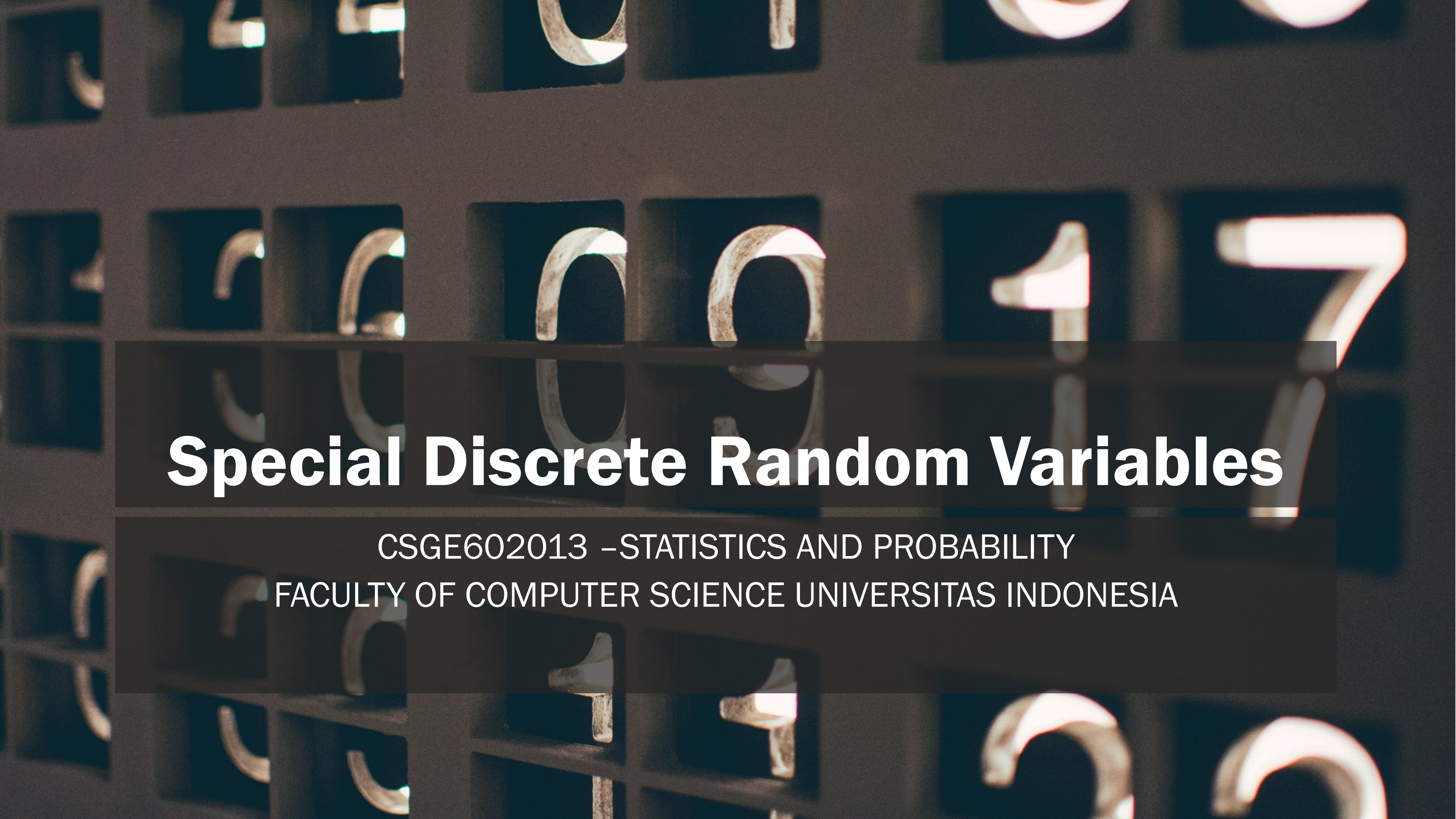


# Special Discrete Random Variables

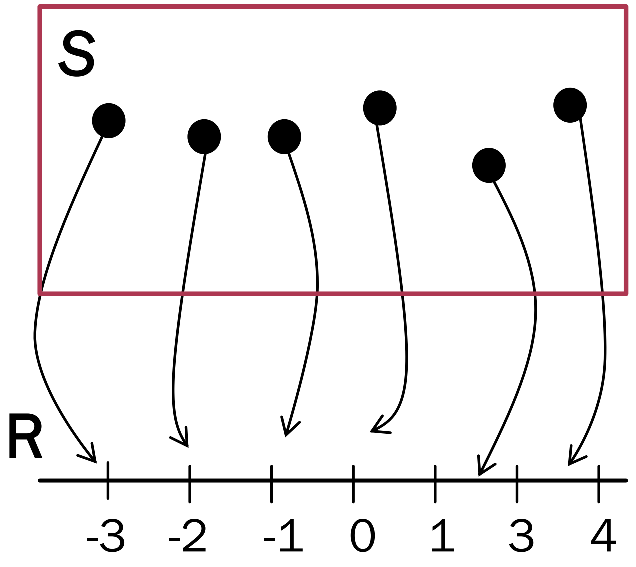
CSGE602013 –STATISTICS AND PROBABILITY

FACULTY OF COMPUTER SCIENCE UNIVERSITAS INDONESIA

# [Recap] Random Variables

$$X : S \rightarrow R \qquad (or\ X(s) \in R, \forall s \in S)$$

All possible outcomes in **S**



# [Recap] Describing the Probabilities of RVs

- The **Probability Mass Function (PMF)** $p_X(x_i)$ of $X$ is defined by

$$p_X(x_i) = P(X = x_i)$$

“the probability that the value of $X$ is *exactly* equal to $x_i$”

- **Cumulative Distribution Function (CDF)**, $F_X(x)$ of the random variable $X$ is defined for any real number $x$ by

$$F_X(x) = P(X \leq x) \qquad x \in \mathbb{R}$$

“the probability that the value of $X$ is less than or equal to $x_i$”

# [Recap] Describing the Probabilities of RVs (2)

- Probability Density Function (PDF), $f_X(x)$ of the random variable $X$ is probability that RV $\boldsymbol{X}$ will be in the interval $a \leq X \leq b$

$$P(a \leq X \leq b) = \int_a^b f_x(x)\, dx$$

"the probability that the value of $X$ is between a and b"

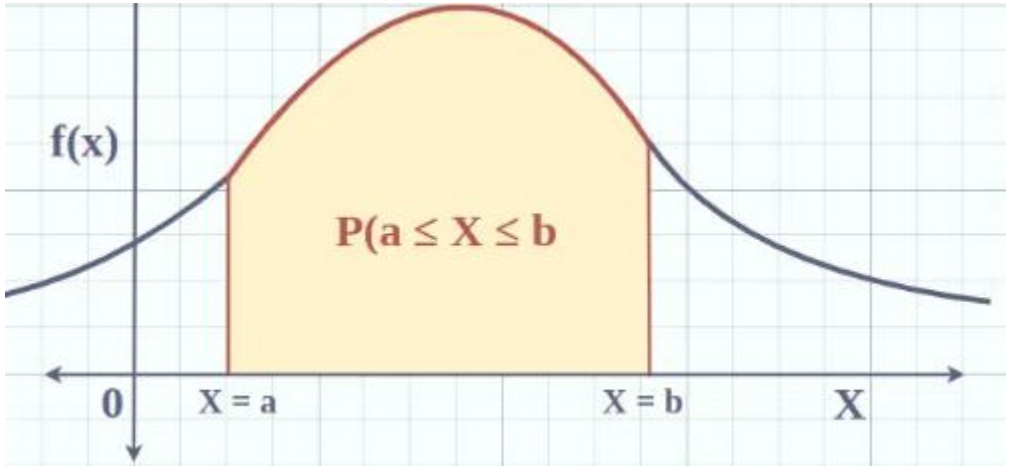


Image from: https://www.geeksforgeeks.org/maths/real-life-applications-of-continuous-probability-distribution/

# Discrete and Continuous RVs

## Discrete Random Variable

Takes only finite or countably infinite number of values

Examples?

There are some special types of discrete variables…

## Continuous Random Variable

Takes infinite and uncountable number of values

Examples?

# Special Discrete Random Variables

01 Bernoulli Random Variables

02 Binomial Random Variables

03 Geometric Random Variables

04 Negative Binomial Random Variables

05 Poisson Random Variables

06 Hyper-geometric Random Variables

# Consider...

- Let X be the random variable with only two possible outcome values.
  - $X = 1$ (e.g., when the outcome is a “success”, or Head in a coin toss)
  - $X = 0$ (e.g., when the outcome is a “failure”, or Tail in a coin toss)

Either yes or no... Either this or that..

- Examples:

1. The outcome of a coin toss {Head, Tail}
2. Whether a valve is open or shut
3. Whether an item is defective or not

# Definition of Bernoulli Random Variables

- The probability mass function (PMF) of $X$ is given by

$$P(X=1)=p \qquad \text{or} \qquad P(X=x)=p^{x}(1-p)^{1-x} \qquad x=0,1$$
$$P(X=0)=1-p$$

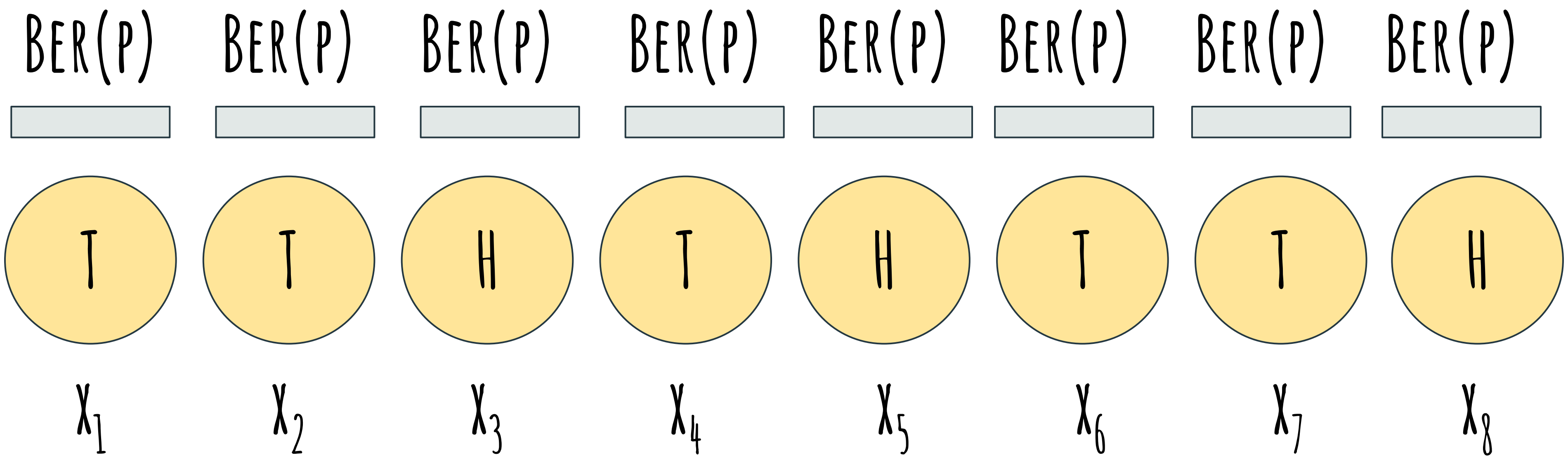
where $p, 0 \leq p \leq 1,$ is the probability that the trial is a “success”.

- If X is a Bernoulli Random Variable (variable that has Bernoulli distribution) with parameter p, we can write

$$X \sim Ber(p)$$

# **Intuition** of Bernoulli Random Variables

- **Bernoulli random variable** models the outcome of a *single binary trial*.
- Suppose we have 8 independent coins, then each coin's toss is a Bernoulli RVs.



# Expectation and Variance of Bernoulli RVs

$$E[X] = \sum_{x \in \{0,1\}} x \cdot p(x) = 1 \cdot p + 0 \cdot (1 - p) = p$$

$$E[X^2] = 1^2 \cdot p + 0^2 \cdot (1 - p) = p$$

$$Var[X] = E[X^2] - (E[X])^2 = p - p^2 = p(1 - p)$$

# Properties of Bernoulli Random Variables $X \sim Ber(p)$

PMF

$$P(X = x) = p^{x}(1-p)^{1-x} \qquad x = 0, 1$$

CDF

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

Expectation / Mean

$$\mu = E[X] = 1.P(X = 1) + 0.P(X = 0) = p$$

Variance

$$\begin{aligned} \sigma^2 &= Var(X) \\ &= E[X^2] - (E[X])^2 \\ &= 1^2.P(X = 1) + 0^2.P(X = 0) - (p)^2 \\ &= p(1-p) \end{aligned}$$

# Now, Consider…

- **Repeat a Bernoulli trial $\boldsymbol{n}$ times..**
  - $\boldsymbol{n}$ Bernoulli Trials $(\boldsymbol{X_1}, \boldsymbol{X_2}, \dots, \boldsymbol{X_n})$ that are **independent**,
  - each $\boldsymbol{X_i}$ has a constant probability $p$ of success.

- For example,
  1. the number of heads in $n$ coin flips
  2. the number of disk drives that crashed in a cluster of 1000 computers
  3. the number of advertisements on a webpage that are clicked by 200 visitors

# Definition of Binomial Random Variables

- If X represents the number of successes that occur in the $n$ trials,

$$X = X_1 + X_2 + \cdots + X_n,$$

then X is said to be a *Binomial random variable* (a variable that has Binomial distribution) with parameters $(n, p)$.

$$X \sim Bin(n, p)$$

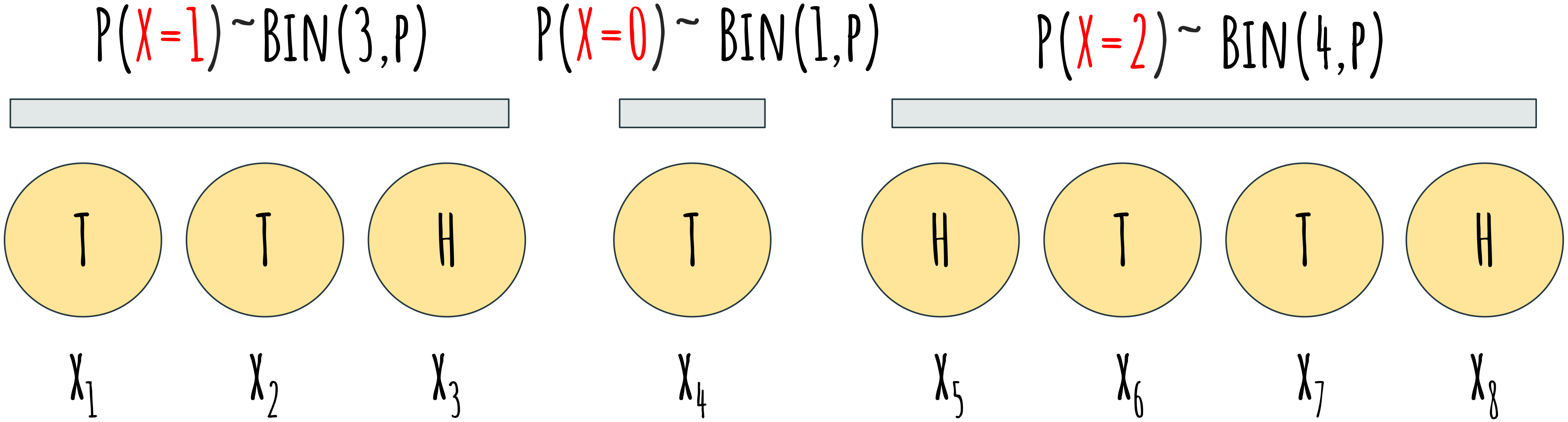
where $n$ is the number of trials, and $p$ is the probability of getting “success” or 1 in each trial.

- A Binomial R.V. can be viewed as the sum of n independent Bernoulli R. V.

$$X = X_1 + X_2 + X_3 + \ldots + X_n \qquad X_i \sim Ber(p)$$

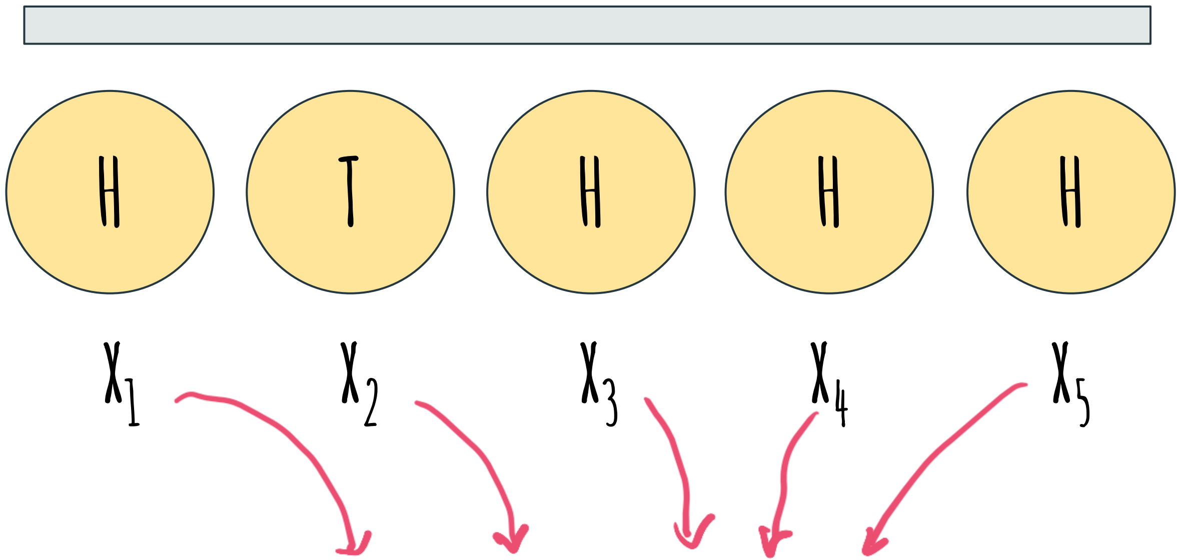
# Intuition of Binomial Random Variables

- **Binomial random variable** models the *total number of successes from $n$ independent Bernoulli trials*.
- Suppose we assume “success” in a coin toss when we get Head.



# Combinatorics in the Binomial Distribution

$P(X=4) \sim BIN(5,p)$



$$P(HTHHH) = p \cdot (1-p) \cdot p \cdot p \cdot p = p^4(1-p)^1$$

$$\boxed{P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}}$$

- Suppose we flip a coin 5 times (i.e, $n = 5$).
- Suppose $X = 4$, i.e., the total number of Head from 5 tosses is 4.
- *Notice we have 5 different ways to get 4 heads in 5 independent trials.*

1. $THHHH$
2. $HTHHH$
3. $HHTHH$
4. $HHHTH$
5. $HHHHT$

$\binom{5}{4}$ ways to get 4 heads out of 5

# Properties of Binomial Random Variables $\boldsymbol{X \sim Bin(n,p)}$

PMF

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{where} \quad k = 0,1,2,\ldots,n$$

CDF

$$F_X(i) = P(X \le i) = \begin{cases} 0 & i < 0 \\ \sum_{k=0}^{\lfloor i \rfloor} \binom{n}{k} p^k (1-p)^{n-k}, & 0 \le i < n \\ 1 & i \ge n \end{cases}$$

Expectation / Mean

$$\begin{aligned} E[X] &= E[X_1 + X_2 + \ldots + X_n] \\ &= E[X_1] + \ldots + E[X_n] = np \end{aligned}$$

Variance

$$\begin{aligned} Var(X) &= Var(X_1 + X_2 + \ldots + X_n) \\ &= Var(X_1) + \ldots + Var(X_n) = np(1-p) \end{aligned}$$

# Example #0

- A factory produces 100 cars per day, but a car is defective with probability 0.02. What is the probability the factory produces 2 or more defective cars?

1. We define the random variable: $X \sim \mathit{Binomial}(n = 100, p = 0.02)$
2. The factory produces 2 or more defective cars: $P(X \geq 2)$
3. From the definition of probability:

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1)$$

4. Compute $P(X = 0)$ and $P(X = 1)$ using Binomial RV's PMF:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

# Conditions for a Binomial RV

1. Fixed number of trials (n): The number of trials must be predetermined.
2. Two possible outcomes per trial: Each trial should be a success or a failure.
3. Constant probability (p) per trial: The probability of success should remain the same across trials.
4. Independence: The outcome of one trial should not affect the others.

# These are not Binomial RV – Why?

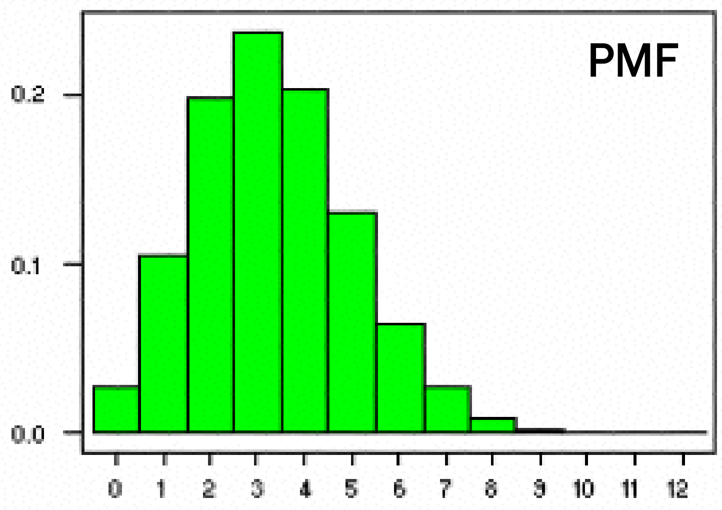
1. You stand in front of FASILKOM. Let, X = number of students passing by in the next 5 minutes.
2. Gather a random sample of 5 men and 5 women. Let, X = number of persons out of 10 who are more than 170 cm tall
3. Draw 4 cards (without replacement) from a deck of 52 cards. Let, X = number of aces among the four

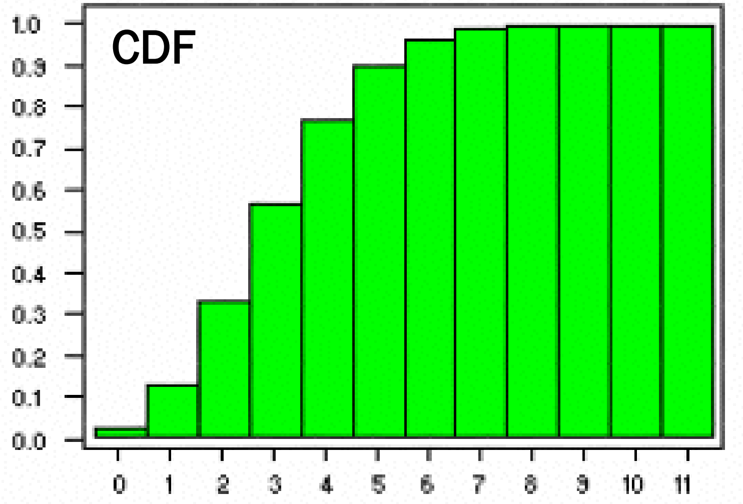
# PMF and CDF of Binomial Random Variable (1)

- CDF of Binomial R.V.

$$F_X(i) = P(X \le i) = \begin{cases} 0 & i < 0 \\ \sum_{k=0}^{\lfloor i \rfloor} \binom{n}{k} p^k (1-p)^{n-k}, & 0 \le i < n \\ 1 & i \ge n \end{cases}$$

- Visualizations of PMF & CDF of $Bin\left(20, \frac{1}{6}\right)$:





# PMF and CDF of Binomial Random Variable (2)

- CDF of Binomial R.V.

$$F_X(i) = P(X \le i) = \begin{cases} 0 & i < 0 \\ \sum_{k=0}^{\lfloor i \rfloor} \binom{n}{k} p^k (1-p)^{n-k}, & 0 \le i < n \\ 1 & i \ge n \end{cases}$$

- Example:
  - Given $X \sim Bin(8, 0.5)$

1. P (X = 3) ?

$$P(X=3) = \binom{8}{3} (0.5)^3 (1-0.5)^5 = 0.219$$

# PMF and CDF of Binomial Random Variable (3)

- CDF of Binomial R.V.

$$F_X(i) = P(X \le i) = \begin{cases} 0 & i < 0 \\ \sum_{k=0}^{\lfloor i \rfloor} \binom{n}{k} p^k (1-p)^{n-k}, & 0 \le i < n \\ 1 & i \ge n \end{cases}$$

- Example:
  - Given $X \sim Bin(8, 0.5)$

2. P (X ≤ 1) = ?

$$P(X \le 1) = P(X = 0) + P(X = 1)$$

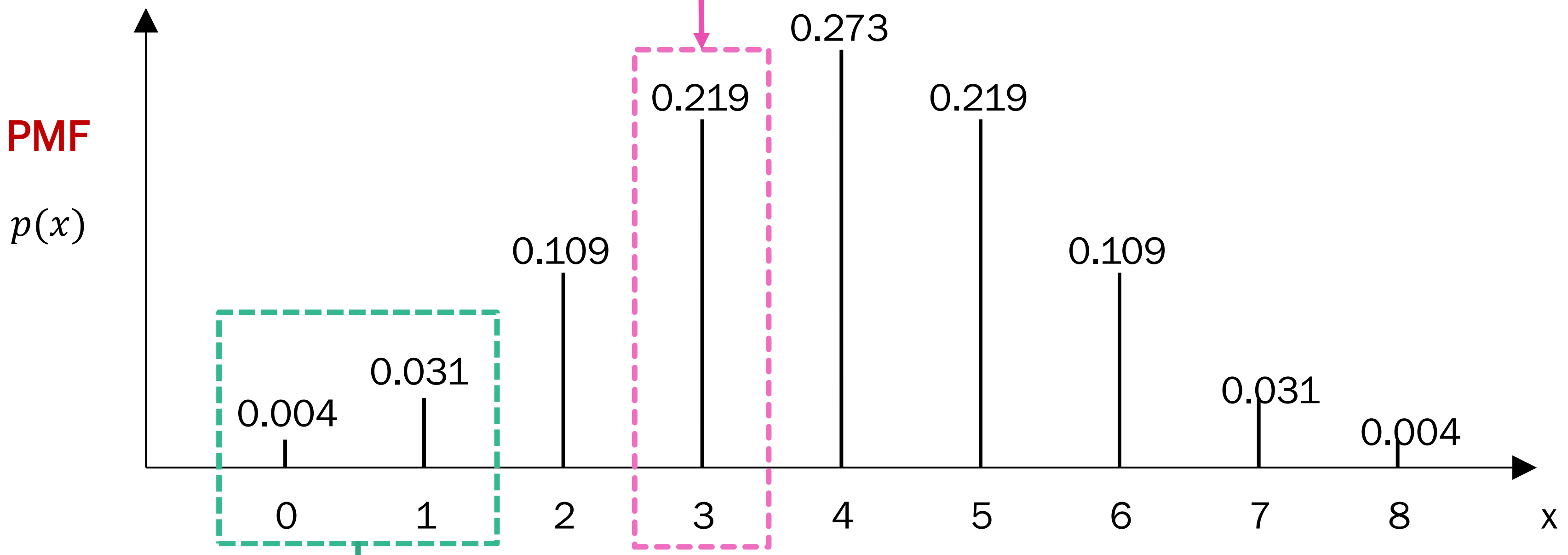
$$= \binom{8}{0}(0.5)^0(1-0.5)^8 + \binom{8}{1}(0.5)^1(1-0.5)^7$$

$$= 0.035$$

# PMF and CDF of Binomial Random Variable (4)

- Given $X \sim Bin(8, 0.5)$

P (X = 3)

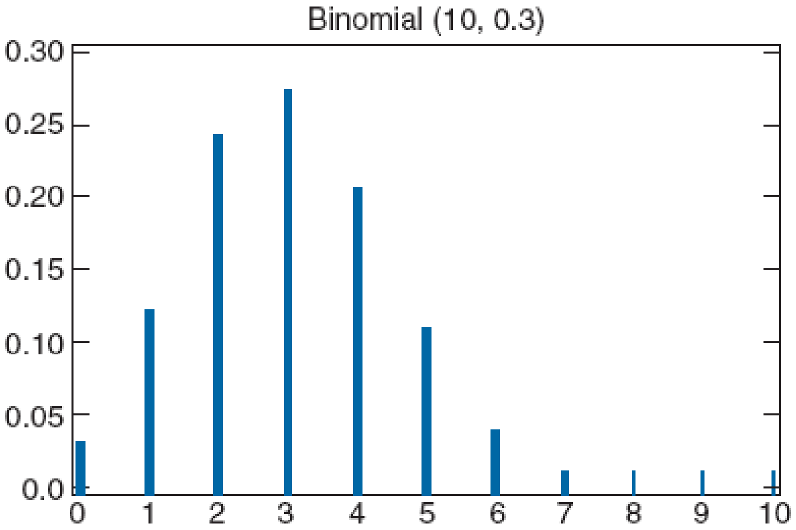


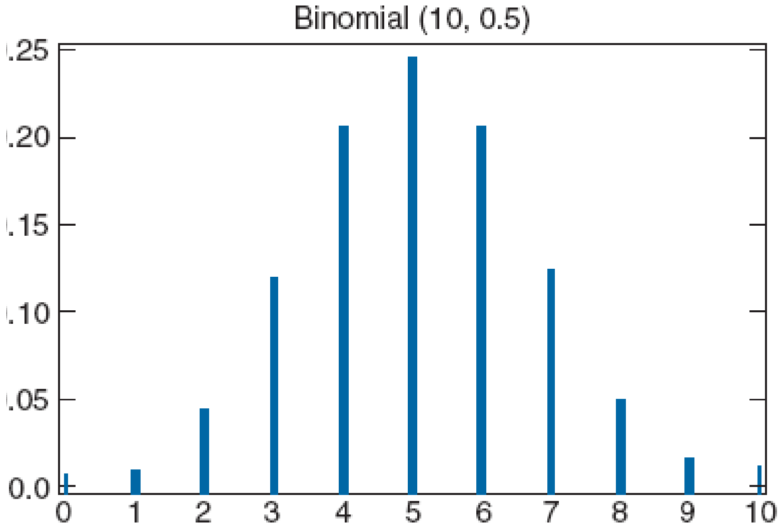
P (X ≤ 1)

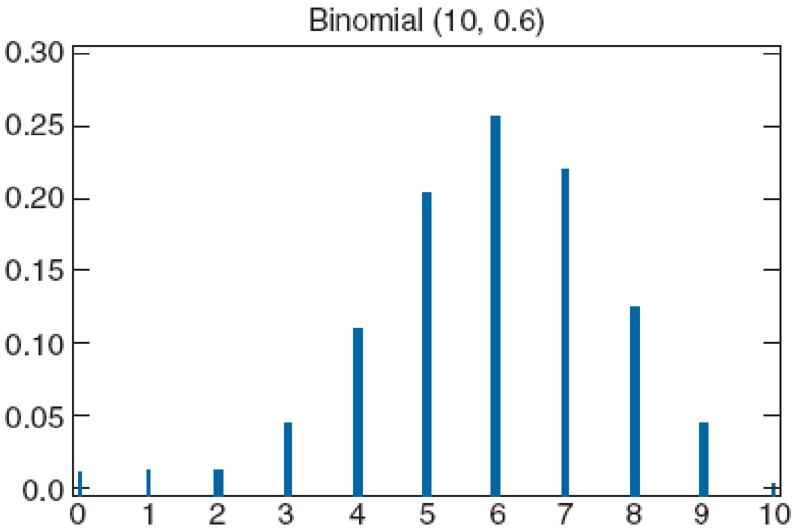
CDF

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| $P(X \leq x)$ | 0.004 | 0.035 | 0.144 | 0.363 | 0.636 | 0.855 | 0.965 | 0.996 | 1.0000 |

# Other PMF Examples of Binomial Random Variable


Binomial (10, 0.3)


Binomial (10, 0.5)


Binomial (10, 0.6)

# Exercise #1

Let random variable **X** be the number of **Heads** obtained from an experiment of tossing **three** coins. Suppose, probability that **Head** occurs when we toss a coin is **p**.

(a) What are possible values for **X** ?

(b) What kind of distribution does **X** follow ?

(c) Determine **PMF** of X !

(d) Determine **CDF** of X !

# Exercise 1

Let random variable **X** be the number of **Heads** obtained from an experiment of tossing **three** coins. Suppose, probability that **Head** occurs when we toss a coin is **p.**

(a) What are possible values for **X** ?

$0,1,2,3$

(b) What kind of distribution does **X** follow ?

$X \sim Bin(3, p)$

# Exercise 1

Let random variable **X** be the number of **Heads** obtained from an experiment of tossing **three** coins. Suppose, probability that **Head** occurs when we toss a coin is **p.**

(c) Determine **PMF** of X !

$$P(X = k) = \binom{3}{k} p^k (1-p)^{3-k} \qquad k = 0,\ 1,\ 2,\ 3$$

(d) Determine **CDF** of X !

$$F_X(a) = P(X \le a) = \begin{cases} 0 & a < 0 \\ \sum_{k=0}^{\lfloor a \rfloor} \binom{3}{k} p^k (1-p)^{3-k}, & 0 \le a < 3 \\ 1 & a \ge 3 \end{cases}$$

# Exercise 2

- A company produces disks with the defect rate of 0.01. One pack of disk contains 10 discs. The company policy is to give money back if one pack contains more than 1 defect disc. If someone buy three packs, what is the probability that exactly one pack is to be returned ?

  1. First, let's calculate the probability of one pack contains more than 1 defect disk. Let us define X be the number of defect disc in one pack. How do we calculate it?

# Exercise 2

- A company produces disks with the defect rate of 0.01. One pack of disk contains 10 discs. The company policy is to give money back if one pack contains more than 1 defect disc. If someone buy three packs, what is the probability that exactly one pack is to be returned ?

1. First, let's calculate the probability of one pack contains more than 1 defect disk. Let us define X be the number of defect disc in one pack. How do we calculate it?

$$X \sim Bin(10, 0.01)$$

$$\begin{aligned} P(X > 1) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - \binom{10}{0}(0.01)^0(0.99)^{10} - \binom{10}{1}(0.01)^1(0.99)^9 \\ &\approx 0.005 \end{aligned}$$

# Exercise 2

- A company produces disks with the defect rate of 0.01. One pack of disk contains 10 discs. The company policy is to give money back if one pack contains more than 1 defect disc. If someone buy three packs, what is the probability that exactly one pack is to be returned ?
- The probability that a package will have to be replaced $is\ \boldsymbol{P}(\boldsymbol{X} > \mathbf{1}) = 0.005$.
- Now, let us define Y be the number of packages that the person will have to return when he/she buys three packs. How do we calculate it?

$$Y \sim Bin(3, 0.005)$$

$$P(Y=1) = \binom{3}{1}(0.005)^1(0.995)^2 = 0.015$$

# Let Use Now Consider…

- Repeat a Bernoulli trial $n$ times..
  - $n$ Bernoulli Trials $(X_1, X_2, \dots, X_n)$ that are **independent**,
  - each $X_i$ has a constant probability $p$ of success.
  - $X$ is the number of trials up to and including the first success

- Examples

1. Toss a coin repeatedly. Let X = number of tosses to *the first head*
2. One percent of bits transmitted through a digital transmission are received in error. Bits are transmitted until the first error. Let X = the number of bits transmitted until *the first error*

# Definition of Geometric Random Variables

- In a sequence of independent Bernoulli trials, each with probability $p$ of success.
- Let $X$ be the number of trials up to and including the first success.
- We say $X$ is a Geometric Random Variable (variable that has geometric distribution) with parameter $p$.

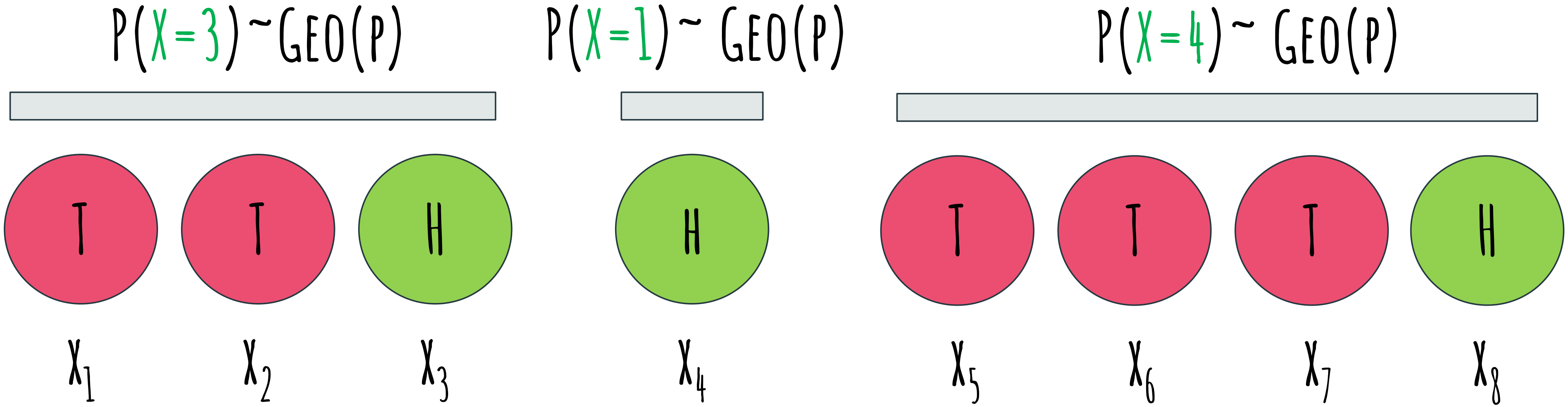
$$X \sim Geo(p)$$

- The PMF of $\boldsymbol{X}$ is

$$P(X = k) = (1 - p)^{k-1} p$$

$$k = 1,2,3,4, \ldots$$

# Visualizations of Geometric Random Variables



1. There are always $k-1$ failure trials! Therefore, their probability value is $(1-p)^{k-1}$
2. The success trial happens in the $k$-th trial! The probability until the first success at $k$-th is

$$P(X=k)=(1-p)^{k-1}p$$

# Properties of Geometric Random Variables $X{\sim}Geo(p)$

PMF

$$P(X = k) = (1-p)^{k-1}p\,, \qquad k = 1,2,3,\ldots$$

CDF

$$F_X(k) = P(X \le k) = \sum_{r=1}^{\lfloor k \rfloor} p(1-p)^{r-1} = 1-(1-p)^{\lfloor k \rfloor}, \qquad x \ge 1$$

$$F_X(k) = \begin{cases} 0 & k < 1 \\ 1-(1-p)^{\lfloor k \rfloor} & k \ge 1 \end{cases}$$

Expectation / Mean

$$\mu = E[X] = \frac{1}{p}$$

Variance

$$\sigma^2 = Var(X) = \frac{1-p}{p^2}$$

# Proof

- Expectation

$$E[X] = \sum_{x=1}^{\infty} x(1-p)^{x-1} p = p\sum_{x=1}^{\infty} x(1-p)^{x-1} = p.\frac{1}{p^2} = \frac{1}{p}$$

- Variance

$$Var(X) = E[X^2] - (E[X])^2$$

$$E[X^2] = \sum_{x=1}^{\infty} x^2(1-p)^{x-1} p = p\sum_{x=1}^{\infty} x^2(1-p)^{x-1}$$

$$\sum_{x=1}^{\infty} x^2(1-p)^{x-1} = \frac{1}{p}\left(\frac{1}{p} + \frac{2(1-p)}{p^2}\right) = \frac{2-p}{p^3}$$

$$Var(X) = p\left(\frac{2-p}{p^3}\right) - \frac{1}{p^2} = \frac{1-p}{p^2}$$

# Example 1

- The probability that Budi can score when he throws his basketball is 0.6. Assume each throw is independent from each other.

1. What is the probability Budi needs 3 throws until he is able to score?
2. What is the probability Budi needs *at least* 3 throws to score?
3. Estimate how many throws he needs to do to score!

# Example 1

- X : The RV that denotes how many throws Budi does until he scores

$$X \sim Geo(0.6)$$

1. What is the probability Budi needs 3 throws until he is able to score?

$$P(X=3)=(1-p)^2 p=(0.4)^2(0.6)=0.096$$

2. What is the probability Budi needs at least 3 throws to score?

$$\begin{aligned} P(X \geq 3) &= 1-P(X=1)-P(X=2) \\ &= 1-(0.4)^0(0.6)-(0.4)^1(0.6) \\ &= 1-0.4-0.24=0.36 \end{aligned}$$

3. Estimate how many throws he needs to do to score!

$$E[X]=\frac{1}{p}=\frac{1}{0.6}=1.67$$

# Example 2

- If a person is unsuccessful in starting the old car’s engines, then he must wait 10 minutes before trying again. In each attempt, the success probability is 0.75.

1. What is the probability that the old car’s engines start on **the third** attempt ?
2. What is the probability that the engines start **within 20 minutes** of the first attempt ?
3. What is the **expected number of attempts required** to start the engines ?

# Example 2

- If a person is unsuccessful in starting the old car's engines, then he must wait 10 minutes before trying again. In each attempt, the success probability is 0.75.

1. What is the probability that the old car's engines start on **the third** attempt ?

- X : the number of trials until the engines start

$$P(X=3)=(0.25)^2.(0.75)=0.047$$

# Example 2

- If a person is unsuccessful in starting the old car's engines, then he must wait 10 minutes before trying again. In each attempt, the success probability is 0.75.

2. What is the probability that the engines start **within 20 minutes** of the first attempt ?

- X : the number of trials until the engines start

$$P(X \leq 3) = 1 - (1 - 0.75)^3 = 0.984$$

# Example 2

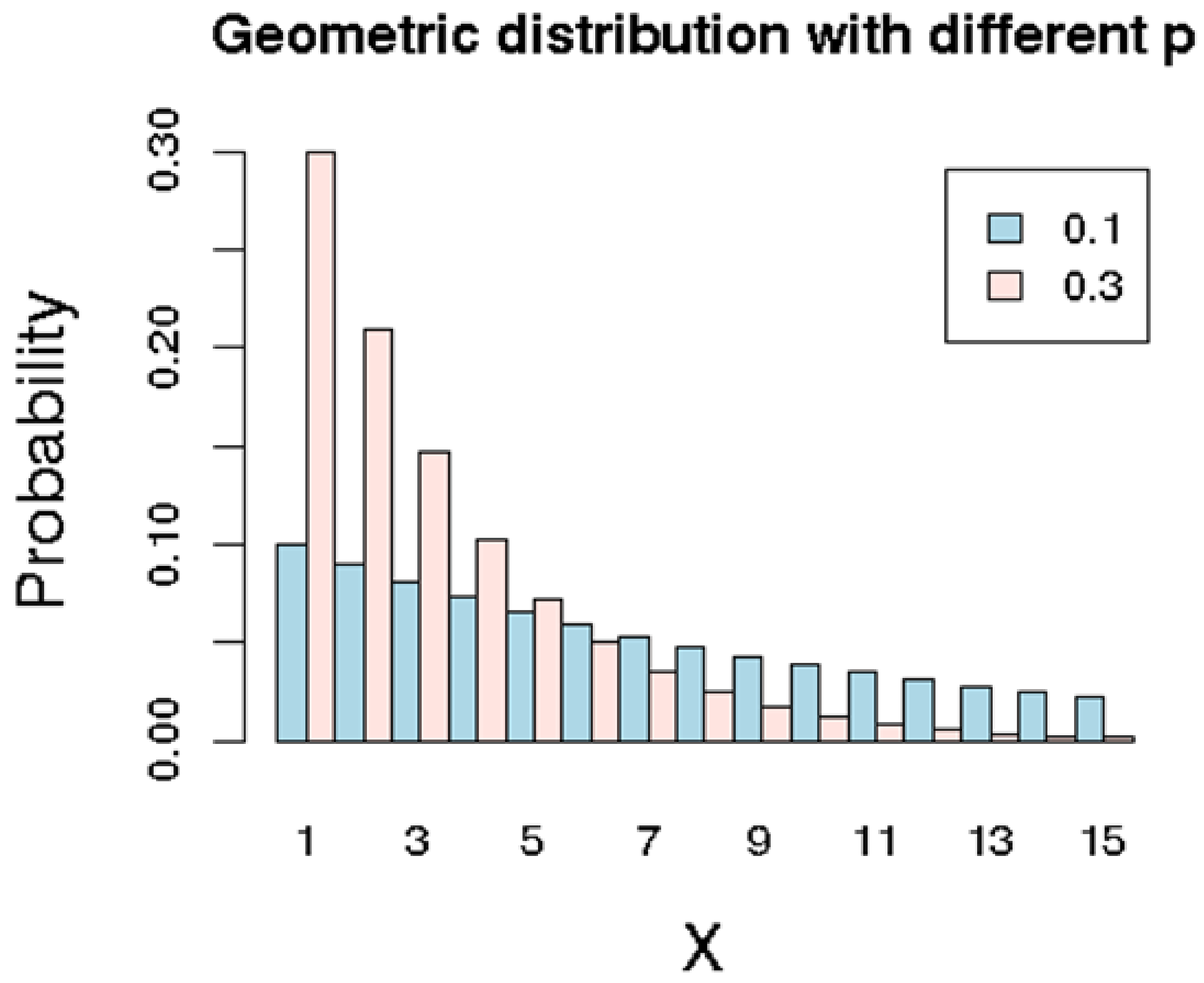
- If a person is unsuccessful in starting the old car’s engines, then he must wait 10 minutes before trying again. In each attempt, the success probability is 0.75.

3. What is the **expected number of attempts required** to start the engines ?

- X : the number of trials until the engines start

$$E[X]=\frac{1}{p}=\frac{1}{0.75}=1.33$$

# PMF of Geometric Random Variable



# Memoryless Property of Geometric Random Variable

- The geometric RV has a memoryless property
- A nonnegative random variable $X$ is memoryless if

$$P(X > r + m \mid X > m) = P(X > r) \qquad r, m \geq 0$$

  - Condition has occurred: failure in all previous $m$ trials, so that "$X > m$" as a prior condition, and
  - We will calculate the probability of success will happen in the $(\mathrm{r+m})^{th}$ experiment with this prior condition.
- This property states that the value of $r$ in the probability $P(X = r)$ is always calculated from current condition (we don't care about how many previous failures).

# Explanations on the Memoryless Property

- What does the “Memoryless Property” mean intuitively?
  - If you have already failed for $m$ trials, the probability of requiring at least $r$ more trials is the same as if you were starting fresh.
  - The process does not "remember" past failures; the probability of success remains unchanged regardless of how long you've already waited.
- Why is this important?
  - The geometric distribution is the only discrete probability distribution that has the memoryless property.
  - This makes it useful in modeling waiting times for the first occurrence of an event, such as (1) number of coin flips until the first heads, (2) number of calls until you reach a busy line, and (3) number of attempts needed to hit a goal in sports.
- Mathematical Justification
  - The probability of failure in each trial is still (1−p).

# Memoryless Property Proof

$$
\begin{aligned}
&P(X > r + m \mid X > m) \\
&= \frac{P(X > r + m, X > m)}{P(X > m)} \\
&= \frac{P(X > r + m)}{P(X > m)} \\
&= \frac{1 - P(X \le r + m)}{1 - P(X \le m)} \\
&= \frac{1 - (1 - (1 - p)^{r+m})}{1 - (1 - (1 - p)^{m})}
\end{aligned}
$$

$$
\begin{aligned}
&= (1 - p)^{r} \\
&= 1 - (1 - (1 - p)^{r}) \\
&= P(X > r)
\end{aligned}
$$

**Other consequence:**

$$P(X > r + m) = P(X > r)P(X > m)$$

# Exercise

- In a sequence of Bernoulli experiments to get number “6” in rolling a die:

1. What is the probability to get 6 successes in 10 experiments?
2. What is the probability to get the first success at the 10th experiment?
3. What is the average number of successes in 10 experiments?
4. What is the average number of experiments to get the first success?

# Consider..

- A random variable that counts the number of trials required (i.e., $k$) to get a fixed number of $r$ successes in independent Bernoulli trials.
- Let us define:
  - Each trial succeeds with probability $p$
  - Define $X$ be the total number of trial (i.e., $k$) on which the $r^{th}$ success occurs
- Example
  1. Shots needed by a basketball player to make 5 baskets.
  2. Numbers of call needed by a sales to close 10 deals.

# Definition of Negative Binomial Random Variables

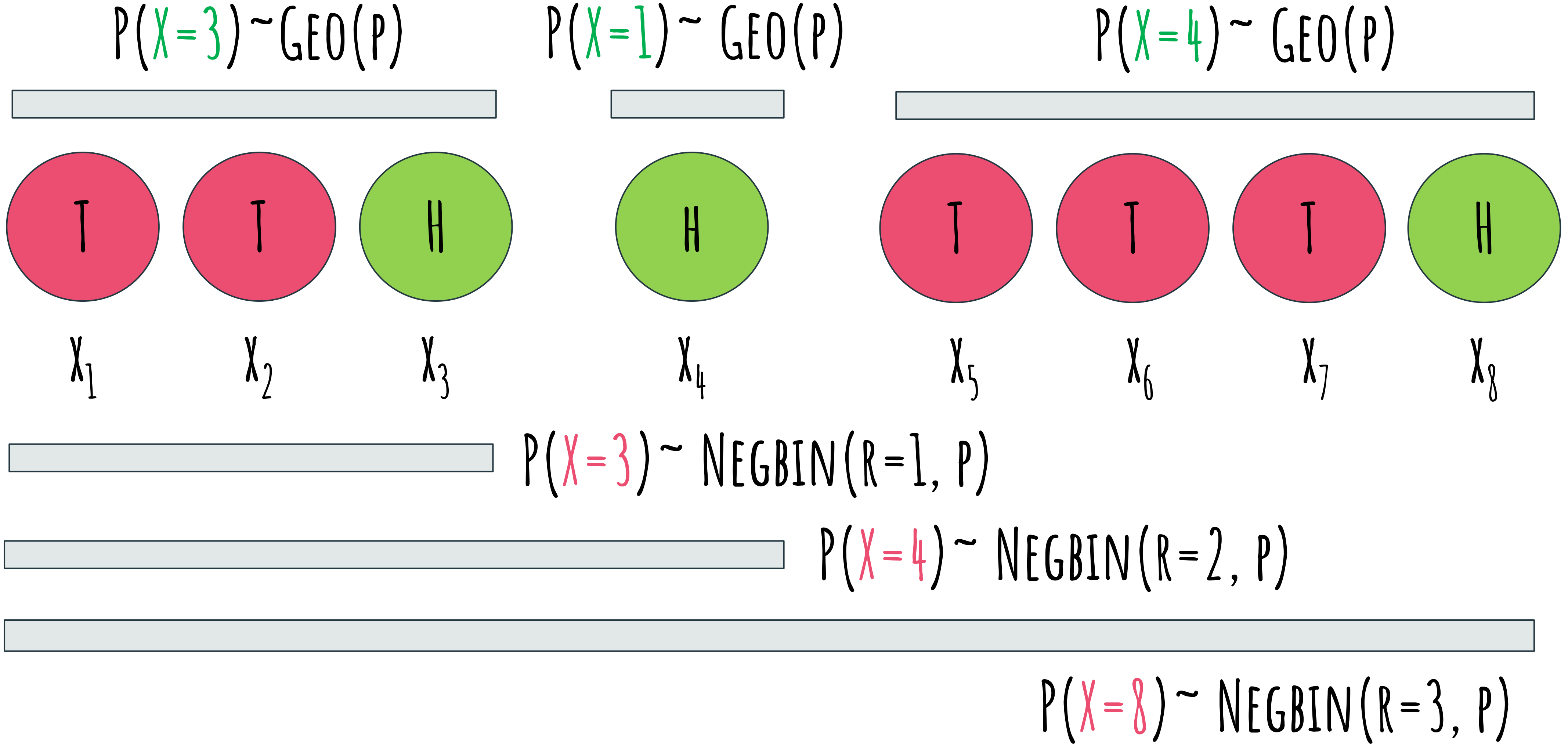
- A Negative Binomial random variable models the number of trials required $k$ to get a fixed number of $r$ successes in independent Bernoulli trials.
- Let $X$ be the total number of trial (i.e., $k$) on which the $r^{th}$ success occurs. Each trial succeeds with probability $p$.
- We say $X$ is a Negative Binomial Random Variable (variable that has Negative Binomial distribution) with parameter $r$ (i.e., number of success) and $p$ (i.e., probability of success).

$$X \sim NegBin(r, p)$$

- The PMF of $X$ is

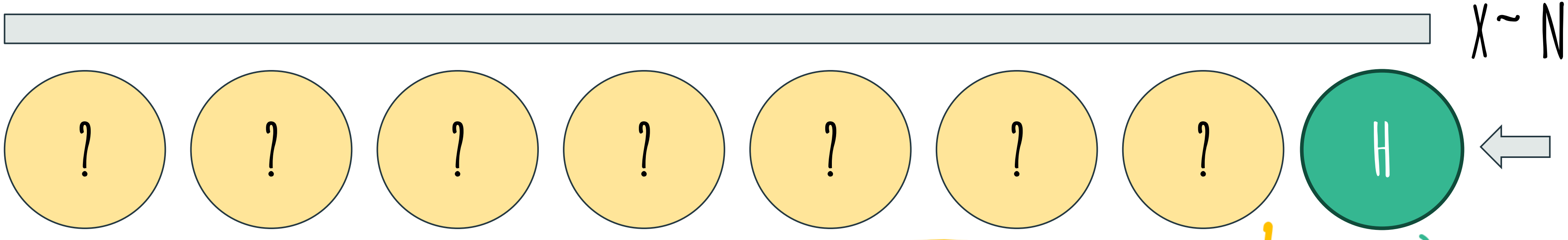
$$P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \qquad k = r, r+1, r+2, \ldots$$

# Negative Binomial RVs vs. Geometric RVs



# Combinatorics in the Negative Binomial Distribution

X~ NegBin(R=3, P)



*Must be H*

- Exactly 2 (i.e., $r-1$) of the first 7 (i.e., $k-1$) must be Heads!
  - In this case, $r=3$ and $k=8$, where $X \sim NegBin(r=3, p)$.

$$P(X=8)=\binom{8-1}{3-1}\overbrace{p^2\cdot(1-p)^5}\cdot p$$

$$\boxed{P(X=k)=\binom{k-1}{r-1}p^r(1-p)^{k-r}}$$

# Properties of Negative Binomial RVs

$X \sim NegBin(r, p)$

PMF

$$P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \qquad k = r, r+1, r+2 \ldots$$

CDF

$$F_X(k) = P(X \le k) = \sum_{j=r}^{k} \binom{j-1}{r-1} p^r (1-p)^{j-r}$$

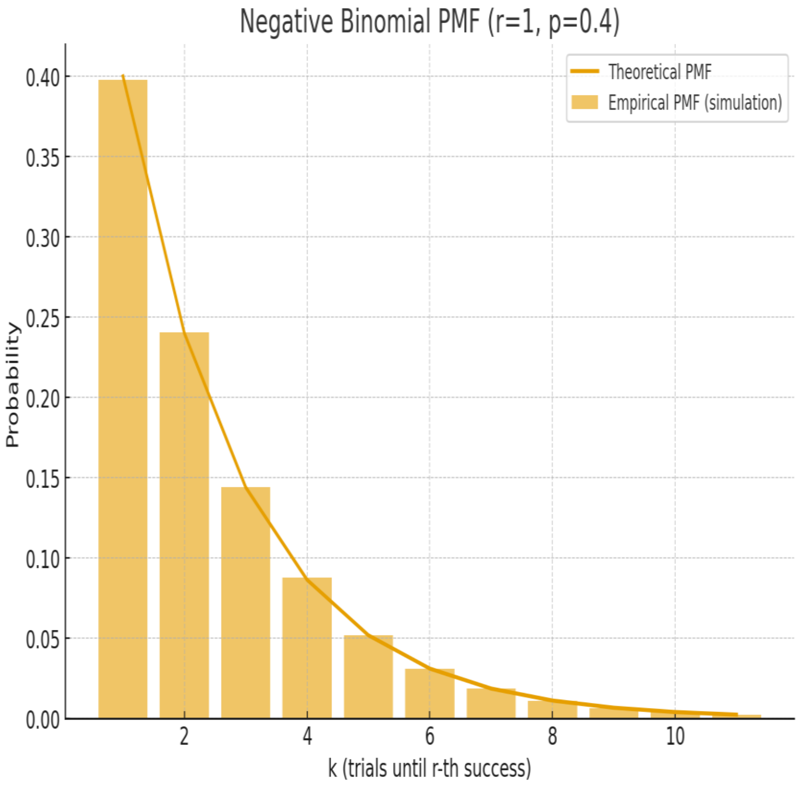
Expectation / Mean

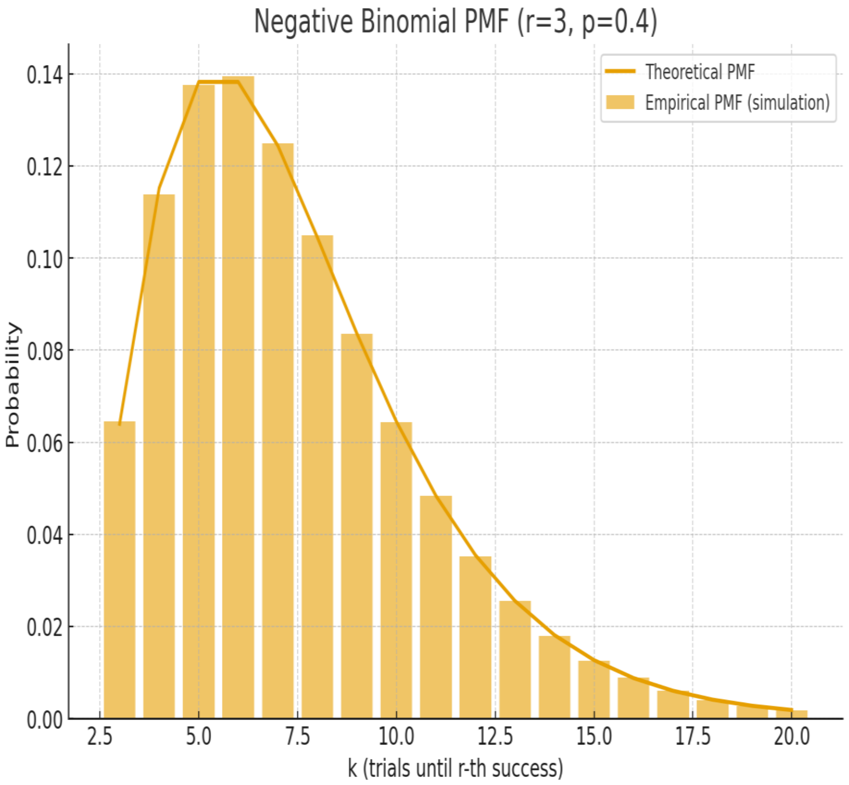
$$\mu = E[X] = \frac{r}{p}$$

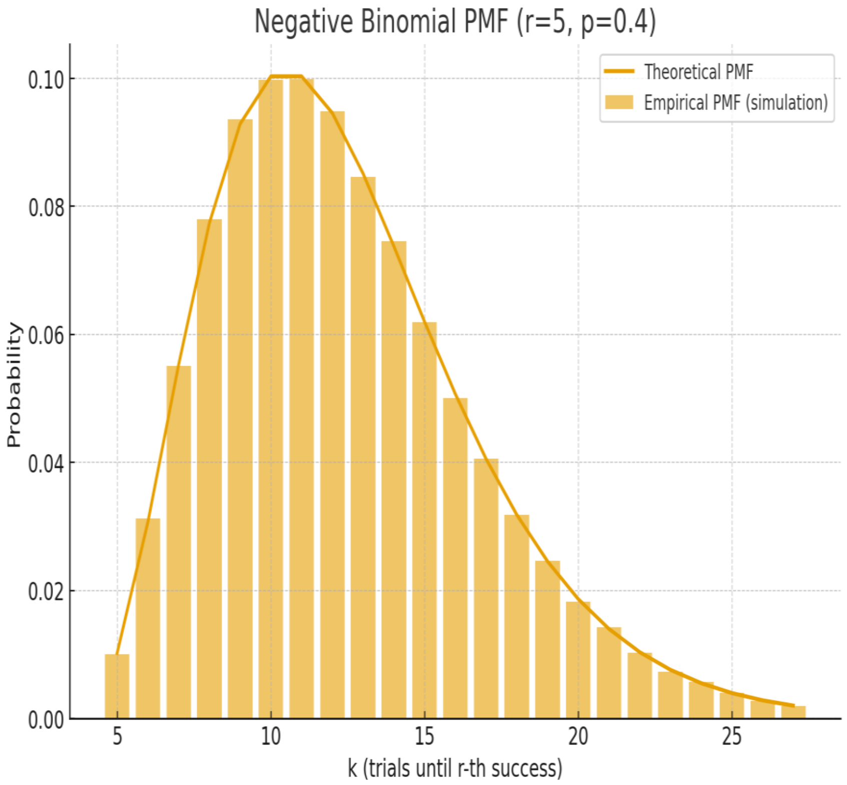
Variance

$$\sigma^2 = Var(X) = \frac{r(1-p)}{p^2}$$

# PMF of Geometric Random Variable







# Exercise

- A taxi driver in Jakarta is trying to reach a target of **5 passengers** before taking a lunch break. Based on past experience, the driver knows that for each stop, there is a **30% chance** ($p = 0.3$) that someone will take the taxi. Each stop is independent of the others.
- Answer the following questions:
  1. What is the probability that the driver meets the 5th passenger exactly on the **12th stop**?
  2. What is the probability that the driver meets the 5th passenger **on or before the 12th stop**?
  3. On average, how many stops does the driver need before he gets 5 passengers?
  4. What is the variance of the number of stops required?

# Exercise

- A taxi driver in Jakarta is trying to reach a target of **5 passengers** before taking a lunch break. Based on past experience, the driver knows that for each stop, there is a **30% chance** ($p = 0.3$) that someone will take the taxi. Each stop is independent of the others.

1. What is the probability that the driver meets the 5th passenger exactly on the **12th stop**?

$$P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \qquad k = r, r+1, r+2 \ldots$$

$$P(X = 12) = \binom{12-1}{5-1} (0.3)^5 (0.7)^{12-5} = \binom{11}{4} (0.3)^5 (0.7)^{12-5}$$

$$= 330 \cdot (0.00243) \cdot (0.08235) \approx 0.066$$

# Exercise

- A taxi driver in Jakarta is trying to reach a target of **5 passengers** before taking a lunch break. Based on past experience, the driver knows that for each stop, there is a **30% chance** ($p = 0.3$) that someone will take the taxi. Each stop is independent of the others.

2. What is the probability that the driver meets the 5th passenger **on or before the 12th stop**?

$$P(X \leq k) = \sum_{j=r}^{k} \binom{j-1}{r-1} p^r (1-p)^{j-r}$$

$$P(X \leq 12) = \sum_{j=5}^{12} \binom{j-1}{4} (0.3)^5 (0.7)^{j-5} \approx 0.647$$

# Exercise

- A taxi driver in Jakarta is trying to reach a target of **5 passengers** before taking a lunch break. Based on past experience, the driver knows that for each stop, there is a **30% chance** $(p = 0.3)$ that someone will take the taxi. Each stop is independent of the others.

3. On average, how many stops does the driver need before he gets 5 passengers?

$$E[X] = \frac{r}{p} = \frac{5}{0.3} \approx 16.67$$

On average, the driver needs **about 17 stops.**

# Exercise

- A taxi driver in Jakarta is trying to reach a target of **5 passengers** before taking a lunch break. Based on past experience, the driver knows that for each stop, there is a **30% chance** $(p = 0.3)$ that someone will take the taxi. Each stop is independent of the others.

4. What is the variance of the number of stops required?

$$\text{Var}(X) = \frac{r(1-p)}{p^2} = \frac{5(0.7)}{0.3^2} = \frac{3.5}{0.09} \approx 38.89$$

# Consider..

- A random variable that counts the number of "events" that occur within certain specified boundaries/interval.
- Example
  1. The number of defects in an item
  2. The number of emails coming in an hour
  3. The number of wrong telephone numbers that are dialed in a day
  4. The number of misprints on a page (or a group of pages) of a book
  5. The number of customers entering a post office on a given day
  6. The number of transistors that fail on their first day of use
  7. The number of people in a community living to 100 years of age

# Definition of Poisson Random Variable

- A **Poisson random variable** models the number of events occurring in a fixed interval of **time or space**, when:
  1. Events occur **independently.**
  2. The **average rate of occurrence** is constant ($\lambda > 0$).
  3. Two events cannot occur at exactly the same instant (no "clumping").
- Formally,

$$X \sim Poisson(\lambda)$$

means that $X$ be the total number of events in the interval, with PMF:

$$P(X = k) = \frac{e^{-\lambda}\lambda^{k}}{k!}, \qquad k = 0,1,2,\ldots$$

# Another Definition of Poisson Random Variable

- We also can define $\alpha = \lambda t$, where
  - $t$ : Length of interval
  - $\lambda$ : Average number (rate) of events occurring per unit of interval $t$

$$X \sim Poi(\alpha) \qquad \alpha = \lambda t$$

- Thus, the PMF can be also written as

$$P(X = k) = \frac{e^{-\alpha}\alpha^k}{k!} = \frac{e^{-\lambda}\lambda^k}{k!}, \qquad x = 0,\ 1,\ 2,\ \ldots$$

# Properties of Poisson Random Variable $\boldsymbol{X \sim Poi(\lambda)}$

PMF

$$P(X = k) = \frac{e^{-\lambda}\lambda^{k}}{k!}, \qquad x = 0,\ 1,\ 2,\ \ldots$$

CDF

$$F_X(k) = P(X \leq k) = \begin{cases} 0 & x < 0 \\ \displaystyle\sum_{j=0}^{k} \frac{e^{-\lambda}\lambda^{j}}{j!} & x \geq 0 \end{cases}$$

Expectation / Mean

$$E[X] = \lambda$$

Variance

$$Var[X] = \lambda$$

# Poisson RV vs. Binomial RV

## Binomial

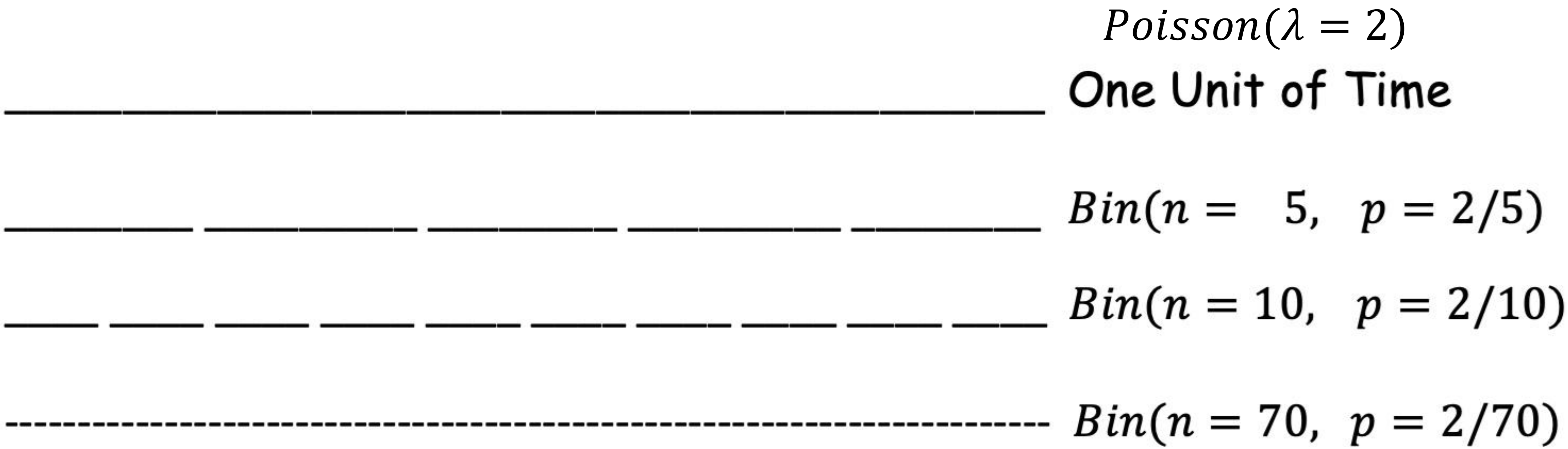
1. Discrete/whole numbers
2. A certain number of **opportunities** for the occurrence are given. You flip a coin certain times and ask how many times Heads appeared.
   “Opportunity = flipping coin”
3. Each trial/“flip” is independent of the others

## Poisson

1. Discrete/whole numbers, but **Infinite**
2. **No special opportunities** for the events. An accident may happen at any times. No special opportunity when an accident happens or not.
3. An occurrence happening at one time interval is independent of another occurrence happening at another time interval

# Poisson RV vs. Binomial RV

- Let’s say we want to model babies born in the next minute, if the historical average rate is 2 babies/min. That is $\lambda = 2$ babies/min.



$Poisson(\lambda = 2)$

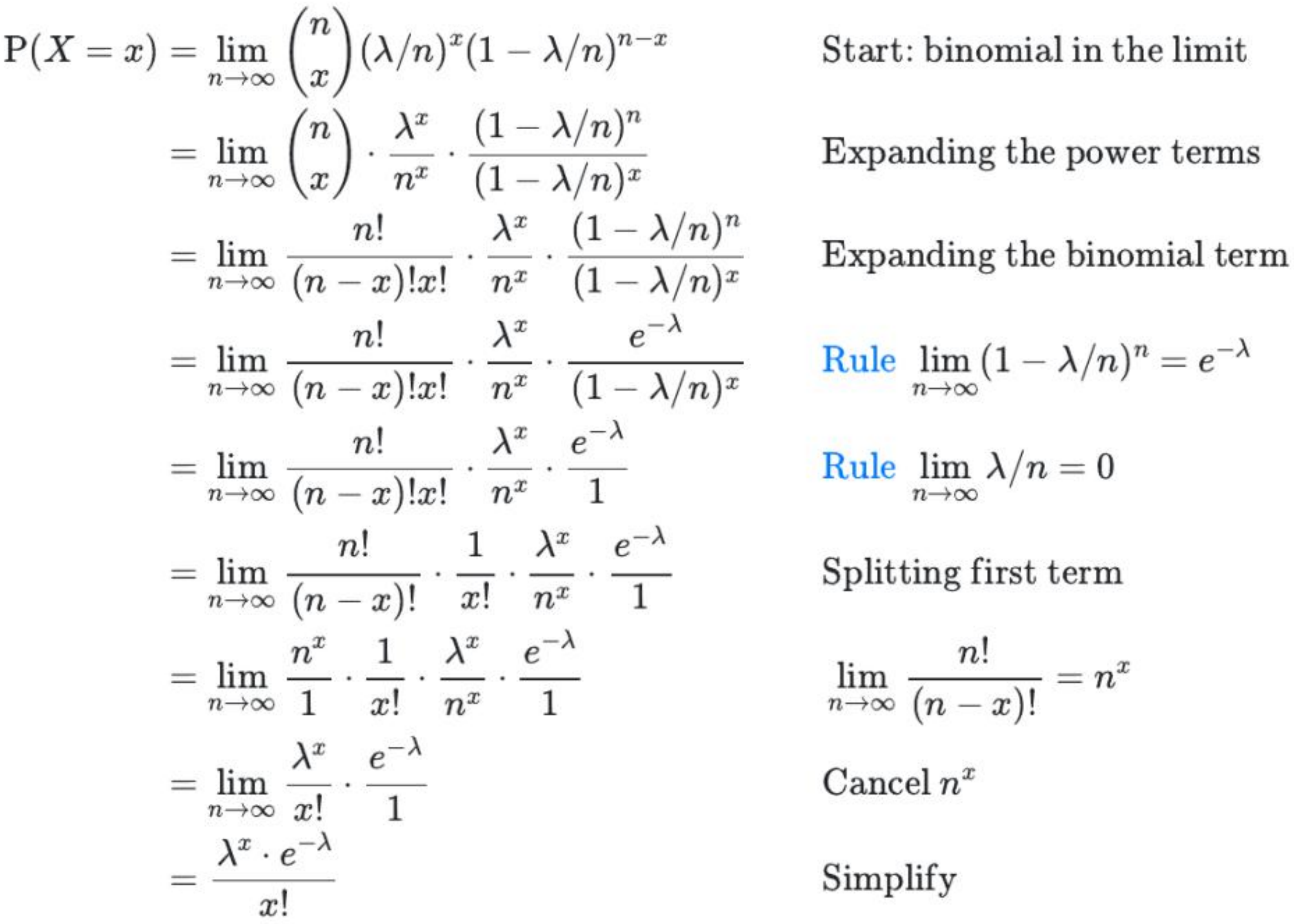
One Unit of Time

$Bin(n = \ 5, \ \ p = 2/5)$

$Bin(n = 10, \ \ p = 2/10)$

$Bin(n = 70, \ p = 2/70)$

# Poisson RVs



$$
\begin{aligned}
\mathrm{P}(X=x) &= \lim_{n\to\infty}\binom{n}{x}(\lambda/n)^x(1-\lambda/n)^{n-x} && \text{Start: binomial in the limit}\\
&= \lim_{n\to\infty}\binom{n}{x}\cdot\frac{\lambda^x}{n^x}\cdot\frac{(1-\lambda/n)^n}{(1-\lambda/n)^x} && \text{Expanding the power terms}\\
&= \lim_{n\to\infty}\frac{n!}{(n-x)!x!}\cdot\frac{\lambda^x}{n^x}\cdot\frac{(1-\lambda/n)^n}{(1-\lambda/n)^x} && \text{Expanding the binomial term}\\
&= \lim_{n\to\infty}\frac{n!}{(n-x)!x!}\cdot\frac{\lambda^x}{n^x}\cdot\frac{e^{-\lambda}}{(1-\lambda/n)^x} && \text{Rule } \lim_{n\to\infty}(1-\lambda/n)^n = e^{-\lambda}\\
&= \lim_{n\to\infty}\frac{n!}{(n-x)!x!}\cdot\frac{\lambda^x}{n^x}\cdot\frac{e^{-\lambda}}{1} && \text{Rule } \lim_{n\to\infty}\lambda/n = 0\\
&= \lim_{n\to\infty}\frac{n!}{(n-x)!}\cdot\frac{1}{x!}\cdot\frac{\lambda^x}{n^x}\cdot\frac{e^{-\lambda}}{1} && \text{Splitting first term}\\
&= \lim_{n\to\infty}\frac{n^x}{1}\cdot\frac{1}{x!}\cdot\frac{\lambda^x}{n^x}\cdot\frac{e^{-\lambda}}{1} && \lim_{n\to\infty}\frac{n!}{(n-x)!} = n^x\\
&= \lim_{n\to\infty}\frac{\lambda^x}{x!}\cdot\frac{e^{-\lambda}}{1} && \text{Cancel } n^x\\
&= \frac{\lambda^x\cdot e^{-\lambda}}{x!} && \text{Simplify}
\end{aligned}
$$

# Example 1

- Suppose that the average number of accidents occurring weekly on a particular stretch of a highway equals 3. Calculate the probability that there is at least one accident this week!
- Let, X = the number of accidents during this week

# Example 1

- Suppose that the average number of accidents occurring weekly on a particular stretch of a highway equals 3. Calculate the probability that there is at least one accident this week!
- Let, X = the number of accidents during this week
- t = 1 week
- We can assume that X is a Poisson random variable. The average number of accidents per week is $3$ $(\lambda = 3)$. so, we can write $X \sim Poi(\lambda t) \sim Poi(3)$.

# Example 1

- Suppose that the average number of accidents occurring weekly on a particular stretch of a highway equals 3. Calculate the probability that there is at least one accident this week!
- Let, X = the number of accidents during this week
- t = 1 week
- We can assume that X is a Poisson random variable. The average number of accidents per week is 3 $(\lambda = 3)$. so, we can write $X \sim Poi(\lambda t) \sim Poi(3)$.

$$\alpha = \lambda t = 3(1) = 3$$

$$P(X \geq 1) = 1 - P(X = 0)$$

$$= 1 - e^{-3}\frac{3^0}{0!}$$

$$\approx 0.9502$$

# Example 2

- A quality inspector at a glass manufacturing company inspects sheets of glass to check for any slight imperfections. Suppose that the number of these flaws X in a sheet of glass has a Poisson distribution with λ = 0.5 flaws per sheet.
- Compute

1. The probability that there is no flaw in a sheet ?
2. The probability that there are two or more flaws in a sheet

# Example 2

- A quality inspector at a glass manufacturing company inspects sheets of glass to check for any slight imperfections. Suppose that the number of these flaws X in a sheet of glass has a Poisson distribution with λ = 0.5 flaws per sheet.
- Compute

1. The probability that there is no flaw in a sheet ?

- λ = 0.5 implies that the expected number of flaws per sheet is 0.5.
- t = 1 sheet

$$\alpha = \lambda t = 0.5$$

$$X \sim P(0.5)$$

$$P(X = 0) = \frac{e^{-0.5} \times (0.5)^0}{0!} = e^{-0.5} = 0.607$$

# Example 2

- A quality inspector at a glass manufacturing company inspects sheets of glass to check for any slight imperfections. Suppose that the number of these flaws X in a sheet of glass has a Poisson distribution with λ = 0.5 flaws per sheet.
- Compute

2. The probability that there are two or more flaws in a sheet

- λ = 0.5 implies that the expected number of flaws per sheet is 0.5.
- t = 1 sheet

$$\alpha = \lambda t = 0.5$$
$$X \sim P(0.5)$$

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1)$$
$$= 1 - \frac{e^{-0.5} \times (0.5)^0}{0!} - \frac{e^{-0.5} \times (0.5)^1}{1!} = 0.090$$

# Exercise

- The arrival of a customer in a store has a Poisson distribution with 5 visitors per hour.
- Compute

1. The probability that visitors coming less than 5 in one hour is ?
2. The number of visitors =10 in a day ?

# Exercise

- The arrival of a customer in a store has a Poisson distribution with 5 visitors per hour.
- Compute

1. The probability that visitors coming less than 5 in one hour is ?

Let **X** : R.V. that describes the number of visitors coming in **one hour**

$$\lambda = 5, t = 1 \text{ hour}, \alpha = \lambda t = 5,\ X \sim P(5)$$

$$P[X<5] = \frac{5^0 e^{-5}}{0!} + \frac{5^1 e^{-5}}{1!} + \frac{5^2 e^{-5}}{2!} + \frac{5^3 e^{-5}}{3!} + \frac{5^4 e^{-5}}{4!}$$

$$= e^{-5}(1+5+25/2+125/6+625/24)$$

$$= 0.440493$$

# Exercise

- The arrival of a customer in a store has a Poisson distribution with 5 visitors per hour.
- Compute

2. The probability that the number of visitors =10 in a day ?

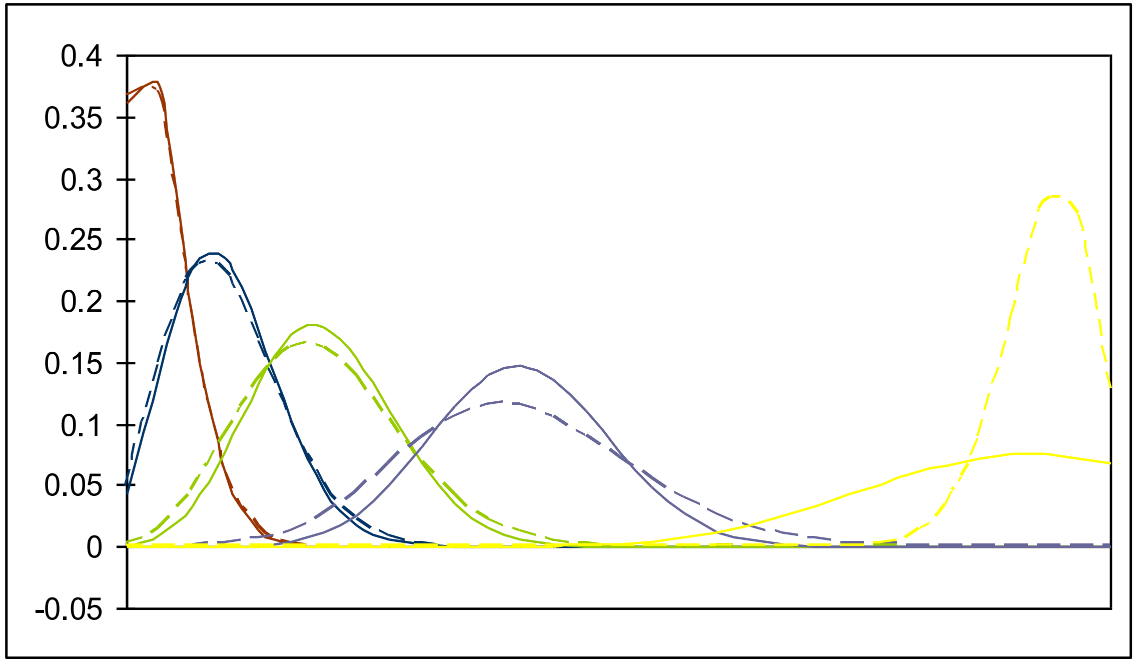
Let **Y** : R.V. that describes the number of visitors coming in **a day**

$$\lambda = 5, t = 24 \text{ hour}, \alpha = \lambda t = 120, Y \sim P(120)$$

$$P[Y = 10] = \frac{120^{10} e^{-120}}{10!}$$

# Approximation for a Binomial R.V.

- Binomial PMF (full line) and Poisson PMF (dashed line ) for $n = 30$; each color for $np = \alpha = 1, 3, 6, 12,$ and $28$.



- Both function graphs are nearly the same for $\alpha = \lambda t \leq 3$, while for other, the bigger the value of α the lower the Poisson pmf graphs compared to Binomial graphs.

# Proof

- Let, X ~ B(n, p) and α = np.

$$P(X=i)=\frac{n!}{(n-i)!i!}p^i(1-p)^{n-i}$$

$$=\frac{n!}{(n-i)!i!}\left(\frac{\alpha}{n}\right)^i\left(1-\frac{\alpha}{n}\right)^{n-i}=\frac{n(n-1)\ldots(n-i+1)}{n^i}\frac{\alpha^i}{i!}\frac{(1-\alpha/n)^n}{(1-\alpha/n)^i}$$

- Now, for n large and p small,

$$\left(1-\frac{\alpha}{n}\right)^n\approx e^{-\alpha},\quad \frac{n(n-1)\ldots(n-i+1)}{n^i}\approx 1,\quad \left(1-\frac{\alpha}{n}\right)^i\approx 1$$

- So

$$P(X=i)\approx e^{-\lambda t}\frac{\lambda t^i}{i!}$$

# Approximation for a Binomial R.V.

- Poisson random variable can be used as an approximation for a binomial random variable with parameters $(n, p)$ when $n$ is large and $p$ is small.
- When we approximate using Poisson R.V., we use

$$\alpha = \lambda t = np$$

in some literatures, it is recommended for $n \geq 30$ and $np \leq 3$. ***

# Exercise Approximation for a Binomial R.V.

- Suppose the probability that an item produced by a certain machine will be defective is 0.1.
- Find the probability that a sample of 10 items will contain at most one defective item !
- Assume that the quality of successive items is independent.

# Exercise

- Suppose the probability that an item produced by a certain machine will be defective is 0.1.
- Find the probability that a sample of 10 items will contain at most one defective item !
- Assume that the quality of successive items is independent.
- $X$ : the number of defective item on the sample
- $X \sim Bin(10, 0.1)$, so

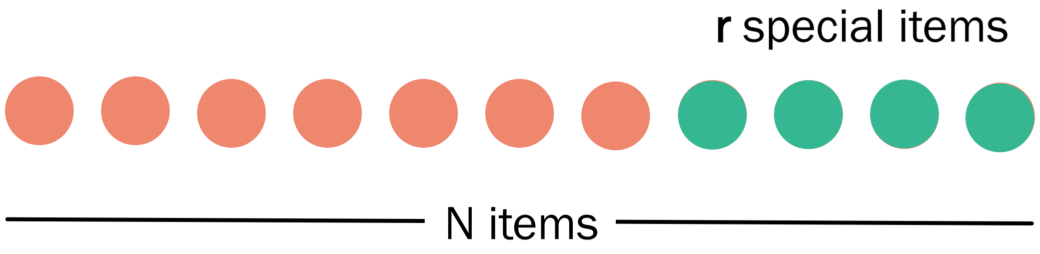
$$P(X \le 1) = \binom{10}{0}(0.1)^0(0.9)^{10} + \binom{10}{1}(0.1)^1(0.9)^9 = 0.7361$$

- Whereas, we can also approximate using Poisson approximation

$$P(X \le 1) \approx e^{-1}\frac{1^0}{0!} + e^{-1}\frac{1^1}{1!} \approx 0.7358$$

# Consider...

- A collection of **N** items of which **r** are of a certain/special kind or “successes”.



- If one of the **N** items is chosen at random, the probability that it is a special kind is:

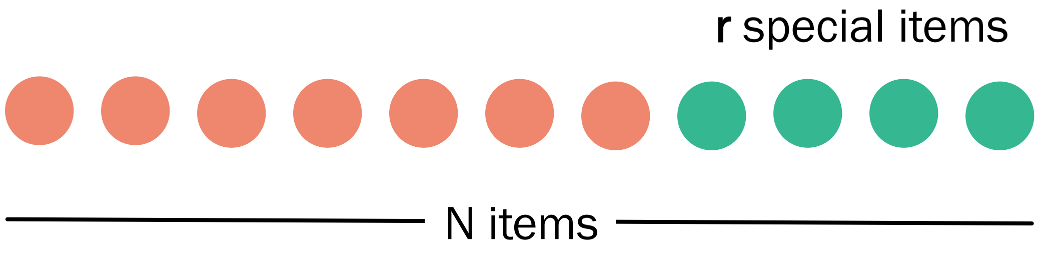
$$p = \frac{r}{N}$$

- If **n** items are chosen at random **with replacement**, it is clear that **X**, the number of special items chosen, is the **Binomial random variable**:

$$X \sim Bin(n, \frac{r}{N})$$

# Consider...

- A collection of **N** items of which **r** are of a certain/special kind or “successes”.



- If one of the **N** items is chosen at random, the probability that it is a special kind is:

$$p = \frac{r}{N}$$

- However, if **n** items are chosen at random **without replacement**, then **X**, the number of special items chosen, is the **Hypergeometric random variable.**

$$X \sim Hypergeometric(N,K,n)$$

# Definition of Hypergeometric Random Variables

- A **Hypergeometric random variable** models the number of "successes" or "special items" in a sample drawn **without replacement** from a finite population.
- Formally:
  - A population of size $N$ contains $K$ "successes"/"special items" and $N - K$ "failures"/"ordinary".
  - You draw $n$ items without replacement.
  - Let $X$ be the number of "successes"/"special items" in your sample (i.e., that we get).
- Then

$$X \sim \textit{Hypergeometric}(N,K,n)$$

- With PMF

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}, \qquad \max(0, n - (N - K)) \leq k \leq \min(n, K)$$

# Properties of Hypergeometric RVs

$X \sim HG(n,r,N)$

PMF

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}, \qquad \max(0, n - (N - K)) \le k \le \min(n, K)$$

CDF

$$F_X(t) = P(X \le t) = \sum_{k=\max(0,n-(N-K))}^{t} \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

Expectation / Mean

$$E[X] = n \cdot \frac{K}{N}$$

Variance

$$Var(X) = n \cdot \frac{K}{N} \cdot \left(1 - \frac{K}{N}\right) \cdot \frac{N - n}{N - 1}$$

# Example 1

- From a box containing 10 ping pong balls, 4 balls are drawn at random. Among the 10 balls, there are 3 red balls and 7 white balls. Determine the probability that among the 4 balls drawn, there is at most 1 red ball!
- This question is related to Hypergeometric R.V. with x=0 or x=1. Given the information that N=10, n=4, and r=3.
- X = the number of red balls taken

# Example 1

- From a box containing 10 ping pong balls, 4 balls are drawn at random. Among the 10 balls, there are 3 red balls and 7 white balls. Determine the probability that among the 4 balls drawn, there is at most 1 red ball!
- This question is related to Hypergeometric R.V. with x=0 or x=1. Given the information that N=10, n=4, and r=3.
- X = the number of red balls taken

$$P(X \leq 1) = P(X=0) + P(X=1) = \frac{\binom{3}{0} \times \binom{7}{4}}{\binom{10}{4}} + \frac{\binom{3}{1} \times \binom{7}{3}}{\binom{10}{4}} = \frac{2}{3}$$

# Binomial Approximation of Hypergeometric RVs

- If the population size N is much bigger than the number of items taken n, then Binomial R.V. will be a reasonably good approximation for hypergeometric R.V.
- Let *p* = r / N,

$$E[X] = n\frac{r}{N} = np$$

$$Var(X) = \frac{nr(N-n)(N-r)}{N^2(N-1)} = np(1-p)\frac{N-n}{N-1}$$

When, N goes to infinity, then

$$Var(X) = np(1-p)$$

# Binomial Approximation of Hypergeometric RVs (2)

- From the previous example, if N = 100, r = 30, and n = 4.
  - Since N >> n ! We can approximate using binomial R.V.

$$P(X \leq 1) = \binom{n}{0} p^0 (1-p)^n + \binom{n}{1} p^1 (1-p)^{n-1}$$

$$= \binom{4}{0} (0.3)^0 (1-0.3)^4 + \binom{4}{1} (0.3)^1 (1-0.3)^3 = 0.6517$$

  - If you use hypergeometric R.V., you will get 0.6516 ! Not much difference

# Example 2

- A small lake contains 50 fish. One day, a fisherman catches 10 of these fish and tags them so that they can be recognized if they are caught again. The tagged fish are released back into the lake. The next day, the fisherman goes out and catches 8 fish, which are kept in the fishing boat until they are all released at the end of the day.
- What is the distribution of X, the number of tagged fish caught on the second day?
- Variable **X** is a hypergeometric R.V. with **N = 50**, **r = 10**, and **n = 8**.

# Example 2

- A small lake contains 50 fish. One day, a fisherman catches 10 of these fish and tags them so that they can be recognized if they are caught again. The tagged fish are released back into the lake. The next day, the fisherman goes out and catches 8 fish, which are kept in the fishing boat until they are all released at the end of the day.
- What is the distribution of X, the number of tagged fish caught on the second day?
- Variable **X** is a hypergeometric R.V. with **N = 50**, **r = 10**, and **n = 8**.
- Probability that 3 tagged fish are caught on the second day:

$$P(X=3)=\frac{\binom{10}{3}\times\binom{40}{5}}{\binom{50}{8}}=0.147$$

# Example 2

- A small lake contains 50 fish. One day, a fisherman catches 10 of these fish and tags them so that they can be recognized if they are caught again. The tagged fish are released back into the lake. The next day, the fisherman goes out and catches 8 fish, which are kept in the fishing boat until they are all released at the end of the day.
- What is the distribution of X, the number of tagged fish caught on the second day?
- Variable **X** is a hypergeometric R.V. with **N = 50**, **r = 10**, and **n = 8**.
- The expected number of tagged fish recaptured:

$$E[X] = \frac{nr}{N} = \frac{8 \times 10}{50} = 1.6$$