5. DISCRETE RANDOM VARIABLES

Introduction

- We introduced random events and probability in the previous chapter.

- Now we will formalize the random events mathematically.

- This chapter focuses on discrete random variables, while the next chapter will focus on continuous random variables.

Random variables

Random variables

Universiti Kebangsaan Malaysia
The National University of Malaysia

- A random variable is a variable whose values are determined by chance.
  - In other words, random variable is the random outcome of an experiment.

- Discrete random variable – random variable that assumes countable values.
  - Example:
    - The number of students getting an A
    - The number of fish caught on a fishing trip

- We usually denote a random variable by X (uppercase letter), and x (lowercase letter) for a specific value of the random variable.

Probability distribution

Universiti Kebangsaan Malaysia
The National University of Malaysia

- The probability distribution is a table or function, that lists all the possible values for a random variable and their corresponding probabilities.
- Also called probability mass function (pmf) for discrete random variables.
- Notation: $P(x)$ or $P(X = x)$ is the probability that the random variable X takes the value x.
- Two properties of probability distribution of discrete random variable:
  - $0 \leq P(X = x) \leq 1$ Probability must be between 0 and 1
  - $\sum_x P(X = x) = 1$ Summation of all probability must equal to 1

Example

- Consider rolling a die.
  - Let X be the observed outcome when rolling a die.
  - The following table is the probability distribution of X:

x	1	2	3	4	5	6
$P(X = x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

  - $P(X = 1) = 1/6$, means that the probability of X takes value 1 is $1/6$
  - Also note that

$$\sum P(X = x) = P(1) + P(2) + \cdots + P(6) = 1$$

Exercise

- Determine whether the following represents a probability distribution. If it does not, state why.

7.

X	15	16	20	25
$P(X)$	0.2	0.5	0.7	−0.8

8.

X	5	7	9
$P(X)$	0.6	0.8	−0.4

9.

X	−5	−3	0	2	4
$P(X)$	0.1	0.3	0.2	0.3	0.1

10.

X	20	30	40	50
$P(X)$	0.05	0.35	0.4	0.2

11.

X	3	6	9	1
$P(X)$	0.3	0.4	0.3	0.1

12.

X	3	7	9	12	14
$P(X)$	$\frac{4}{13}$	$\frac{1}{13}$	$\frac{3}{13}$	$\frac{1}{13}$	$\frac{2}{13}$

Two properties:

1) $0 \leq P(X=x) \leq 1$

2) $\sum_{x} P(X=x) = 1$

7) $P(X=25) = -0.8 < 0$.

Not a probability distribution

8) $P(X=9) = -0.4 < 0$

Not a probability distribution

9) $0 \leq P(X=x) \leq 1$

$$\sum P(X=x) = 0.1+0.3+0.2+0.3+0.1 = 1$$

Yes, it is a probability distribution

10) $0 \leq P(X=x) \leq 1$

$$\sum P(X=x) = 0.05+0.35+0.4+0.2 = 1$$

Yes, it is a probability distribution

11) $$\sum P(X=x) = 0.3+0.4+0.3+0.1 = 1.1 \neq 1$$

Not a probability distribution

12) $$\sum P(X=x) = \frac{4}{13}+\frac{1}{13}+\frac{3}{13}+\frac{1}{13}+\frac{2}{13} = \frac{11}{13} \neq 1$$

Not a probability distribution.

UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

Exercise

- An industrial psychologist administered a personality inventory test for passive-aggressive traits to 150 employees. Individuals were given a score from 1 to 5, where 1 was extremely passive and 5 extremely aggressive. A score of 3 indicated neither trait. Construct a probability distribution for the random variable X.

Score, x	Frequency, f	$P(X=x)$
1	24	$24/150 = 0.16$
2	33	0.22
3	42	0.28
4	30	0.2
5	21	0.14
Total	150	

Exercise

- At a drop-in mathematics tutoring center, each teacher sees 4 to 8 students per hour. The probability that a tutor sees 4 students in an hour is 0.117; 5 students, 0.123; 6 students, 0.295; and 7 students, 0.328.
  a) Find the probability that a tutor sees 8 students in an hour
  b) Construct the probability distribution.
  c) Find the probability that a tutor sees 6 or less students in an hour.

a) $\sum P(X=x) = 1$

$P(4)+P(5)+P(6)+P(7)+P(8) = 1$

$P(8) = 1 - P(4) - P(5) - P(6) - P(7)$

$= 1 - 0.117 - 0.123 - 0.295 - 0.328$

$= 0.137$

b)

x	4	5	6	7	8
$P(X=x)$	0.117	0.123	0.295	0.328	0.137

c) $P(X \leq 6) = P(4) + P(5) + P(6)$

$= 0.117 + 0.123 + 0.295$

$= 0.535$

$= 1 - P(7) - P(8) = 1 - P(X > 6)$

Exercise

$\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$

- The number 1, 2, 3 and 4 are printed one each on one side of card. The cards are placed face down and mixed. Choose two cards at random; and let X be the sum of the two numbers. Construct the probability distribution for this random variable X.

Possible outcomes	:	1 & 2	1 & 3	1 & 4	2 & 3	2 & 4	3 & 4
X	:	3	4	5	5	6	7
Probability	:	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

x	3	4	5	6	7
$P(X=x)$	$1/6$	$1/6$	$2/6 = 1/3$	$1/6$	$1/6$

eg: $P(X<5) = P(3)+P(4)$

$$= \frac{1}{6} + \frac{1}{6}$$

$$= \frac{1}{3}$$

cdf

Cumulative distribution function

- Another important concept in the distribution of random variable is the cumulative distribution function (also called distribution function).

cumulative distribution function = distribution function

- If X is a discrete random variable, the cumulative distribution function is given by

$$F(x) = P(X \leq x) = \sum_{t \leq x} P(X = t)$$

for $-\infty < x < \infty$.

- $F(x)$ is the probability of X taking value less than or equal to x.

Cumulative distribution function

- Properties of a discrete distribution function:
  - $0 \leq F(x) \leq 1$ for all x
  - If $a < b$, then $F(a) \leq F(b)$
  - $F(x)$ is a step function
  - $F(-\infty) = 0$ and $F(\infty) = 1$

  $P(X \leq -\infty) = 0$ $\quad$ $P(X \leq \infty) = 1$

- Additionally:
  - $P(a < X \leq b) = F(b) - F(a)$ $= P(X \leq b) - P(X \leq a)$
  - $P(X > a) = 1 - P(X \leq a) = 1 - F(a)$
  - $P(X < a) = P(X \leq a - 1) = F(a - 1)$
  - $P(X = a) = P(X \leq a) - P(X \leq a - 1) = F(a) - F(a - 1)$

Example – CDF

- Using the following probability distribution, find its cumulative distribution function $F(x)$.

x	0	1	2	3
$P(X = x)$	1/8	3/8	3/8	1/8

- From the table we get:

$$F(0) = \frac{1}{8} = P(X \leq 0) = P(0)$$

$$F(1) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2} = P(x \leq 1) = P(0) + P(1)$$

$$F(2) = \frac{1}{2} + \frac{3}{8} = \frac{7}{8} = P(0) + P(1) + P(2)$$

$$F(3) = \frac{7}{8} + \frac{1}{8} = 1 = P(0) + P(1) + P(2) + P(3)$$

Example – CDF

- Additionally, $F(x) = 0$ for $x < 0$, and $F(x) = 1$ for $x > 3$.

- Therefore:

$$F(x) = \begin{cases} 0, & x < 0 \\ 1/8, & 0 \leq x < 1 \\ 1/2, & 1 \leq x < 2 \\ 7/8, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

Exercise

- You are given the following table of probability distribution of X. Find the distribution function $F(x)$.

x	1	2	3	4	5
$P(X = x)$	0.16	0.22	0.28	0.20	0.14

$F(x) = P(X \leq x)$ 0.16 0.38 0.66 0.86 1

0.22 0.28 0.2 0.14

$$F(x) = \begin{cases} 0, & x < 1 \\ 0.16, & 1 \leq x < 2 \\ 0.38, & 2 \leq x < 3 \\ 0.66 & 3 \leq x < 4 \\ 0.86 & 4 \leq x < 5 \\ 1 & x \geq 5 \end{cases}$$

Exercise

- Given the following cumulative distribution function for a discrete random variable X:

cdf/probability function:

$$F(x) = \begin{cases} 0, & x < 1 \\ {}^{1}/_{6}, & 1 \le x < 2 \\ {}^{1}/_{2}, & 2 \le x < 3 \\ {}^{5}/_{6}, & 3 \le x < 4 \\ 1, & x \ge 4 \end{cases}$$

$P(X=1) = \frac{1}{6} - 0 = \frac{1}{6}$

$P(X=2) = \frac{1}{2} - \frac{1}{6} = \frac{1}{3}$

$P(X=3) = \frac{5}{6} - \frac{1}{2} = \frac{1}{3}$

$P(X=4) = 1 - \frac{5}{6} = \frac{1}{6}$

- Find
  a) $P(1 < X < 3)$
  b) $P(X > 2)$
  c) $P(2 \le X \le 4)$
  d) The probability distribution of X

a) $P(1<X<3) = P(1<X\leq 2) = P(X\leq 2) - P(X\leq 1)$

$$= F(2) - F(1) = \frac{1}{2} - \frac{1}{6} = \frac{1}{3}$$

b) $P(X>2) = 1 - P(X\leq 2) = 1 - F(2) = 1 - \frac{1}{2} = \frac{1}{2}$

c) $P(2\leq X\leq 4) = P(1<X\leq 4) = P(X\leq 4) - P(X\leq 1)$

$$= F(4) - F(1) = 1 - \frac{1}{6} = \frac{5}{6}$$

Probability distribution:

d)

x	1	2	3	4
$P(X=x)$	$1/6$	$1/3$	$1/3$	$1/6$

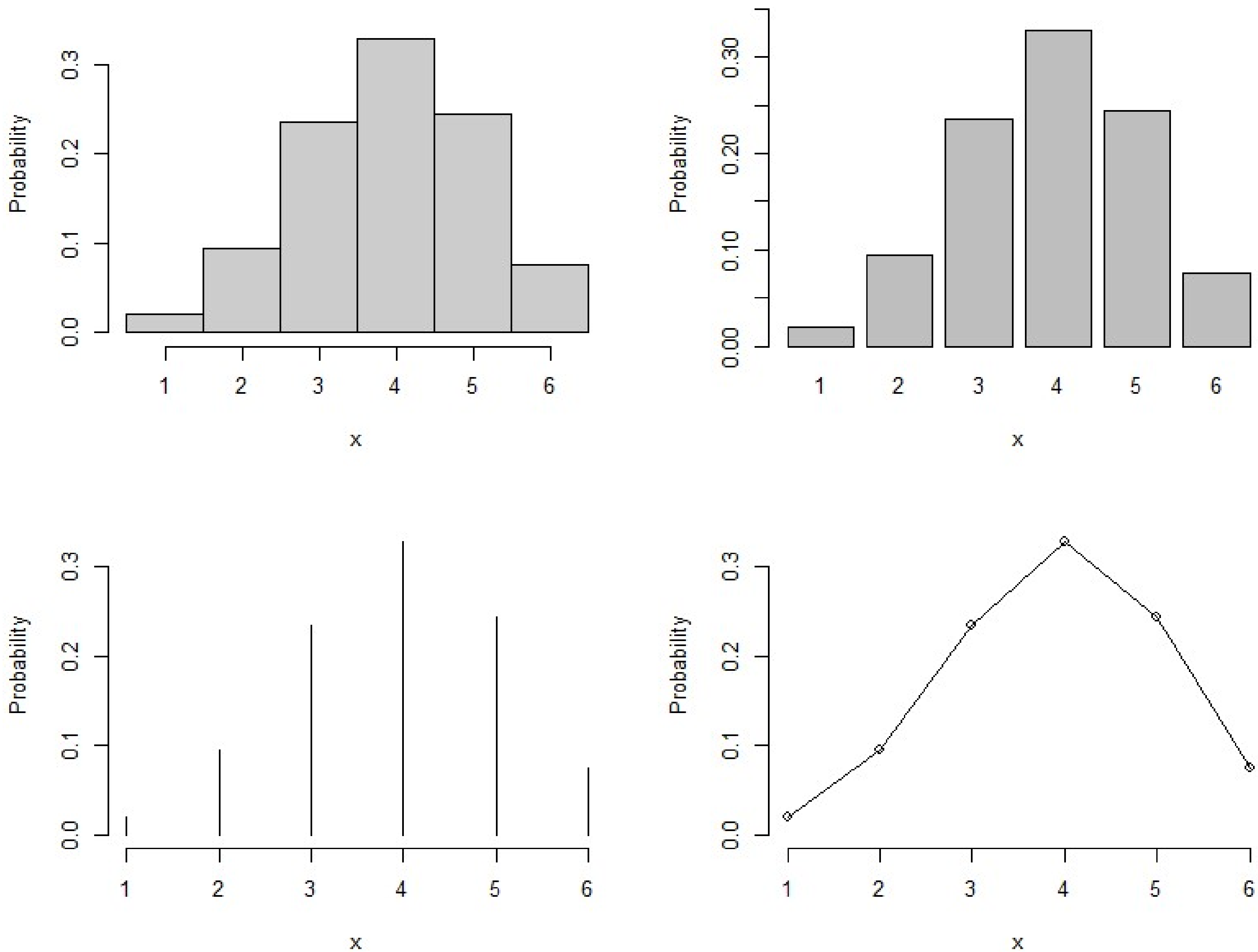
Graphing probability distribution

- The probability distribution can be shown visually using for example:
  - Histogram
  - Line plot

- The distribution function can be shown visually using:
  - Line plot

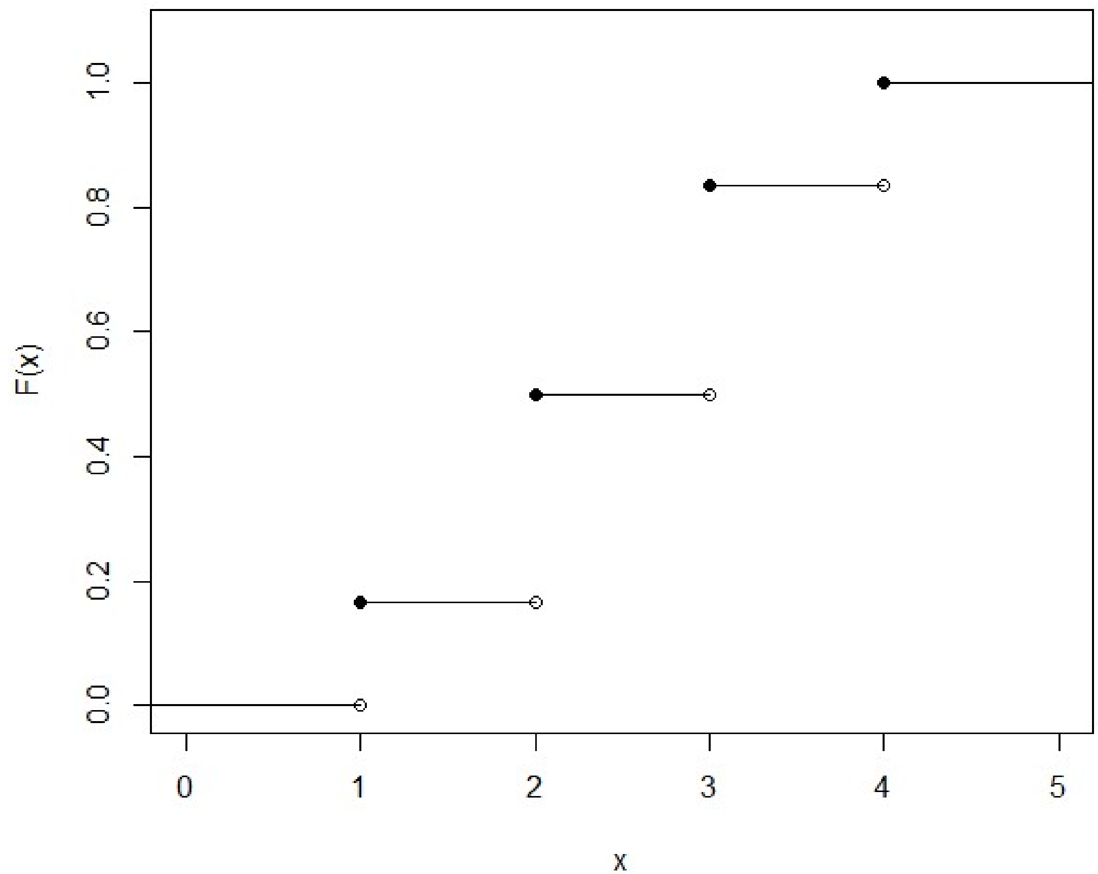
- Presenting the probability distribution visually helps with describing the distribution.

Example – graphing probability distribution



Example – graphing distribution function

$$F(x) = \begin{cases} 0, & x < 1 \\ {}^{1}/_{6}, & 1 \leq x < 2 \\ {}^{1}/_{2}, & 2 \leq x < 3 \\ {}^{5}/_{6}, & 3 \leq x < 4 \\ 1, & x \geq 4 \end{cases}$$



Mean, variance, standard deviation of discrete random variables

Mean of discrete random variables

- The mean, variance and standard deviation for random variable are calculated in a similar way to grouped data.

- We use the probability $P(X = x)$ to replace f/N (frequency over total population).

- The mean of random discrete variable is defined as:

$$\mu = \sum\left(xP(X = x)\right)$$

Variance and standard deviation of discrete random variables

- Similarly, the variance and standard deviation can be calculated using:

$$\sigma^2 = \sum(x-\mu)^2 P(X=x), \qquad \sigma = \sqrt{\sum(x-\mu)^2 P(X=x)}$$

- As before, for easier computation:

$$\sigma^2 = \left(\sum x^2 P(X=x)\right) - \mu^2, \qquad \sigma = \sqrt{\sum x^2 P(X=x) - \mu^2}$$

Example

- Three coins are tossed and let X be the number of heads that occur.

x	0	1	2	3
$P(X = x)$	$1/8$	$3/8$	$3/8$	$1/8$

$$\mu = \sum xP(X = x)$$
$$= \left(0 \times \frac{1}{8}\right) + \left(1 \times \frac{3}{8}\right) + \left(2 \times \frac{3}{8}\right) + \left(3 \times \frac{1}{8}\right)$$
$$= 1.5$$

$$\sigma^2 = \sum x^2 P(X = x) - \mu^2$$
$$= \left(0^2 \times \frac{1}{8}\right) + \left(1^2 \times \frac{3}{8}\right) + \left(2^2 \times \frac{3}{8}\right) + \left(3^2 \times \frac{1}{8}\right) - 1.5^2$$
$$= 3 - 2.25 = 0.75$$

$$\sigma = \sqrt{0.75} = 0.8660$$

Another example

- The probability distribution for the number of batteries sold over the weekend at a convenience store is given below:

x	2	4	6	8
$P(x)$	0.20	0.40	0.32	0.08

$$\begin{aligned}\mu &= \sum xP(X=x) \\ &= 2(0.2)+4(0.4)+6(0.32)+8(0.08) \\ &= 4.56\end{aligned}$$

$$\begin{aligned}\sigma^2 &= \sum x^2P(X=x)-\mu^2 \\ &= 2^2(0.2)+4^2(0.4)+6^2(0.32)+8^2(0.08)-4.56^2 \\ &= 23.84-20.7936=3.046\end{aligned}$$

$$\sigma=\sqrt{3.046}=1.745$$

Exercise

UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

5.29 Let x be the number of heads obtained in two tosses of a coin. The following table lists the probability distribution of x.

x	0	1	2
$P(x)$	.25	.50	.25

Calculate the mean and standard deviation of x. Give a brief interpretation of the value of the mean.

$$\mu = \sum x P(X=x) = (0 \times 0.25) + (1 \times 0.5) + (2 \times 0.25)$$

$$= 1$$

$$\sigma^2 = \sum x^2 P(X=x) - \mu^2$$

$$= (0^2 \times 0.25) + (1^2 \times 0.5) + (2^2 \times 0.25) - 1^2$$

$$= 0.5$$

$$\sigma = \sqrt{0.5} = 0.7071$$

If the experiment is repeated many times, the average number of heads in two tosses is 1.

Expectation

- Another concept related to the mean for a probability distribution is that of expected value or expectation

- The expected value of a random variable is the theoretical average of the variable.

eg: $E[X^n]$

$= \sum_x x^n P(X=x)$

- For discrete random variable, the formula is

$E(X) =$

$$E[X] = \sum xP(X = x) = \mu$$

$E[f(x)]$

$= \sum_x f(x) P(X=x)$

- Also note that the variance

$$\sigma^2 = E[(X - \mu)^2] = E[X^2] - E[X]^2$$

$\sum x^2 P(X=x) - \mu^2$

Exercise

- A person pays \$2 to play a certain game by rolling a single die once. If a 1 or a 2 comes up, the person wins nothing. If, however, the player rolls a 3, 4, 5, or 6, he or she wins the difference between the number rolled and \$2. Find the expectation and variance for this game. Is the game fair?

Die outcome :	1	2	3	4	5	6
Amount win :	0	0	1	2	3	4

Let X = amount win.

x	0	1	2	3	4
$P(X=x)$	$1/3$	$1/6$	$1/6$	$1/6$	$1/6$

$$E[X] = \mu = \sum x P(X=x) = \left(0 \times \frac{1}{3}\right) + \left(1 \times \frac{1}{6}\right) + \left(2 \times \frac{1}{6}\right) + \left(3 \times \frac{1}{6}\right) + \left(4 \times \frac{1}{6}\right)$$

$$= \frac{5}{3} = 1.667$$

$$\sigma^2 = \sum x^2 P(X=x) - \mu^2$$

$$= \left(0^2 \times \frac{1}{3}\right) + \left(1^2 \times \frac{1}{6}\right) + \left(2^2 \times \frac{1}{6}\right) + \left(3^2 \times \frac{1}{6}\right) + \left(4^2 \times \frac{1}{6}\right) - \left(\frac{5}{3}\right)^2$$

$$= \frac{20}{9} = 2.222$$

The game is not fair because the expected amount of winning (\$1.67) is less than the cost (\$2)

Exercise

UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

- The probability distribution for the number of batteries sold over the weekend at a convenience store is given below:

x	2	4	6	8
$P(X = x)$	0.20	0.40	0.32	0.08

- Calculate
  a) $E[X]$
  b) $E[2X]$
  c) $E[X^2]$

a) $E[X] = \sum x P(X=x)$

$= (2 \times 0.20) + (4 \times 0.40)$

$+ (6 \times 0.32) + (8 \times .08)$

$= 4.56$

b) $E[\underline{2X}] = \sum \underline{2x} P(X=x)$

$= (2 \times 2 \times 0.2) + (2 \times 4 \times 0.4) + (2 \times 6 \times 0.32)$

$+ (2 \times 8 \times 0.08)$

$= 9.12 \qquad (= 2E[X])$

c) $E[\underline{X^2}] = \sum \underline{x^2} P(X=x)$

$= (2^2 \times 0.2) + (4^2 \times 0.4) + (6^2 \times 0.32)$

$+ (8^2 \times 0.08)$

$= 23.84 \qquad (\neq E[X]^2)$

Binomial distribution

Binomial experiment

- Consider a probability experiment that only has two outcomes – “success” or “failure”.
  - Eg: Toss a coin. If head is observed, then it is a success.
- This experiment is called a Bernoulli trial.
- On the other hand, suppose we repeat the trial a few times and count how many success we observed.
- These are called binomial experiments.

Bernoulli = 1 trial
Binomial = multiple trials

Binomial experiment

- A binomial experiment is a probability experiment that satisfies the following four requirements:
  - There must be a fixed number of trials.
  - Each trial can have only two outcomes. These outcomes can be considered as either "success" or "failure".
  - The outcomes of each trial must be independent of one another.
  - The probability of a success must remain the same for each trial.

- "Success" does not imply that something good or positive has occurred.

Binomial distribution

- Let X be the number of successes observed in a binomial experiment with n number of trials.

- Then the probability distribution of X is called the binomial distribution.

Assume independent

- Example:
  - Toss a coin 10 times, and let X be the number of times head is observed.
  - Roll a die 5 times, and let X be the number of times odd numbers are observed.

Notation for binomial distribution

- We usually use these notations:
  - n: Total number of trials.
  - p: The probability of success.
  - q: The probability of failure. $q = 1 - p$.
  - X: Number of successes in n trials.

- Note that $0 \leq X \leq n$. X can take the values 0, 1, 2, …, n

$X \sim$ Binomial (n,p)

Probability distribution for binomial

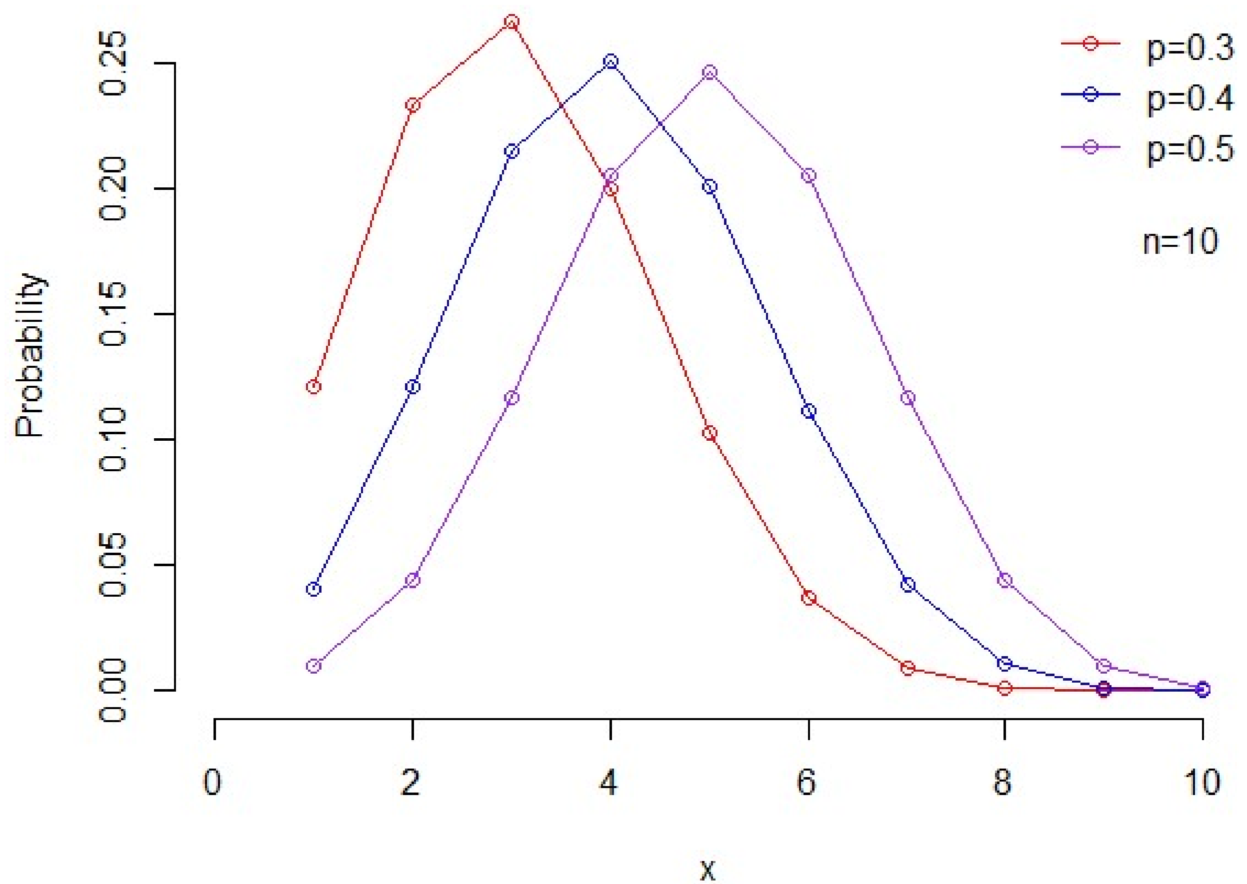
- The probability distribution of X:

$$P(X = x) = {}^{n}C_{x} p^{x} q^{n-x}$$

  for $x = 0, 1, \dots, n$

- How do we get this probability?
  - Out of the n trials, x number of them are successes.
    - Total possible way to get x successes out of n trials is ${}^{n}C_{x}$
  - There are x number of successes, each with probability p.
    - The probability for this is p^{x}
  - There are $n - x$ number of failures, each with probability q.
    - The probability for this is q^{n-x}
  - Since they are independent, we can multiply their probabilities.

Probability distribution for binomial



Mean and variance for binomial distribution

- Mean:

$$\mu = np$$

- Variance:

$$\sigma^2 = npq = np(1-p)$$

- Standard deviation:

$$\sigma = \sqrt{npq} = \sqrt{np(1-p)}$$

Example

- A coin is tossed 3 times.
  - Find the probability that we observe 2 heads.
  - Find the mean, variance, and standard deviation of the number of heads that will be obtained.

- Let X be the number of times we observe heads. X has a binomial distribution.
- In this case, "success" is defined as "observing head".

$$p = P(\text{head}) = 0.5, \qquad q = P(\text{tail}) = 0.5$$

- The number of trials, n is 3.
- Probability of observing 2 heads:

$$P(X = 2) = {}^{3}C_{2}(0.5)^{2}(0.5)^{1} = 0.375$$

Example

- Mean:

$$\mu = np = 3(0.5) = 1.5$$

- Variance:

$$\sigma^2 = npq = 3(0.5)(0.5) = 0.75$$

- Standard deviation:

$$\sigma = \sqrt{npq} = \sqrt{0.75} = 0.8660$$

Another example

- A die is rolled 3 times.
  - Find the probability that we observe numbers 1 or 2 two times.
  - Find the mean and variance of the number of times 1 or 2 are observed.

- Let X be the number of times 1 or 2 are observed. X has a binomial distribution.
- Number of trials: $n = 3$.
- "Success" is observing 1 or 2.
- Probability of success and failure:

$$p = \frac{2}{6} = \frac{1}{3}, \qquad q = 1 - p = \frac{2}{3}$$

- Probability $X = 2$:

$$P(2) = {}^{3}C_{2}\left(\frac{1}{3}\right)^{2}\left(\frac{2}{3}\right)^{1} = 0.2222$$

Another example

- Mean:

$$\mu = np = 3\left(\frac{1}{3}\right) = 1$$

- Variance:

$$\sigma^2 = npq = 3\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{2}{3}$$

Exercise

- Forty percent of prison inmates were unemployed when they entered prison. If 5 inmates are randomly selected, find these probabilities:
  a) Exactly 3 were unemployed.
  b) At most 4 were unemployed.
  c) At least 3 were unemployed.
  d) Fewer than 2 were unemployed.

Let X be the number of inmates that were unemployed out of 5 selected inmates

X has a binomial distribution with

$n = 5, \quad p = 0.4, \quad q = 0.6$

$X \sim Binom(n, p)$

$X \sim Binom(5, 0.4)$

X can take value 0, 1, 2, 3, 4, 5

a) $P(X=3) = {}^nC_x p^x q^{n-x} = {}^5C_3 (0.4)^3 (0.6)^{5-3}$

$= 0.2304$

b) $P(X \leq 4) = 1 - P(X=5) = 1 - {}^5C_5 (0.4)^5 (0.6)^0$

$= 0.9898$ (or $P(0)+P(1)+P(2)+P(3)+P(4)$)

c) $P(X \geq 3) = P(X=3) + P(X=4) + P(X=5)$

$= {}^5C_3 (0.4)^3 (0.6)^2 + {}^5C_4 (0.4)^4 (0.6)^1$

$+ {}^5C_5 (0.4)^5 (0.6)^0$

$= 0.3174$

d) $P(X<2) = P(X=0) + P(X=1) = {}^5C_0 (0.4)^0 (0.6)^5$

$+ {}^5C_1 (0.4)^1 (0.6)^4$

$= 0.3370$

Exercise

- Thirty-two percent of adult Internet users have purchased groceries online. For a random sample of 200 adult Internet users, find the mean, variance, and standard deviation for the number who have purchased groceries online.

X = number of adult who have purchased groceries online, out of 200 adults

X has a binomial distribution with

$n = 200$, $p = 0.32$, $q = 1 - p = 0.68$

$X \sim \text{Binom}(\underset{n}{200}, \underset{p}{0.32})$

$$\mu = np = 200 \times 0.32 = 64$$

$$\sigma^2 = npq = 200 \times 0.32 \times 0.68 = 43.52$$

$$\sigma = \sqrt{43.52} = 6.597$$

Poisson distribution

Poisson experiment

- A Poisson experiment is a probability experiment that satisfies the following requirements:
  - The random variable X is the number of occurrences of an event over some interval (i.e., length, area, volume, period of time, etc.).
  - The occurrences occur randomly.
  - The occurrences are independent of one another.
  - The average number of occurrences over an interval is known.

- Note that $X \geq 0$ and can take the values 0, 1, 2, …

Poisson distribution

- Let X be the outcome of a Poisson experiment.
- Then the distribution of X is a Poisson distribution.
- Example:
  - Number of patients admitted in a hospital in a day.
  - Number of customers in 5 hours interval.
  - Number of typographical error in a page.

Probability distribution for Poisson

- Denote λ be the mean number of occurrences in that interval.

- Then probability distribution of X:

$X \sim \text{Poisson}(\lambda)$

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

for $x = 0,1,2,\ldots$

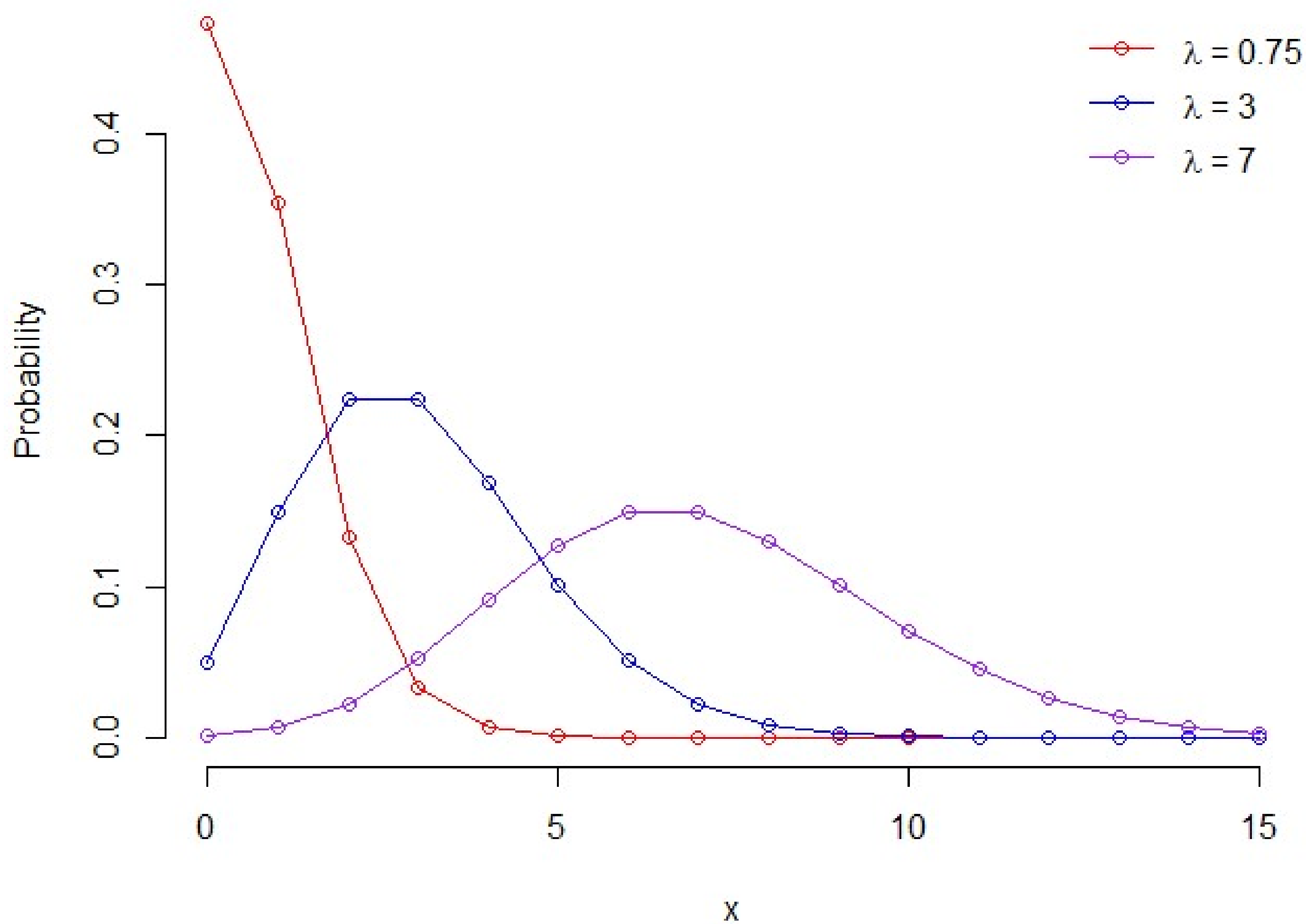
- Mean:

$$\mu = \lambda$$

- Variance and standard deviation:

$$\sigma^2 = \lambda, \qquad \sigma = \sqrt{\lambda}$$

Probability distribution for Poisson



Example

- A sales firm receives, on average, 3 calls per hour on its toll-free number. For any given hour, find the probability that it receives exactly 5 calls.

- Let X be the number calls received in an hour. X has a Poisson distribution.

$X \sim \text{Poisson}(3)$

- Mean number of calls per hour, $\lambda = 3$.
- Probability receiving exactly 5 calls:

$$P(5) = \frac{3^5 e^{-3}}{5!} = 0.1008$$

Another example

- A sales firm receives, on average, 3 calls per hour on its toll-free number. Find the probability that it receives exactly 5 calls in a two-hour period.

- Let X be the number calls received in a two-hour period. X has a Poisson distribution.

$X \sim \text{Poisson}(6)$

- Mean number of calls in a two-hour period:

$$\lambda = \text{mean number per hour} \times 2 = 3 \times 2 = 6$$

- Probability receiving exactly 5 calls:

$$P(5) = \frac{6^5 e^{-6}}{5!} = 0.1606$$

Exercise

Universiti Kebangsaan Malaysia
The National University of Malaysia

- A recent study of robberies for a certain geographic region showed an average of 1 robbery per 20,000 people. In a city of 80,000 people, find the probability of the following.
  a) 0 robberies
  b) 1 robbery
  c) 2 robberies

X = number of robberies in a city of 80 000 people.

X has a Poisson distribution with

$$\lambda = \frac{1}{20\,000} \times 80000 = 4$$

$X \sim \text{Poisson}(4)$

a) $P(X=0) = \frac{\lambda^{x} e^{-\lambda}}{x!} = \frac{4^{0} e^{-4}}{0!} = 0.01832$

b) $P(X=1) = \frac{4^{1} e^{-4}}{1!} = 0.07326$

c) $P(X=2) = \frac{4^{2} e^{-4}}{2!} = 0.1465$

Exercise

- The mean number of accidents per month at a certain intersection is three. Find the probability that in any given month,
  - a) No accidents will occur.
  - b) At least one accident will occur.
  - c) Less than three accident will occur.

X = number of accidents in a month

X has a Poisson distribution with $\lambda = 3$

$X \sim \text{Poisson}(3)$

a) $P(X=0) = \dfrac{\lambda^x e^{-\lambda}}{x!} = \dfrac{3^0 e^{-3}}{0!} = 0.04979$

b) $P(X \geq 1) = 1 - P(X=0) = 1 - 0.04979$

$= 0.9502$

c) $P(X<3) = P(X=0) + P(X=1) + P(X=2)$

$= \dfrac{3^0 e^{-3}}{0!} + \dfrac{3^1 e^{-3}}{1!} + \dfrac{3^2 e^{-3}}{2!}$

$= 0.4232$

Geometric distribution

Geometric distribution

- A geometric experiment is a probability experiment such that:
  - Each trial has two outcomes, either success or failure
  - The outcomes are independent of each other
  - The probability of a success is the same for each trial
  - The experiment continues until a success is obtained.

- Geometric distribution is related to the binomial.
  - Binomial random variable counts number of successes in n trials.
  - Geometric random variable counts number of trials until success.

- Note: in some books, the geometric r.v. may refer to the number of failures until success

Probability distribution for geometric

- Probability distribution:

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

for $x = 1,2,3, \ldots$

$X \sim \text{Geometric}(p)$

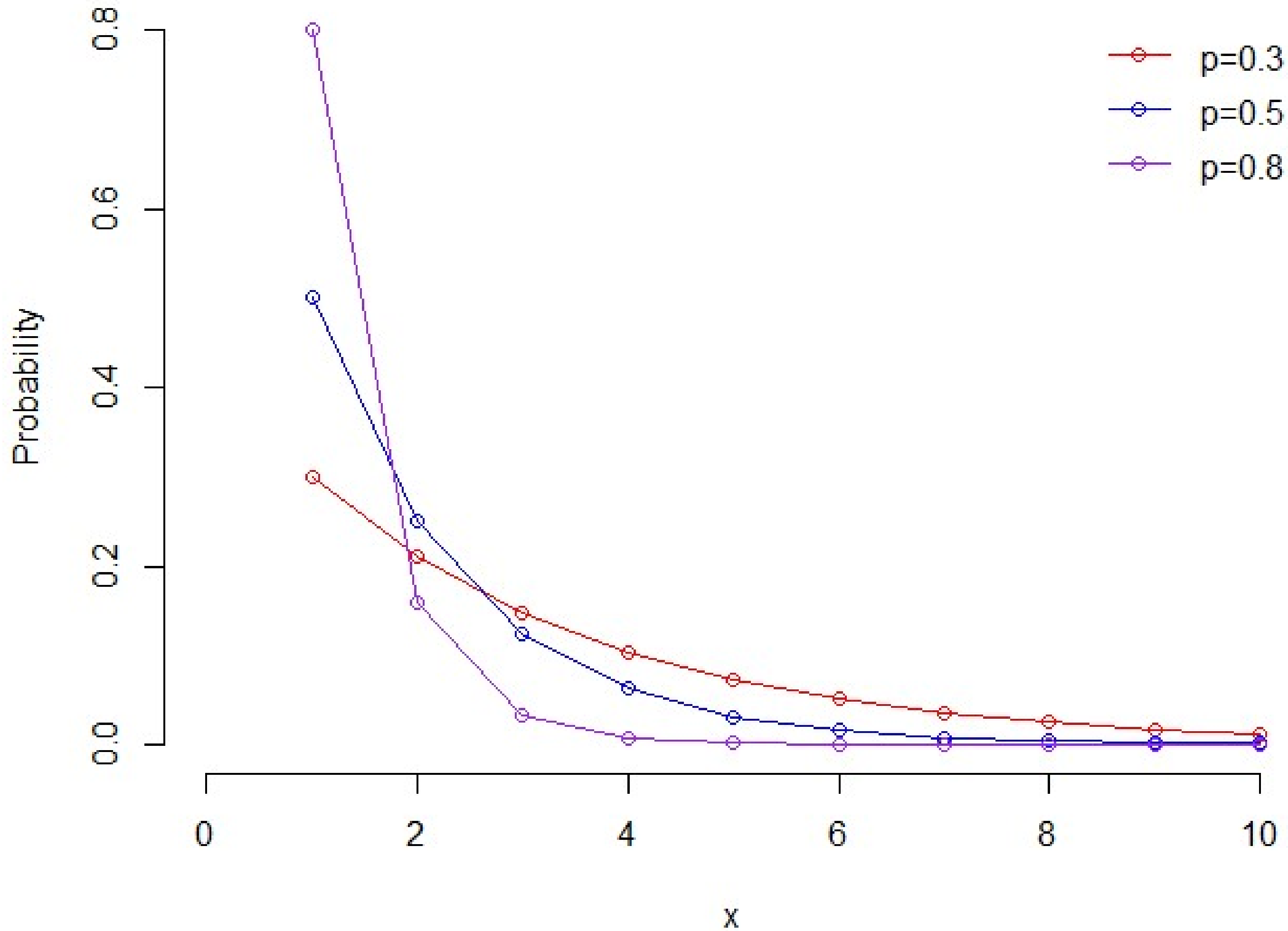
- Mean:

$$\mu = \frac{1}{p}$$

- Variance:

$$\sigma^2 = \frac{1-p}{p^2}$$

Probability distribution for geometric



Example

- Suppose that the probability that an applicant for a driver's license will pass the road test on any given try is 0.75, and assume the tests are independent.

- Let X be the number of road test taken by an applicant until they pass the test.

- In this case, X can take values 1, 2, 3, … and X follows the geometric distribution with $p = 0.75$.

- For example, the probability that it will take an applicant passing the test for the first time at the 4th attempt is

$$P(X = 4) = (0.25)^3 0.75 = 0.01171$$

- Additionally, the mean number of attempts it will take for an applicant to pass the test is $\frac{1}{p} = 1.333$

Exercise

UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

- The probability that you will make a sale on any given telephone call is 0.19. Find the probability that you:
  a) make your first sale on the fifth call.
  b) make your first sale on the first, second, or third call.
  c) do not make a sale on the first three calls.

X = number of call until the first sale

X has a geometric distribution with $p = 0.19$

a) $P(X=5) = (1-p)^{x-1} p = (1-0.19)^4 (0.19)$

$= 0.08179$

b) $P(1 \leq X \leq 3) = P(X=1) + P(X=2) + P(X=3)$

$= (0.81)^0(0.19)^1 + (0.81)^1(0.19)^1$

$+ (0.81)^2(0.19)$

$= 0.4686$

c) $P(X>3) = 1 - P(X \leq 3)$

$= 1 - 0.4686$

$= 0.5314$

Summary

- This chapter introduces discrete random variables.
- We first introduced what random variables are.
- Then we focused on discrete random variables including
  - Probability distribution
  - Distribution function
  - Mean
  - Variance
- We looked at three special case of discrete random variables:
  - Binomial distribution – number of successes in n trials.
  - Poisson distribution – number of events in an interval.
  - Geometric distribution – number of trials until success