldots, \infty \} \]
Geometric Distribution

- Consider a Bernoulli trials process continued for an infinite number of trials.
- For example, a coin tossed an infinite sequence of times.
- We can determine the distribution for any random variable $X$ relating to the experiment provided $P(X = a)$ can be computed in terms of a finite number of trials.
- For example, let $T$ be the number of trials up to and including the first success.

<table>
<thead>
<tr>
<th>Geometric Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(T = n) = q^{n-1}p$</td>
</tr>
</tbody>
</table>

![Diagram showing geometric distribution example]
Geometric Distribution

Toss a coin

| first head |

\[
P(T = n) = q^{n-1}p = \left(\frac{1}{2}\right)^{n-1}\frac{1}{2} = \left(\frac{1}{2}\right)^n
\]
Geometric Distribution

Roll a dice

**first 6**

$$P(T = n) = q^{n-1}p = \left(\frac{5}{6}\right)^{n-1} \frac{1}{6}$$

Roll a dice

**first even number**

$$P(T = n) = q^{n-1}p = \left(\frac{1}{2}\right)^{n-1} \frac{1}{2}$$

$$= \left(\frac{1}{2}\right)^n$$
An experiment has *two* possible outcomes: success and failure.

The *same* experiment is repeated until the *first* success.

Any two experiments are *independent*.

**Geometric Distribution**

\[ P(T = n) = q^{n-1}p \]
Geometric Distribution

\[ P(T = n) = q^{n-1}p \]
Colorful Balls in a Jar

Experiment
Blindly pick up balls until get a brown one.

Random variable
Number of balls.

Distribution
Geometric distribution or not?
Geometric Distribution

- Consider a Bernoulli trials process continued for an infinite number of trials.
- Let $T$ be the number of trials up to and including the first success.

Geometric Distribution

\[ P(T = n) = q^{n-1}p \]

Geometric Distribution

\[ P(T > k) = \sum_{j=k+1}^{+\infty} q^{j-1}p = q^k p (1 + q + q^2 + \cdots) = q^k \]
What is the formula for memoryless in probability theory?
Exponential Distribution

The amount of time we have to wait for an occurrence does not depend on how long we have already waited. The memoryless property says that knowledge of what has occurred in the past has no effect on future probabilities.

**Memoryless Property**

\[ P(X > r + s | X > r) = P(X > s) \]
Geometric Distribution

Geometric Distribution

\[ P(T > k) = \sum_{j=k+1}^{+\infty} q^{j-1}p = q^k p (1 + q + q^2 + \cdots) = q^k \]

\[ P(T > r + s | T > r) = \frac{P(T > r + s)}{P(T > r)} = \frac{q^{r+s}}{q^r} = q^s = P(T > s) \]

Memoryless Property

\[ P(T > r + s | T > r) = P(T > s) \]
Same independent trial 1

Trial 2

two

same

$k\text{th}$

independent

XC 2020
Negative Binomial Distribution

- Suppose we are given a coin which has probability $p$ of coming up heads when it is tossed.
- We fix a positive integer $k$, and toss the coin until the $k$th head appears.
- We let $X$ represent the number of tosses.
- When $k = 1$, $X$ is ...

Geometric Distribution

$$P(T = n) = q^{n-1}p$$
We let $X$ represent the number of tosses. For a general $k$, we now calculate the probability distribution of $X$.

### Negative Binomial Distribution

- number of heads: $k = 3$
- number of tosses: $x = 5$
- distribution: $u(x, k, p)$

---

**Examples:**

1. **1 2 3 4 5**
   - ✓ ✓ ✗ ✗ ✓
   - distribution: $u(5, 3, p)$

2. **1 2 3 4 5**
   - ✗ ✓ ✓ ✗ ✓
   - distribution: $u(5, 3, p)$

3. **1 2 3 4 5**
   - ✗ ✗ ✓ ✓ ✓
   - distribution: $u(5, 3, p)$

XC 2020
If $X = x$, then it must be true that there were exactly $k - 1$ heads thrown in the first $x - 1$ tosses, and a head must have been thrown on the $x$th toss.
### Negative Binomial Distribution

- number of heads: \( k \)
- number of tosses: \( x \)
- distribution: \( u(x, k, p) \)

There are

\[
\binom{x - 1}{k - 1}
\]

sequences of length \( x \) with these properties, and each of them is assigned the same probability, namely

\[ p^k q^{x-k}. \]
Negative Binomial Distribution

- number of heads: \( k \)
- number of tosses: \( x \)
- distribution: \( u(x, k, p) \)

There are
\[
\binom{x - 1}{k - 1}
\]
sequences of length \( x \) with these properties, and each of them is assigned the same probability, namely
\[
p^k q^{x-k}.
\]
# Negative Binomial Distribution

Roll a dice until the 6\(^{th}\) 6

- probability of success: \( p = \frac{1}{6} \)
- number of successes: \( k = 6 \)
- number of trials: \( x \)
- distribution: \( u(x, k, p) \)

\[
u(x, k, p) = \binom{x - 1}{k - 1} p^k q^{x-k}
\]

\[
= \binom{x - 1}{5} \left(\frac{1}{6}\right)^6 \left(\frac{5}{6}\right)^{x-6}
\]
# Negative Binomial Distribution

**Roll a dice**

until the 2\textsuperscript{nd} even number

- probability of success: \( p = \frac{1}{2} \)
- number of successes: \( k = 2 \)
- number of trials: \( x \)
- distribution: \( u(x, k, p) \)

\[
u(x, k, p) = \binom{x - 1}{k - 1} p^k q^{x-k}
\]

\[
= \binom{x - 1}{1} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{x-2} = \frac{x - 1}{2^x}
\]
Colorful Balls in a Jar

Experiment
Blindly pick up balls until get the 3rd brown one.

Random variable
Number of balls.

Distribution
Negative binomial distribution or not?
When it comes to time…

The Poisson distribution can be viewed as arising from the binomial distribution or from the exponential density.

**Binomial distribution**

\[ b(n, p, k) = \binom{n}{k} p^k q^{n-k} \]

**Poisson Distribution**

\[ P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \]

**Exponential Distribution**

\[ f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \]
Suppose that we have a situation in which a certain kind of occurrence happens at random over a period of time.

For example, the occurrences that we are interested in might be incoming telephone calls to a police station in a large city.

We want to model this situation so that we can consider the probabilities of events such as more than 3 phone calls occurring in a 12-hour time interval.
We want to model this situation so that we can consider the probabilities of events such as more than 3 phone calls occurring in a 12-hour time interval.

- We assume that the average rate, i.e., the average number of occurrences per hour, is a constant. This rate we will denote by $\lambda$.
- On the average, there are $\lambda t$ occurrences in a time interval of length $t$. 
Number of occurrences

Number of telephone calls

- average rate: $\lambda$
- time: $t$

$k = \lambda t$
Bernoulli trials

We can use the binomial distribution to model this situation.

We imagine that a given time interval is broken up into $n$ subintervals of equal length.

If the subintervals are sufficiently short, we can assume that two or more occurrences happen in one subinterval with a probability which is negligible in comparison with the probability of at most one occurrence.

Thus, in each subinterval, we are assuming that there is either 0 or 1 occurrence.
The sequence of subintervals can be thought of as a sequence of Bernoulli trials, with a success corresponding to an occurrence in the subinterval. If this time interval is divided into $n$ subintervals, then we would expect, using the Bernoulli trials interpretation, that there should be $np$ occurrences.
Bernoulli trials

Number of telephone calls

- probability of success: $p$
- number of trials: $n$

$$k = np$$
Number of occurrences

Number of telephone calls

- average rate: \( \lambda \)
- time: \( t \)

\[ k = \lambda t \]

Number of telephone calls

- probability of success: \( p \)
- number of trials: \( n \)

\[ k = np \]

\[ np = \lambda t \]

\[ p = \frac{\lambda t}{n} \]
Number of occurrences

### Binomial Distribution

\[
b(n, p, k) = \binom{n}{k} p^k q^{n-k}
\]

\[
P(X = 0) = b(n, p, 0) = (1 - p)^n
\]

\[
\frac{P(X = k)}{P(X = k - 1)} = \frac{b(n, p, k)}{b(n, p, k - 1)} = \frac{\binom{n}{k} p^k q^{n-k}}{\binom{n}{k-1} p^{k-1} q^{n-k+1}} = \frac{(n - k + 1)p}{kq}
\]
Number of occurrences

\[ p = \frac{\lambda t}{n}, \quad t = 1, \quad n \to \infty \]

### Binomial Distribution

\[ b(n, p, k) = \binom{n}{k} p^k q^{n-k} \]

\[ P(X = 0) = (1 - p)^n = \left(1 - \frac{\lambda}{n}\right)^n \]

\[ \frac{P(X = k)}{P(X = k - 1)} = \frac{(n - k + 1)p}{kq} = \frac{\lambda - (k - 1)p}{kq} \]
Number of occurrences

\[ p = \frac{\lambda t}{n}, \quad t = 1, \quad n \to \infty, \quad p \to 0 \]

Binomial Distribution

\[
\binom{n}{k} p^k q^{n-k}
\]

\[
P(X = 0) = (1 - p)^n = (1 - \frac{\lambda}{n})^n \approx e^{-\lambda}
\]

\[
\frac{P(X = k)}{P(X = k - 1)} = \frac{(n - k + 1)p}{kq} = \frac{\lambda - (k - 1)p}{kq} \approx \frac{\lambda}{k}
\]

\[
P(X = 1) \approx \lambda e^{-\lambda}
\]

\[
P(X = 2) \approx \frac{\lambda^2}{2!} e^{-\lambda}
\]
### Number of occurrences

#### Poisson Distribution

- **Circuit**

- **Diagram**

- **Formula**

\[
p = \frac{\lambda t}{n}, \quad t = 1, \ n \to \infty, \ p \to 0
\]

- **Table**

<table>
<thead>
<tr>
<th>Occurrence</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( P(X = 0) = e^{-\lambda} )</td>
</tr>
<tr>
<td>( k )</td>
<td>( \frac{P(X = k)}{P(X = k - 1)} = \frac{\lambda}{k} )</td>
</tr>
<tr>
<td>( k )</td>
<td>( P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} )</td>
</tr>
</tbody>
</table>
### Poisson Distribution

The Poisson Distribution is given by:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

#### Examples

<table>
<thead>
<tr>
<th>Value</th>
<th>Formula</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$e^{-\lambda}$</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td>at least one</td>
</tr>
<tr>
<td></td>
<td></td>
<td>....</td>
</tr>
<tr>
<td>1</td>
<td>$\lambda e^{-\lambda}$</td>
<td>one</td>
</tr>
<tr>
<td></td>
<td></td>
<td>less than two</td>
</tr>
<tr>
<td></td>
<td></td>
<td>at least two</td>
</tr>
<tr>
<td></td>
<td></td>
<td>...</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{\lambda^2}{2} e^{-\lambda}$</td>
<td>two</td>
</tr>
<tr>
<td></td>
<td></td>
<td>less than three</td>
</tr>
<tr>
<td></td>
<td></td>
<td>at least three</td>
</tr>
<tr>
<td></td>
<td></td>
<td>...</td>
</tr>
</tbody>
</table>
Binomial Distribution and Poisson Distribution

**Binomial Distribution**

\[ b(n, p, k) = \binom{n}{k} p^k q^{n-k} \]

**Poisson Distribution**

\[ P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \]

Graph showing the comparison of Binomial and Poisson distributions with parameters \( n = 10, p = 0.2, \lambda = 2.0 \).
$p = \frac{\lambda t}{n}$, $t = 1$, $n \to \infty$, $p \to 0$
Which one to use?

Binomial distribution

\[ b(n, p, k) = \binom{n}{k} p^k q^{n-k} \]

Poisson Distribution

\[ P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \]

two parameters

\[ n \quad p \]

one parameter

\[ \lambda \]
Example

Real Estate

- The average number of homes sold by the Acme Realty company is 2 homes per day. What is the probability that exactly 3 homes will be sold tomorrow?
Example

Poisson Distribution

\[ P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \]

- 2 homes per day: \( \lambda = 2 \)
- 3 homes: \( k = 3 \)

\[ P(X = 3) = \frac{2^3}{3!} e^{-2} = \frac{4}{3e^2} \]
Which one to use?

Binomial distribution

\[ b(n, p, k) = \binom{n}{k} p^k q^{n-k} \]

Poisson Distribution

\[ P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \]

For

\[ n < +\infty \]

Use the Binomial distribution.

For

\[ n \rightarrow +\infty \]

Use the Poisson distribution.
Which one to use?

Binomial distribution

\[ b(n, p, k) = \binom{n}{k} p^k q^{n-k} \]

Poisson Distribution

\[ P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \]

- \( n = 2, 5, \ldots \)
- \( n = 50, 100, \ldots \)
- \( np = \lambda \)
Nuclear Power Plant

Assume that the probability that there is a significant accident in a nuclear power plant during one year’s time is .001. If a country has 100 nuclear plants, estimate the probability that there is at least one such accident during a given year.
Nuclear Power Plant

Parameters (binomial)

- number of power plants $n: 100$
- probability of an accident $p: 0.001$

Parameter (Poisson)

$$\lambda = np = 100 \times 0.001 = 0.1$$
Nuclear Power Plant

Event

at least one such accident during a given year

Random variable

number of accidents

Probability

\[ P(X \geq 1) = 1 - P(X = 0) \]
Nuclear Power Plant

Parameters (binomial)
- number of power plants $n$: 100
- probability of an accident $p$: 0.001
- $b(n, p, k) = \binom{n}{k}p^kq^{n-k}$

Parameter (Poisson)
- $\lambda = np = 100 \times 0.001 = 0.1$
- $P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$

Probability
- $P(X \geq 1) = 1 - P(X = 0)$
  $= 1 - e^{-0.1}$

Probability
- $P(X \geq 1) = 1 - P(X = 0)$
  $= 1 - e^{-0.1}$
More than just time!

**Binomial distribution**

\[ b(n, p, k) = \binom{n}{k} p^k q^{n-k} \]

**Poisson Distribution**

\[ P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \]

- \( n = 2, 5, \ldots \)
- \( n = 50, 100, \ldots \)
- \( np = \lambda \)
An advertiser drops 10,000 videos on Instagram under 2000 hashtags. Assume that each video has an equal chance of landing on each hashtag. What is the probability that a particular hashtag will receive no videos?
10,000 videos on Instagram under 2000 hashtags

**Parameters (binomial)**

- number of videos $n$: 10000
- number of hashtags $\frac{1}{p}$: 2000

**Parameter (Poisson)**

$$\lambda = np = 10000 \times \frac{1}{2000} = 5$$
Instagram Ads

Event

- a particular hashtag will receive no videos

Random variable

- number of videos the hashtag will receive

Probability

- \(P(X = 0)\)
Instagram Ads

Parameters (binomial)

- number of videos $n: 10000$
- number of hashtags $\frac{1}{p}: 2000$

Parameter (Poisson)

\[ \lambda = np = 10000 \times \frac{1}{2000} = 5 \]

Probability

\[ P(X = 0) = e^{-5} \approx 0.00674 \]
Colorful Balls in a Jar

Experiment
Blindly pick up 5 balls (without replacement).

Random variable
Number of brown balls.

Distribution
Hypergeometric distribution.
Colorful Balls in a Jar

Experiment
Blindly pick up 5 balls (without replacement).

Event
Get 3 brown balls

\[
N \text{ choose } n \text{ samples} \\
N = 12, n = 5 \\
(N \choose n) = \binom{12}{5}
\]
Colorful Balls in a Jar

Random variable

Number of brown balls.

<table>
<thead>
<tr>
<th>$N$ choose $n$ samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N = 12$, $n = 5$</td>
</tr>
<tr>
<td>$\binom{N}{n} = \binom{12}{5}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$k$ choose $x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k = 6$, $x = 3$</td>
</tr>
<tr>
<td>$\binom{k}{x} = \binom{6}{3}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N - k$ choose $n - x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N - k = 6$, $n - x = 2$</td>
</tr>
<tr>
<td>$\binom{N - k}{n - x} = \binom{6}{2}$</td>
</tr>
</tbody>
</table>
Colorful Balls in a Jar

\[ \begin{array}{|c|c|}
\hline
\text{N choose } n \text{ samples} & \text{k choose } x \\
\hline
N = 12, n = 5 & k = 6, x = 3 \\
\hline
\binom{N}{n} = \binom{12}{5} & \binom{k}{x} = \binom{6}{3} \\
\hline
\text{N – k choose } n – x \\
\hline
N – k = 6, n – x = 2 & \\
\hline
\binom{N – k}{n – x} = \binom{6}{2} \\
\hline
\end{array} \]

Hypergeometric Distribution

\[ H(N, k, n, x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} = \frac{\binom{6}{3} \binom{6}{2}}{\binom{12}{5}} \]

XC 2020
\[ H(N, k, n, x) = \frac{{k \choose x} {N-k \choose n-x}}{{N \choose n}} \]
Hypergeometric Distribution

- A finite population of size $N$.
- Objects with the desired feature of size $k$.
- The probability of $x$ successes in $n$ draws.

Hypergeometric Distribution

\[ H(N, k, n, x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} \]
Binomial Distribution VS Hypergeometric Distribution

1 2 3 4 5
✗ ✓ ✗ ✗ ✓

without replacement

with replacement

without replacement
Binomial Distribution VS Hypergeometric Distribution

- Let $p = k/N$ remain constant as $k$ and $N$ approach $\infty$.
- It recovers to binomial distribution with parameters $n$ and $p$. 
Let \( p = k/N \) remain constant as \( k \) and \( N \) approach \( \infty \).

- It recovers to binomial distribution with parameters \( n \) and \( p \).
Example

Voting

A shareholders’ meeting has 105 female voters and 95 male voters. A random sample of 10 voters is drawn. What is the probability exactly 7 of the voters will be female?

Hypergeometric Distribution

\[ H(N, k, n, x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} \]
A shareholders’ meeting has 105 female voters and 95 male voters. A random sample of 10 voters is drawn. What is the probability exactly 7 of the voters will be female?

**Voting**

**Hypergeometric Distribution**

\[ N = 200, \ k = 105, \ n = 10, \ x = 7 \]

\[ H(N, k, n, x) = \binom{k}{x} \frac{\binom{N-k}{n-x}}{\binom{N}{n}} = H(200, 105, 10, 7) = \frac{\binom{105}{7} \binom{95}{3}}{\binom{200}{10}} \]
“I learned very early the difference between knowing the name of something and knowing something.”
Binomial distribution

\[ b(n, p, k) = \binom{n}{k} p^k q^{n-k} \]

Integers or non-integers?

- Can \( n \) be a non-integer?
- Can \( k \) be a non-integer?
- Can \( \lambda \) be a non-integer?
- Can \( \lambda \) be negative?
Binomial distribution

\[ b(n, p, k) = \binom{n}{k} p^k q^{n-k} \]

Poisson Distribution

\[ P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \]

Integers or non-integers?

- Can \( n \) be a non-integer? no
- Can \( k \) be a non-integer? no
- Can \( \lambda \) be a non-integer? yes
- Can \( \lambda \) be negative? no

XC 2020
I really can’t do a good job, any job, of explaining **magnetic force** in terms of something else you’re more familiar with, because I don’t understand it in terms of anything else you’re more familiar with.

I really can’t do a good job, any job, of explaining **Poisson distribution** in terms of something else you’re more familiar with, because I don’t understand it in terms of anything else you’re more familiar with.
Assume that the number of cars that arrive at the fork in unit time has a Poisson distribution with parameter $\lambda = 4$.

Cars coming along Wolverine Street come to a fork in the road and have to choose either Mystique Street or Magneto Street continue.

Let $X$ be the random variable which counts the number of cars that, in a given unit of time, pass by Professor X’s Book Shop on Magneto Street.

What is the distribution of $X$?
Poisson Process

Poisson distribution

\[ P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \]

Exponential distribution

\[ f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \]

### 10 Generating Functions

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.1</td>
<td>Discrete Distributions</td>
<td>365</td>
</tr>
<tr>
<td>10.2</td>
<td>Branching Processes</td>
<td>376</td>
</tr>
<tr>
<td>10.3</td>
<td>Continuous Densities</td>
<td>393</td>
</tr>
</tbody>
</table>
RICHARD FEYNMAN

“I think I can safely say that nobody understands quantum mechanics.”
HOLLYWOOD SCREENWRITERS

When you cannot explain something: use quantum mechanics!
MATH 20 BABY PROBABILISTS

When you cannot explain something: use Poisson distribution (Poisson process)!