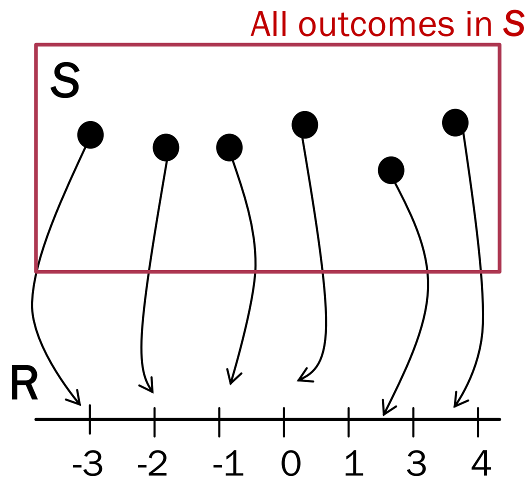
Special Discrete Random Variables

CSGE602013 –STATISTICS AND PROBABILITY

FACULTY OF COMPUTER SCIENCE UNIVERSITAS INDONESIA

Random Variables

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



Describing the Probabilities of RVs

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

$$p(x_i) = P(X = x_i)$$

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

- Cumulative Distribution Function, $F_X(x)$ 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”

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 $a \le X \le b$

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

“the probability that the value of X is between a and b”

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
Hyper-geometric Random Variables

05
Poisson Random Variables

Consider…

- Let X be the random variable with only two possible values.
- Outcomes
  - X = 1 when the outcome is a “success”
  - X = 0 when the outcome is a “failure”

Either yes or no… Either this or that..

- Equivalently:
  - The outcome of a coin toss {Head, Tail}
  - Whether a valve is open or shut
  - Whether an item is defective or not

Bernoulli Random Variables

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

$$P(X=1)=p \qquad P(X=0)=1-p$$

or

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

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

- If X is a Bernoulli R.V. (variable that has Bernoulli distribution) with parameter p, we can write

$$X \sim Ber(p)$$

Not **really** useful alone, **but** serves as a basis for other important RVs

Bernoulli Random Variables (2)

Expectation / Mean	$\mu = E[X] = 1.P(X=1) + 0.P(X=0) = p$
Variance	$\sigma^2 = Var(X)$ $= E[X^2] - (E[X])^2$ $= 1^2.P(X=1) + 0^2.P(X=0) - (p)^2$ $= p(1-p)$
CDF	$F_X(x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$

Consider...

- Repeat a Bernoulli trial 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
  - the number of heads in n coin flips
  - the number of disk drives that crashed in a cluster of 1000 computers
  - the number of advertisements on a webpage that are clicked by 200 visitors

Binomial Random Variables

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

$$\boldsymbol{X} = \boldsymbol{X}_1 + \boldsymbol{X}_2 + \cdots + \boldsymbol{X}_n,$$

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

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

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

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

Binomial Random Variables (2)

Expectation / Mean

$$E[X] = E[X_1 + X_2 + \ldots + X_n]$$
$$= E[X_1] + \ldots + E[X_n] = np$$

Variance

$$Var(X) = Var(X_1 + X_2 + \ldots + X_n)$$
$$= Var(X_1) + \ldots + Var(X_n) = np(1-p)$$

PMF

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, 2, \ldots, n$$

Example

- 1000 people are polled in a survey and asked if they are able to drive
  - The responses are YES or NO
  - The probability of a person saying YES is 0.8
- If X represents the number of YESes that occur, then X is a binomial random variable with parameters:
  - $n = 1000$
  - $p = 0.8$

$$X \sim Bin(1000, p)$$

These are not Binomial RV – Why?

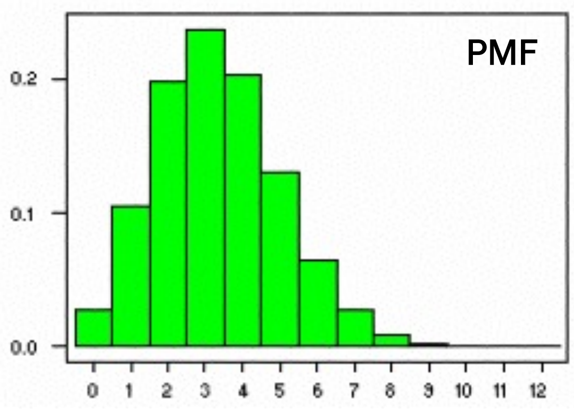
- You stand in front of FASILKOM. Let, X = number of students passing by in the next 5 minutes
- Gather a random sample of 5 men and 5 women. Let, X = number of persons out of 10 who are more than 170 cm tall
- Draw 4 cards (without replacement) from a deck of 52 cards. Let, X = number of aces among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

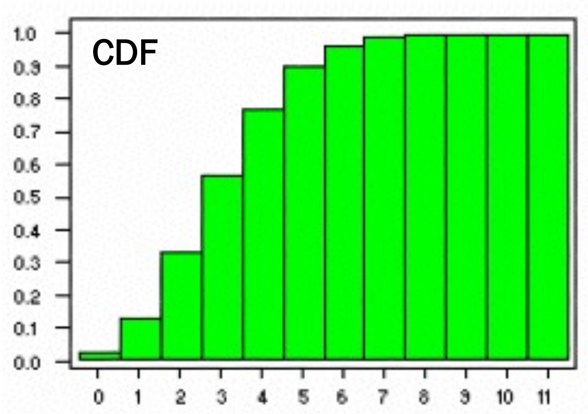
Example

- **CDF** of Binomial R.V.

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

- PMF & CDF of $\boldsymbol{Bin}(\mathbf{20}, \frac{\mathbf{1}}{\mathbf{6}})$:





Example

- **CDF** of Binomial R.V.

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

- Given $X \sim Bin(8, 0.5)$
- $P\ (X\ =\ 3)$?

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

Example

- **CDF** of Binomial R.V.

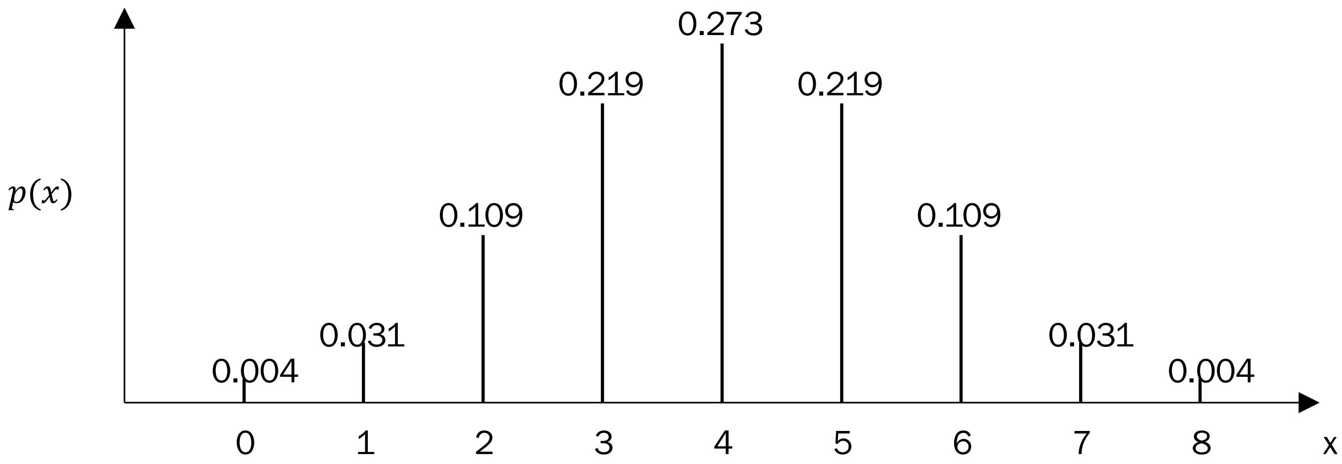
$$F_X(i) = P(X \le i) = \begin{cases} 0 & i < 0 \\ \sum_{k=0}^{\lfloor i \rfloor} \binom{n}{k} p^k (1-p)^{n-k}, & 0 \le i < n \\ 1 & i \ge n \end{cases}$$

- Given $X \sim Bin(8, 0.5)$
- P (X ≤ 1) = ?

$$P(X \le 1) = P(X = 0) + P(X = 1)$$
$$= \binom{8}{0}(0.5)^0(1-0.5)^8 + \binom{8}{1}(0.5)^1(1-0.5)^7$$
$$= 0.035$$

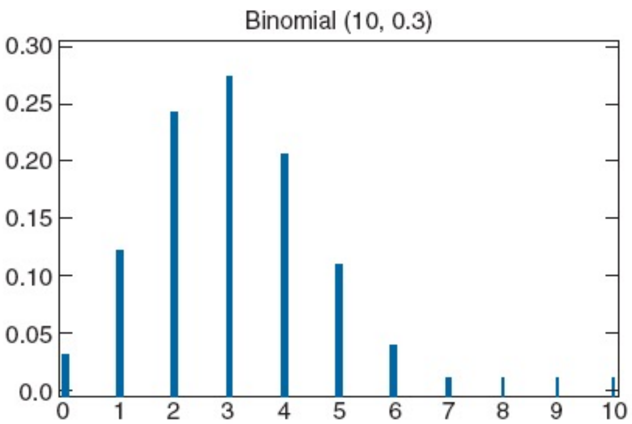
Example

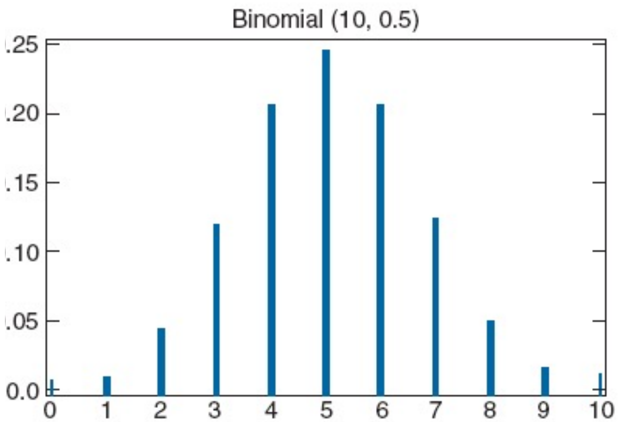
- Given $X \sim Bin(8, 0.5)$

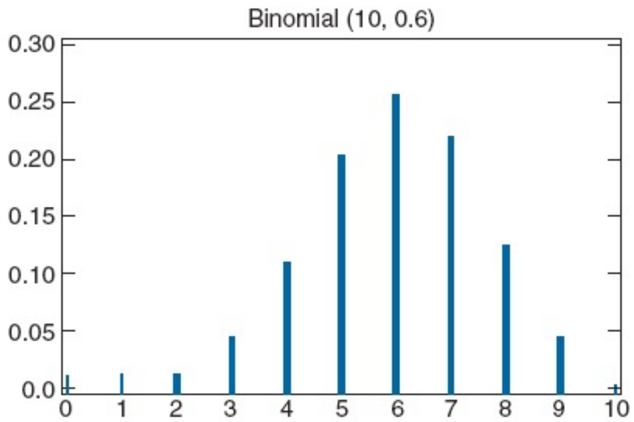


x	0	1	2	3	4	5	6	7	8
$P(X \leq x)$	0.004	0.035	0.144	0.363	0.636	0.855	0.965	0.996	1.0000

Other Examples







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 !

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

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

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

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 ?
- Let X = the number of defect disc in one pack

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 ?
- Let X = the number of defect disc in one pack

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

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

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 Y = the number of packages that the person will have to return when he buys three packs.

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

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

Consider…

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

- Examples
  - Toss a coin repeatedly. Let X = number of tosses to first head
  - 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

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.

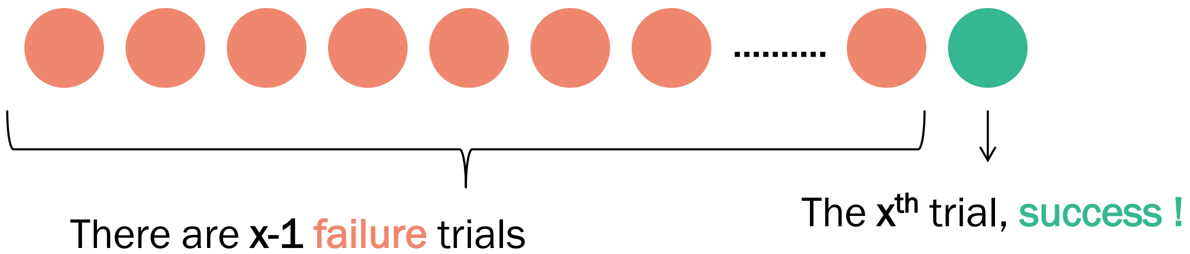
$$X \sim Geo(p)$$

- The PMF of **X** is

$$P(X = x) = (1 - p)^{x-1} p$$
$$x = 1,2,3,4,\ldots$$

Geometric Random Variables (2)

- The number of trials until the first success, with the probability of success in each trial p.
- **We also count the first success**



- The probability until the **first success** is $(1-p)(1-p)\ldots(1-p)p$
- Hence

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

Geometric Random Variables (3)

Expectation / Mean

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

Variance

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

CDF

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

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

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?

$$P(X \geq 3) = 1 - P(X = 1) - P(X = 2)$$
$$= 1 - 0.4^0 0.6 \quad - 0.4^1 0.6$$
$$= 1 - 0.6 - 0.24 = 1 - 0.84 = 0.16$$

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

- X : the number of trials until the engines start

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

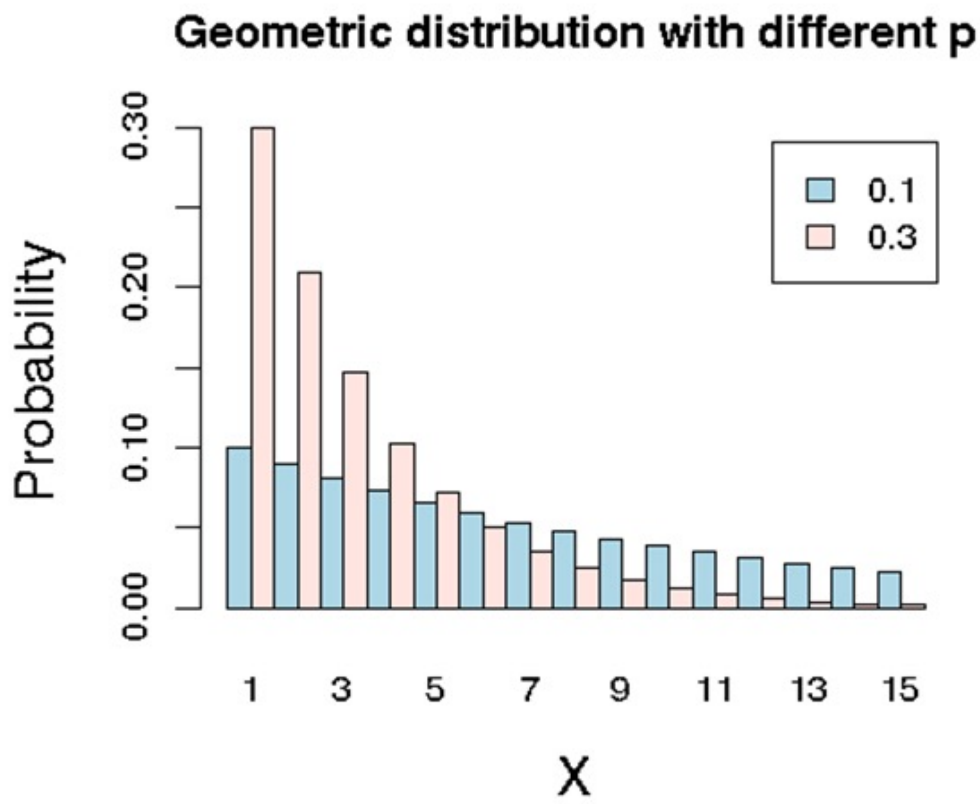
Example 2

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

- X : the number of trials until the engines start

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

Geometric Random Variable



Memoryless Property

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

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

- 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).
  - 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 $(m+r)^{th}$ experiment with this prior condition.

Consider...

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

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

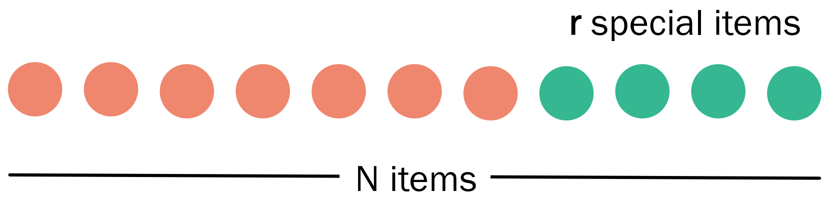
$$p = \frac{r}{N}$$

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

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

Consider...

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



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

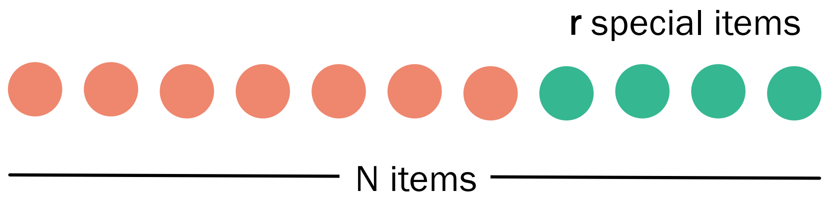
$$p = \frac{r}{N}$$

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

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

Hypergeometric Random Variables

- From N items,



- if n items are chosen at random without replacement, then X, the number of special items chosen, is the hypergeometric R. V.

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

Hypergeometric Random Variables (3)

Expectation / Mean

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

Variance

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

PMF

$$P(X = x) = \frac{\binom{r}{x} \times \binom{N-r}{n-x}}{\binom{N}{n}}$$

for $\max\{0, n+r-N\} \leq x \leq \min\{n, r\}$

Example 1

- From inside a box containing 10 ping pong balls, 4 balls are taken randomly. Among those 10 balls, there are 3 red and 7 white balls. Find the probability of the 4 balls have been taken, there is 1 red ball at the most !
- 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 inside a box containing 10 ping pong balls, 4 balls are taken randomly. Among those 10 balls, there are 3 red and 7 white balls. Find the probability of the 4 balls have been taken, there is 1 red ball at the most !
- 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

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

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:

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

Consider..

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

Poisson Random Variable

- A random variable X distributed as a **Poisson** random variable with parameter **α**, which is written

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

- The PMF

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

X : Random variable that describes the number of events/arrivals within interval *t*.
λ : Average number of events occurring per unit of interval (rate)
t : Length of interval
α : Average number of events occurring in any interval of length *t*.

Poisson Random Variable (2)

Expectation / Mean

Variance

$$E[X] = Var[X] = \alpha = \lambda t$$

CDF

$$F_X(x) = P(X \le x) = \begin{cases} 0 & i < 0 \\ \sum_{i=0}^{\lfloor x \rfloor} e^{-\alpha} \dfrac{\alpha^i}{i!} & i \ge 0 \end{cases}$$

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 $\lambda = 0.5$ 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**

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

$$\begin{aligned} P[X < 5] &= \frac{5^0 e^{-5}}{0!} + \frac{5^1 e^{-5}}{1!} + \frac{5^2 e^{-5}}{2!} + \frac{5^3 e^{-5}}{3!} + \frac{5^4 e^{-5}}{4!} \\ &= e^{-5}(1 + 5 + 25/2 + 125/6 + 625/24) \\ &= 0.440493 \end{aligned}$$

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 ?

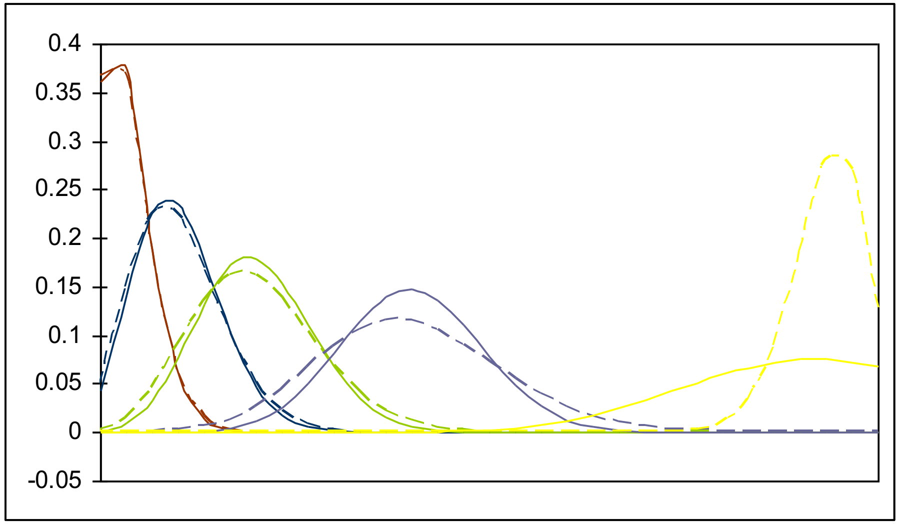
Let **Y** : R.V. that describes the number of visitors coming in **a day**

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

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

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.

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

- 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

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

- Whereas, we can also approximate using Poisson approximation

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