بسم الله الرحمن الرحيم

QUA 107

INTRODUCTION TO BUSINESS STATISTICS

Prepared by

Mohammad Althobaiti

King Saud University
Dept. of Quantitative Analysis
Email: malthobaiti1@ksu.edu.sa

18 Nov 2024

Chapter 1:
What is Statistics?

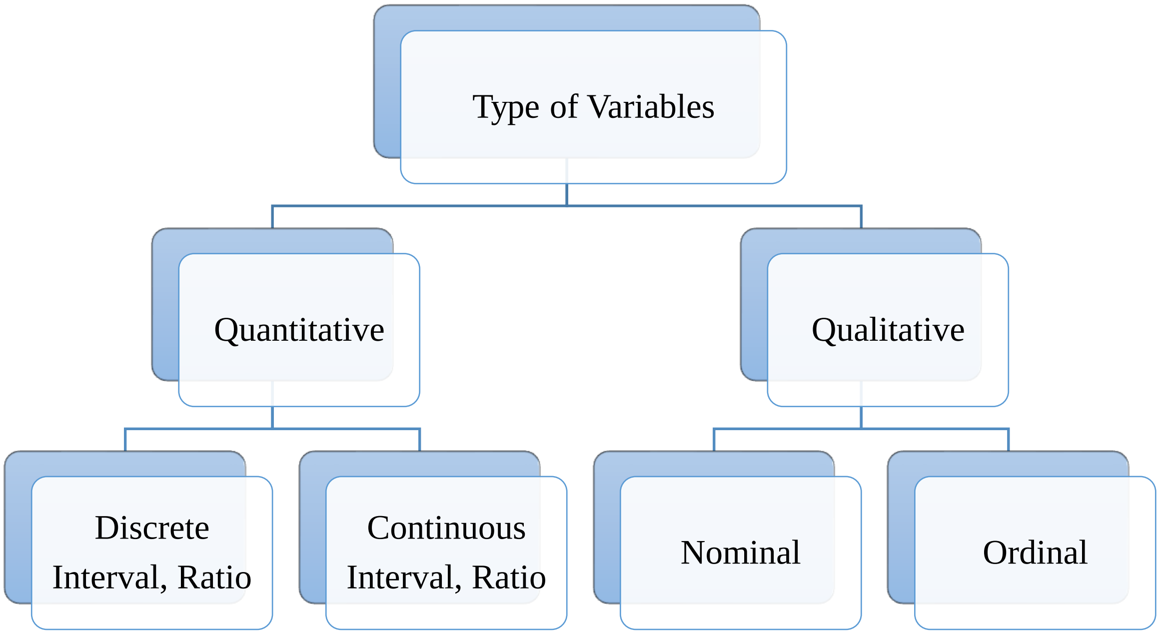
- **Some Basic Concepts:**

Statistics: The science of collecting, organizing, presenting, analyzing, and interpreting data to assist in making more effective decisions.
Type of Statistics:
1. Descriptive Statistics Methods of organizing, summarizing, and presenting data in an informative way.
2. Inferential Statistics The methods used to estimate a property of a population on the basis of a sample.

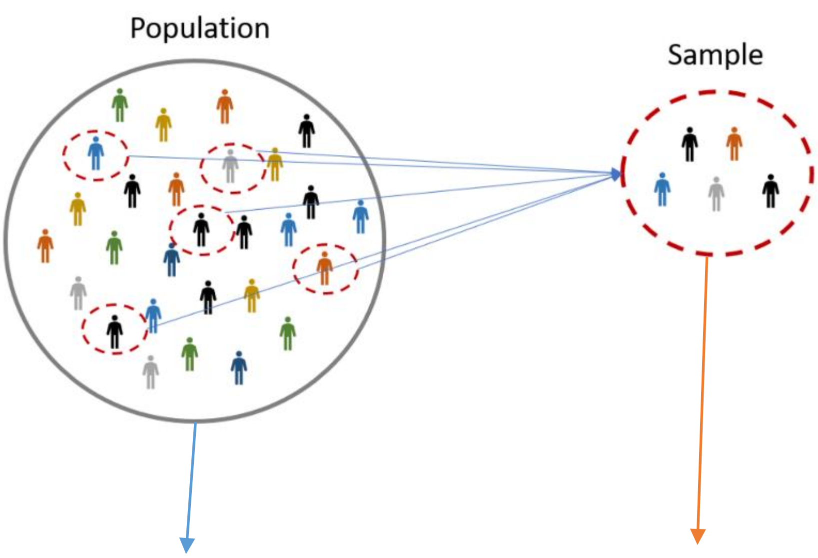
Population: The entire set of individuals or objects of interest or the measurements obtained from all individuals or objects of interest.
Sample: A portion or part of the population of interest.
Types of variables:
Qualitative Variable: An object or individual is observed and recorded as a non-numeric characteristic or attribute.
Examples: Gender, state of birth, eye color.
Quantitative Variable: A variable that is reported numerically.
Examples: Balance in your checking account, the life of a car battery, the number of people employed by a company.

Nominal: Data recorded at the nominal level of measurement is represented as labels or names. They have no order. They can only be classified and counted.
Ordinal: Data recorded at the ordinal level of measurement is based on a relative ranking or rating of items based on a defined attribute or qualitative variable. Variables based on this level of measurement are only ranked and counted.
Interval: For data recorded at the interval level of measurement, the interval or the distance between values is meaningful. The interval level of measurement is based on a scale with a known unit of measurement.
Ratio Data recorded at the ratio level of measurement are based on a scale with a known unit of measurement and a meaningful interpretation of zero on the scale.

- **<u>Some Basic Concepts:</u>**



- <u>Statistical Inference:</u>



Parameter	Statistic
Population mean μ Population variance σ^2	Sample mean $\bar{X}$ Sample variance S^2

Question 1:

1. The number of students admitted in College of Business Administration in King Saud University is a variable of type

A	Discrete	B	Qualitative	C	Continuous	D	Nominal

2. A mean of a population is called:

A	Parameter	B	Statistic	C	Median	D	Mode

3. The measure that obtained from the population is called

A	Parameter	B	Sample	C	Population	D	Statistic

4. The measure that obtained from the sample is called

A	Parameter	B	Sample	C	Population	D	Statistic

5. Which of the following is an example of a statistic?

A	Population variance	B	Sample median	C	Population mean	D	Population mode

6. A sample is defined as:

A	The entire population of values.
B	A measure of reliability of the population.
C	A subset of data selected from a population.
D	Inferential statistics.

7. The continuous variable is a

A	Variable takes on values within intervals.
B	Variable which can’t be measured.
C	Variable with a specific number of values.
D	Variable with no mode.

8. The nominal variable is a

A	Qualitative variable which can't be ordered.
B	Quantitative variable.
C	Qualitative variable which can be ordered.
D	Variable with a specific number of values.

9. One of the following is an example of an ordinal variable:

A	Socio-economic level.
B	Place of birth.
C	The time of finish the exam.
D	The number of students enrolled QUA 107.

10. One of the following is an example of a statistic:

A	The sample mode.
B	The population median.
C	The population variance.
D	None of these.

11. One of the following is a part of a population:

A	Sample.
B	Statistic
C	Variable
D	None of these

12. The variable is a

A	Characteristic of the population to be measured.
B	Subset of the population.
C	Parameter of the population.
D	None of these.

Question 2:

From customers aged more than 20 years living in Qaseem, we select 200 customers. It was found that the average spending of the customers was 150 SAR per day.

1. The variable of interest is:

A	Age	B	Spending	C	200 men	D	76 kg

2. The sample size is:

A	150	B	20	C	200	D	1520

Question 3:

A study of 250 customers who made purchases during the past year revealed that, on the average (mean), the customers lived 15 miles from the store.

1. The sample in the study is:

A	250 customers	B	250 hospitals	C	250 houses	D	15 miles

2. The population in this study is:

A	Some customers who made purchases during the past year.
B	All customers who made purchases during the past year.
C	250 customers who made purchases during the past year.
D	500 customers who made purchases during the past year.

3. If a researcher is interested in studying the customer satisfaction level (High, Normal, Low) for 13 customers, what is the type of variables?

A	Qualitative nominal.
B	Quantitative nominal.
C	Qualitative ordinal.
D	Quantitative ordinal.

<u>Question 4:</u>

What is the level of measurement for each of the following variables?

1. Student IQ rating scale:

A	Interval	B	Ratio	C	Nominal	D	Ordinal

2. The distance student travel to class is:

A	Interval	B	Ratio	C	Nominal	D	Ordinal

3. The jersey number of Alhilal football team is:

A	Interval	B	Ratio	C	Nominal	D	Ordinal

4. The academic degree (Bachelor, Master and PhD) is:

A	Interval	B	Ratio	C	Nominal	D	Ordinal

5. Study time for university student per week is:

A	Interval	B	Ratio	C	Nominal	D	Ordinal

Chapter 2:
Describing Data: Frequency Tables, Distribution and Graphic Presentation.

- **Creating a frequency distribution:**

Step 1:

Decide on the number of classes.

Use the $2^k > n$ rule,

Where,

k is the number of classes.

n is the number of values in the data set.

Step 2:

Determine the class interval.

The classes all taken together must cover the distance from the minimum to the maximum.

The class interval: $i = \frac{\text{Maximum value} - \text{Minmum value}}{k}$

Step 3:

Set the individual class limits.

Step 4:

Tally the data into classes and determine the number of the observations in each class.

Question 1:

The following data is for stock market profit during last month in SAR.

17	34	1	46	23
21	39	12	54	29
22	41	14	58	31
19	35	2	52	26
19	35	5	54	27
21	38	6	54	27
23	42	15	56	31
23	43	16	60	32
17	35	1	47	25
18	35	2	52	26

Create a frequency distribution for the data set.

- Deciding on the number of classes:

$n = 50$

$2^k > n = 2^k > 50$ $2^5 = 32$, $2^6 = 64$ since $2^6 > 50$. So, we use 6 classes.

- Determine the class interval:

$$i = \frac{\text{Maximum value} - \text{Minmum value}}{k} = \frac{60-1}{6} = 9.83 \approx 10$$

- The frequency distribution is:

Classes	Frequency	Tally
0 up to 10	6	𝍸 \
10 up to 20	9	𝍸 \\\\
20 up to 30	12	𝍸 𝍸 \\
30 up to 40	10	𝍸 𝍸
40 up to 50	5	𝍸
50 up to 60	8	𝍸 \\\
Total	50	

Question 2:

The following data is for stock market profit during last month in SAR.

37	51	23	44	59
38	52	26	69	63
42	55	31	34	46
33	45	22	37	49
43	55	31	41	53
32	65	22	22	27

Create a frequency distribution for the data set.

- Deciding on the number of classes:

$n = 30$

$2^k > n = 2^k > 30$ $\quad 2^4 = 16$, $2^5 = 32$ $\quad$ since $2^5 > 30$. So, we use 5 classes.

- Determine the class interval:

$$i = \frac{\text{Maximum value} - \text{Minmum value}}{k} = \frac{69-22}{5} = 9.4 \approx 10$$

- The frequency distribution is:

Classes	Frequency	Tally
20 up to 30	6	~~\\\\\\\\~~ \
30 up to 40	8	~~\\\\\\\\~~ \\\
40 up to 50	7	~~\\\\\\\\~~ \\
50 up to 60	6	~~\\\\\\\\~~ \
60 up to 70	3	\\\
Total	30	

- **Frequency Tables:**

Example:

Let's consider a dataset of ages in a group, and we want to organize the data into class intervals:

Class Interval (Age)	Midpoint	Frequency	Relative frequency	Cumulative frequency	Relative cumulative frequency
10 up to 20	$\left(\frac{10+20}{2}\right) = 15$	2	$\frac{2}{25} = 0.08$	2	$\frac{2}{25} = 0.08$
20 up to 30	$\left(\frac{20+30}{2}\right) = 25$	4	$\frac{4}{25} = 0.16$	$(2+4) = 6$	$\frac{6}{25} = 0.24$
30 up to 40	$\left(\frac{30+40}{2}\right) = 35$	10	$\frac{10}{25} = 0.40$	$(2+4+10) = 16$	$\frac{16}{25} = 0.64$
40 up to 50	45	8	$\frac{8}{25} = 0.32$	$(2+4+10+8) = 24$	$\frac{24}{25} = 0.96$
50 up to 60	55	1	$\frac{1}{25} = 0.04$	$(2+4+10+8+1) = 25$	$\frac{25}{25} = 1$
Total		25	1	-	-

class width =10 sample size (n) always =1 always = n always = 1

Question 1:

The “life” of 40 similar car batteries recorded to the nearest tenth of a year. The batteries are guaranteed to last 3 years.

Class Interval	Midpoint	Frequency	Relative Frequency
1.45 up to 1.95	1.7	2	0.050
1.95 up to 2.45	2.2	B	0.025
2.45 up to 2.95	A	4	D
2.95 up to 3.45	3.2	15	0.375
3.45 up to 4.95	3.7	C	0.250
3.95 up to 4.45	4.2	5	0.125
4.45 up to 4.95	4.7	3	0.075

1. The value of A: $A = \frac{2.45+2.95}{2} = 2.7$

2. The value of B: $\frac{B}{40} = 0.025 \Longrightarrow B = 40 \times 0.025 = 1$

3. The value of C: $\frac{C}{40} = 0.25 \Longrightarrow C = 40 \times 0.25 = 10$

4. The value of D: $D = \frac{4}{40} = 0.10$

Question 2:

Fill in the table given below. Answer the following questions.

Class Interval	Frequency	Cumulative Frequency	Relative Frequency	Cumulative Relative Frequency
5 up to 10	8			
10 up to 15	15		C	
15 up to 20	11	B		D
20 up to 25	A	40	0.15	

1) The value of A is: $A = 40 - (8 + 15 + 11) = 40 - 34 = 6$

2) The value of B is: $B = 8 + 15 + 11 = 34$

3) The value of C is: $C = \frac{15}{40} = 0.375$

4) The value of D is: $D = \frac{34}{40} = 0.85$

5) The midpoint of 2nd class is: $\frac{10+15}{2} = 12.5$

6) The number of observations less than 20 is: $8 + 15 + 11 = 34$

Question 3:

The table shows the weight loss (kg) of a sample of 40 healthy adults who fasted in Ramadan.

Class interval	Frequency	Cumulative Frequency
1.20 up to 1.30	2	2
1.30 up to 1.40	6	8
1.40 up to 1.50	10	K
1.50 up to 1.60	C	34
1.60 up to 1.70	6	40

1) The value of the missing value K is 18
2) The value of the missing value C is 16

Question 4:

The following table gives the distribution of the ages of a sample of 50 patients who attend a dental clinic.

Class limits	Frequency	Relative Frequency
15 - 20	4	-
20 - 25	8	-
25 - 30	z	0.32
30 - 35	-	-
35 - 40	10	-

Age	Cumulative frequency
Less than 20	4
Less than 25	y
Less than 30	-
Less than 35	-
Less than 40	x

1. The class width is: $20 - 15 = 5$

A	5	B	10	C	150	D	19

2. The value of x is:

A	22	B	28	C	50	D	10

3. The value of y is: $4 + 8 = 12$

A	4	B	12	C	19	D	150

4. The value of z is: $\frac{z}{50} = 0.32 \implies z = 50 \times 0.32 = 16$

A	14	B	12	C	50	D	16

5. Percent of the patients with age between 20 and 25 is: $\frac{8}{50} \times 100 = 16$

A	16%	B	8%	C	20%	D	32%

6. The 5th interval midpoint is: $\frac{35+40}{2} = 37.5$

A	38	B	52	C	27	D	37.5

Question 5:

Consider the following Table showing a frequency distribution of weights in a sample of 20 cans of fruits:

Class interval	Midpoint	Frequency	Relative Frequency	Cumulative Frequency
19.2 – 19.4		1		
19.4 – 19.6			0.1	
19.6 – 19.8		8		
19.8 – 20.0		4		

1. The fifth-class interval is:

A	20.2 - 20.4	B	20.1-20.3	C	21.0 - 21.2	D	20.0 - 20.2

2. The midpoint of the fourth-class interval is: $\frac{19.8+20}{2} = 19.9$

A	20.5	B	19.9	C	19.9	D	20.1

3. The frequency of the second-class interval is: $\frac{x}{20} = 0.1 \implies x = 20 \times 0.1 = 2$

A	10	B	4	C	2	D	3

4. The relative frequency of the fourth-class interval is: $\frac{4}{20} = 0.20$

A	0.20	B	0.15	C	0.13	D	0.40

5. The cumulative frequency of the final class interval is:

A	13	B	4	C	20	D	100

Question 6:

The following table gives the distribution of weight (in Kg) of a sample of 200 persons.

Weights	Frequency	Cumulative frequency	Relative frequency
50 - 55	40		
55 - 60	A		0.15
60 - 65		B	
65 - 70	60		C
70 - 75	20	D	

[1] The value of A is $\frac{A}{200} = 0.15 \implies A = 200 \times 0.15 = 30$

A	20	B	30	C	40	D	50

[2] The value of B is

A	90	B	100	C	110	D	120

[3] The value of C is $C = \frac{60}{200} = 0.3$

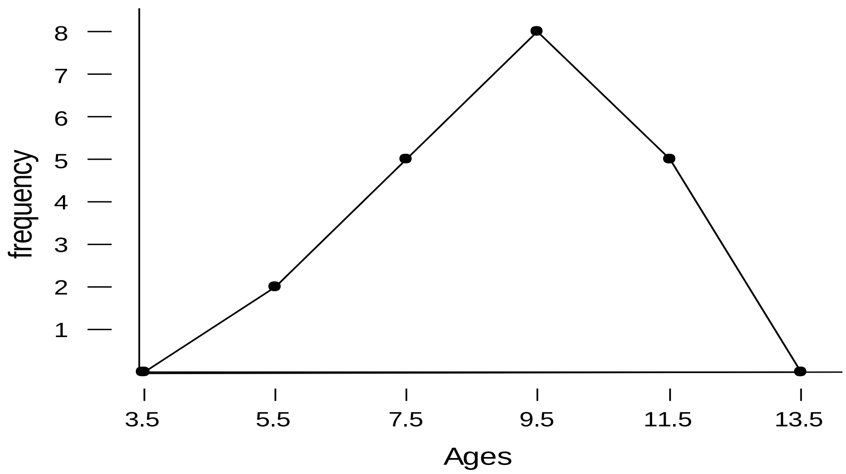
A	0.3	B	0.4	C	0.5	D	0.6

[4] The value of D is

A	150	B	85	C	80	D	200

Question 7:

Consider the following frequency polygon of ages of 20 students in a certain school.



The frequency distribution of ages corresponding to above polygon is

(a)

Class limits	4.5- 6.5	6.5-8.5	8.5- 10.5	10.5 -12.5
frequency	2	5	8	5

(b)

Class limits	3.5- 5.5	5.5-7.5	7.5- 9.5	9.5 -11.5
frequency	2	5	8	4

(c)

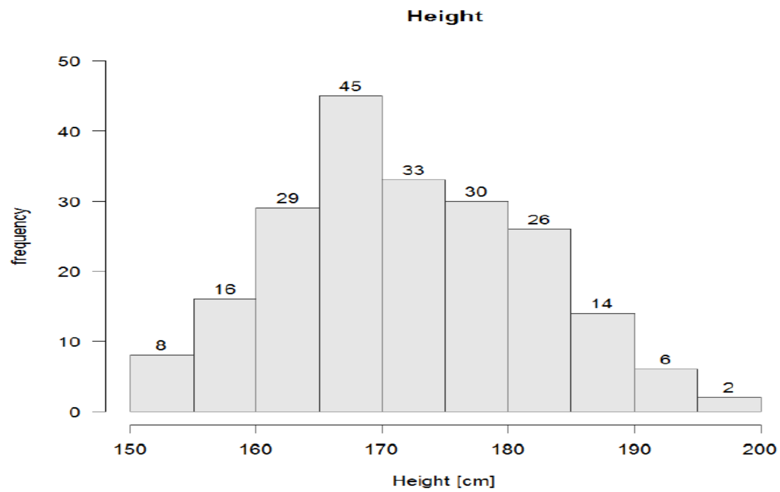
Class interval	5- 6	7-8	9- 10	11 -12
frequency	1	7	8	4

(d)

Class interval	5- 6	7-8	9- 10	11 -12
frequency	4	7	8	6

Question 8:

For a sample of students, we obtained the following graph for their height in (cm).



1. The variable under study is:

A	Patients	B	Graph	C	Height	D	Discrete

2. The type of variable:

A	Continuous	B	Discrete	C	Frequency	D	Height

3. The number of students with the lowest level height:

A	14	B	2	C	115	D	8

4. The sample size is: $8 + 16 + 29 + 45 + 33 + 30 + 26 + 14 + 6 + 2 = 209$

A	28	B	209	C	156	D	130

5. The midpoint of the interval with highest frequency is: $\frac{165+170}{2} = 167.5$

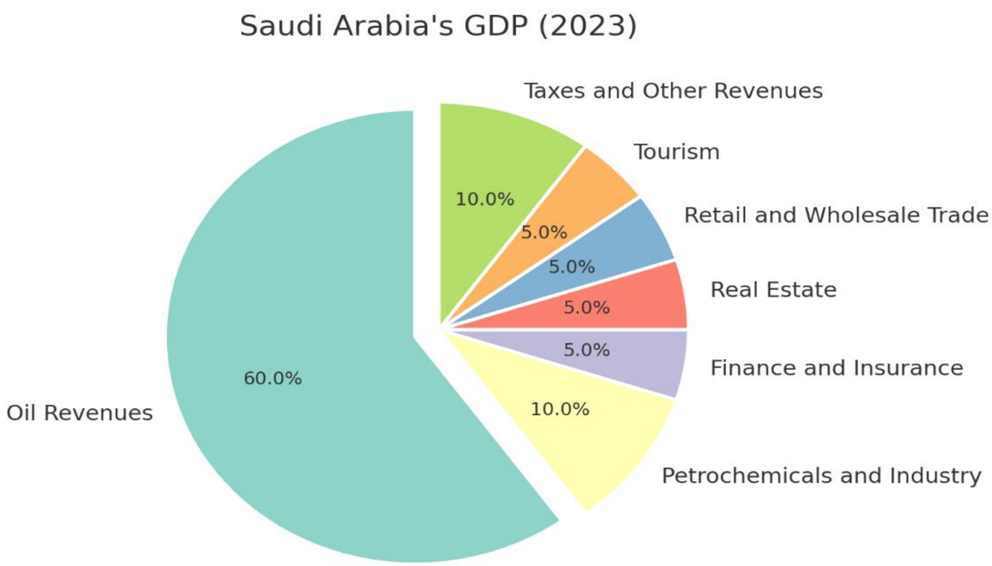
A	182.5	B	130.5	C	167.5	D	30

6. The relative frequency of the interval with highest frequency is: $\frac{45}{209} = 0.215$

A	0.283	B	0.215	C	0.241	D	0.262

Question 9:

The following chart is for the approximate GDP (الناتج المحلي)
of Saudi Arabia 2023



a. What percentage of state resources is accounted for by "Real estate" and "Petrochemicals and industry"?

A	%5	B	%10	C	%15	D	%20

b. What source will generate more GDP for Saudi Arabia?

A	Real estate	B	Oil revenues	C	Tourism	D	Insurance

c. The Saudi Arabia's GDP in 2023 is estimated to 4,160 billion SR. Estimate the amount of revenue in billions of riyals for “Tourism”.

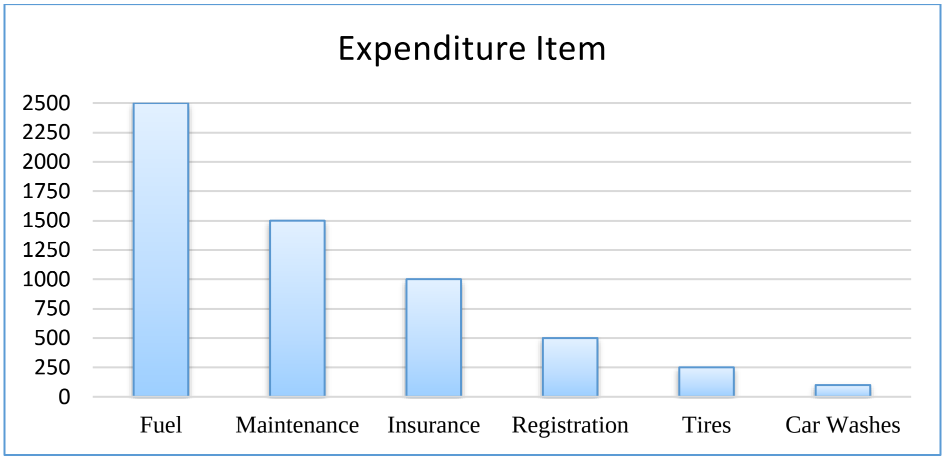
A	208 billion	B	416 billion	C	624 billion	D	2496 billion

d. The Saudi Arabia's GDP in 2023 is estimated to 4,160 billion SR. Estimate the amount of revenue in billions of riyals for “Oil revenues”.

A	208 billion	B	416 billion	C	624 billion	D	2496 billion

Question 10:

A recent survey in Saudi, showed that the typical Korean car owner spends 5850 SAR per year on operating expenses. Following is a breakdown of the various expenditure items.



a. According to the study results, what is the amount spent annually on car fuel?

A	2500 SAR	B	1500 SAR	C	1000 SAR	D	500 SAR

b. According to the study results, what is the percentage of amount spent annually on insurance?

A	1000	B	17.1%	C	8.5%	D	25.6%

c. According to the study results, which expenditure item has the lowest annual cost?

A	Fuel	B	Maintenance	C	Tires	D	Car Washes

d. What is the combined annual expenditure on Registration and Tires according to the survey?

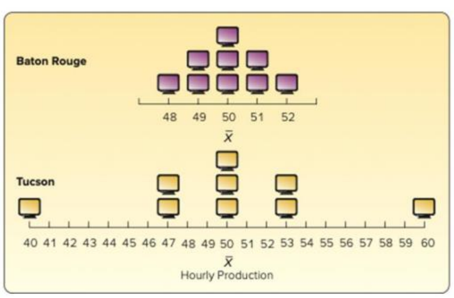
A	1000 SAR	B	250 SAR	C	750 SAR	D	500 SAR

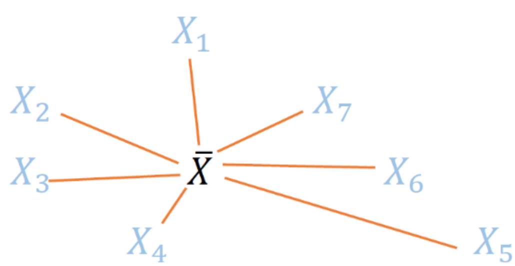
Chapter 3:
Describing Data: Numerical Measures.

- **Measures of Central tendency and Dispersion**

Measures of central tendency (Location)		
Mean	$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$	Unit
Median		Unit
Mode	The value with the highest frequency	Unit
Measures of dispersions (Shape)		
Range	$R = max - min$	Unit
Variance	$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$	Unit^2
Standard deviation	$S = \sqrt{S^2}$	Unit

- Example: Hourly production at two computer monitor plants.





- **For median:**

If the sample size (n) is an odd number, the rank of the median is

Ordered set	y_1	y_2	...	y_m	...	y_n
Rank	1	2	...	m	...	n

$$Rank = \frac{n+1}{2} = m$$

If the sample size (n) is an even number, the rank of the median is

Ordered set	y_1	y_2	...	y_m	y_{m+1}	...	y_n
Rank	1	2	...	m	$m+1$	...	n

$$Rank = \frac{n+1}{2} = m + 0.5$$

Find the median for the following cases:

Observations	Odd or Even	Ordering data	Median
Student's ages: 4, 5, 2, 9, 10, 8, 4	Odd	2 4 4 [5] 8 9 10	5
Student's ages: 10, 13, 9, 20, 11, 100	Even	9 10 [11 13] 20 100	$\frac{11+13}{2} = 12$
Student's grades: A, C, B, C, F, B, B	Odd	A B B [B] C C F	B
Student's grades: A, C, B, C, F, B, B, B	Even	A B B [B B] C C F	B
Student's grades: A, C, B, C, F, B, C, B	Even	A B B [B C] C C F	No median

Question 1:

If the number of visits to the clinic made by 8 pregnant women in their pregnancy period is:

12 15 16 12 15 16 12 14

1. The type of the variable is:

   discrete

2. The sample mean is:

   $$\bar{X} = \frac{12+15+16+12+15+16+12+14}{8} = 14$$

3. The sample standard deviation is:

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

$$= \frac{(12-14)^2+(15-14)^2+(16-14)^2+(12-14)^2+(15-14)^2+(16-14)^2+(12-14)^2+(14-14)^2}{8-1}$$

$$= 3.14 \Longrightarrow S = 1.77$$

4. The sample median is:

   12 12 12 14 15 15 16 16 $\Longrightarrow \frac{14+15}{2} = 14.5$

5. The range is:

   $16 - 12 = 4$

Question 2:

Consider the following marks for a sample of students carried out on 10 quizzes:

6, 7, 6, 8, 5, 7, 6, 9, 10, 6

1. The mean mark is:

$$\overline{X} = \frac{6+7+6+8+5+7+6+9+10+6}{10} = 7$$

2. The median mark is:

$$5\ 6\ 6\ 6\ \boxed{6\ 7}\ 7\ 8\ 9\ 10 \Longrightarrow \frac{6+7}{2} = 6.5$$

3. The mode for this data is: $\boxed{6}$
4. The range for this data is $\boxed{5}$
5. The standard deviation for this data is:

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1} = \frac{(5-7)^2+(6-7)^2+\cdots+(10-7)^2}{10-1} = 2.434 \Longrightarrow S = 1.56$$

Question 3:

Temperature recorded at 2 am in Hail on 8 days randomly chosen in a year were as follows:

21	6	-2	11	15	6	0	19

1) The average temperature for the sample is:

A	248	B	9.5	C	6	D	48

2) The median temperature for the sample is:

A	8.5	B	-21	C	-8.5	D	17

3) The mode of temperature for the sample is:

A	-2	B	44	C	2	D	6

4) The standard deviation for the sample data is:

A	35.319	B	30.904	C	71.714	D	8.468

5) The range of the sample is:

A	19	B	21	C	23	D	15

Question 4:

The following sample represents the number of years spent by senior-level employees working in one of the major companies in Riyadh:

5, 3, 5, 2, 5, 4

[1] The sample mean is

A	4	B	5	C	3	D	6

[2] The sample median is

A	4	B	5	C	4.5	D	3

[3] The sample mode is

A	4	B	3	C	4.5	D	5

[4] The sample standard deviation is

A	3.2649	B	8.2649	C	1.2649	D	2.2649

Question 5:

The following data represents the number of vacant positions (الوظائف الشاغرة) in some departments of the General Tourism Authority:

5, 8, 0, 8, 3, 7, 8, 9

1) The sample size is:

A	9	B	6	C	8	D	5

2) The sample mode is:

A	9	B	0	C	8	D	No mode

3) The sample mean is:

A	48	B	6	C	8	D	0

4) The sample variance is:

A	2.915	B	8.5	C	9.714	D	3.117

5) The sample median is:

A	5.5	B	7.5	C	8	D	7

6) The range of data is:

A	8	B	0	C	3	D	9

Question 6:

1. The …... is a measure of location?

A	Median	B	Range	C	Variance	D	CV

2.The …... is not affected by extreme values?

A	Range	B	Mean	C	Median	D	Variance

3.Which of the following measures can be used for the blood type in a given sample?

A	Median	B	Mean	C	Variance	D	Mode

Question 7:

The frequency table for daily number of car accidents during a month is:

Number of car accidents	Frequency
3	2
4	3
5	1
6	2
7	2
Total	10

3,3,4,4,4,5,6,6,7,7

1. The type of variable:

A	Nominal	B	Discrete	C	Ordinal	D	Continuous

2. The mean for the number of accidents is:

A	4.07	B	4.90	C	3.75	D	2.98

3. The median is:

A	5.5	B	5	C	4.5	D	4

4. The mode is:

A	4	B	5	C	6	D	3

5. The variance for the number of accidents is:

A	8.45	B	6.43	C	2.32	D	1.05

Question 8:

1. The biggest advantage of the standard deviation over the variance is:

A	The standard deviation is always greater than the variance.
B	The standard deviation is calculated with the median instead of the mean.
C	The standard deviation is better for describing the qualitative data.
D	The standard deviation has the same units as the original data.

2. Which of the following location (central tendency) measures is affected by extreme values?

A	Median
B	Mean
C	Variance
D	Range

3. Which of the following measures can be used for gender in a given sample?

A	Mode
B	Mean
C	Variance
D	Range

4. If x_1,x_2 and x_3 has mean $\overline{\mathrm{x}} = 4$, then x_1,x_2, x_3 and $x_4 = 4$ has mean:

A	equal 4
B	less than 4
C	greater than 4
D	None of this

5. The sample mean is a measure of

A	Relative position.
B	Dispersion.
C	Central tendency.
D	all of the above

6. The sample standard deviation is a measure of

A	Relative position.
B	Central tendency.
C	Dispersion.
D	all of the above.

7. Which of the following are examples of measures of dispersion?

A	The median and the mode.
B	The range and the variance.
C	The parameter and the statistic.
D	The mean and the variance.

- **Weighted mean:**

The weighted mean is found by multiplying each observation by its corresponding weight.
A convenient way to compute the mean when there are several observations with the same value.

$$\bar{x}_w = \frac{w_1 x_1 + w_2 x_2 + \cdots + w_n x_n}{w_1 + w_2 + \cdots + w_n}$$

- $\bar{x}_w$ is read "x bar sub w".
- $x_1, x_2, \dots, x_n$ is the set of numbers.
- $w_1, w_2, \dots, w_n$ are the corresponding weights.

The denominator is always the sum of the weights.

<u>Question 1:</u>
A student has the following grades in three different courses, with each course having a different credit weight:

- Accounting: 85 (4 credits)
- Marketing: 90 (3 credits)
- Business Ethics: 75 (2 credits)

What is the student's weighted mean (average) grade across all courses?

The weighted mean = $\frac{(85\times4)+(90\times3)+(75\times2)}{9} = 84.44$

A	82	B	83.5	C	84.44	D	86

<u>Question 2:</u>
A company buys two types of raw materials to produce a product. The costs and the quantities purchased of each material are as follows:

- 5 SAR per unit for material A, the company buys 200 units.
- 8 SAR per unit for material B, the company buys 150 units.

What is the weighted mean cost per unit of raw materials for the company?

The weighted mean = $\frac{(5\times200)+(8\times150)}{350} = 6.29$

A	175 SAR	B	6.29 SAR	C	6.50 SAR	D	SAR

Question 3:

In June, an investor purchased 400 shares of stock at 66 SAR per share. In August, she purchased an additional 330 shares at 92 SAR per share. In November, she purchased an additional 330 shares at 85 SAR. What is the weighted mean price per share?

$$\text{The weighted mean} = \frac{(66\times 400)+(92\times 330)+(85\times 330)}{1060} = 80.0094$$

A	80 SAR	B	353.33 SAR	C	81.50 SAR	D	77.63 SAR

Question 4:

A Saudi company sells three types of products: A, B, and C. Each product has a different price and contributes differently to the total revenue based on the number of units sold. The company wants to calculate the average price of all products sold, taking into account the quantity sold for each.
Find the weighted average price.

Product	Price per Unit (SAR)	Units Sold
A	10	100
B	15	150
C	20	50

$$\text{The weighted mean} = \frac{(10\times 100)+(15\times 150)+(20\times 50)}{300} = 14.17$$

A	14.91 SAR	B	12.64 SAR	C	10.50 SAR	D	14.17 SAR

- **Geometric mean:**

• The geometric mean is useful in finding average rates of change over time. • The rates can be expressed as percentages (or ratios). • A percentage change is always (1+change) • Wide applications in business and economics (example: GDP (الناتج المحلي الإجمالي)). • The geometric mean will always be no more than the arithmetic mean. • The data values must be positive. $GM = \sqrt[n]{(x_1)(x_2)\dots(x_n)}$

Question 1:

An investor is analyzing the annual growth rates of a stock over three years. The stock's growth rates for the three years are as follows:

- Year 1: 10%
- Year 2: 15%
- Year 3: 5%

What is the geometric mean of the growth rates over these three years?

- Year 1: 10% $\Rightarrow$ 10% $= 0.10 \Rightarrow 1 + 0.10 = 1.10$
- Year 2: 15% $\Rightarrow$ 15% $= 0.15 \Rightarrow 1 + 0.15 = 1.15$
- Year 3: 5% $\Rightarrow$ 5% $= 0.05 \Rightarrow 1 + 0.05 = 1.05$

The geometric mean $= \sqrt[3]{1.10 \times 1.15 \times 1.05} = 1.0992$

$1.0992 - 1 = 0.0992$

$0.0992 \times 100 = 9.92\%$

A	9.5%	B	11%	C	9.92%	D	11.65%

Question 2:

A farmer wants to calculate the average growth rate of his crop yields over four years. The yield increased by the following factors each year:

Year 1	Year 2	Year 3	Year 4
20%	30%	10%	40%

What is the geometric mean growth rate over these four years?

- Year 1: 20% $\Rightarrow$ 20% $= 0.20 \Rightarrow 1 + 0.20 = 1.20$
- Year 2: 30% $\Rightarrow$ 30% $= 0.30 \Rightarrow 1 + 0.30 = 1.30$
- Year 3: 10% $\Rightarrow$ 10% $= 0.10 \Rightarrow 1 + 0.10 = 1.10$
- Year 4: 40% $\Rightarrow$ 40% $= 0.40 \Rightarrow 1 + 0.40 = 1.40$

The geometric mean $= \sqrt[4]{1.20 \times 1.30 \times 1.10 \times 1.40} = 1.245$

$1.245 - 1 = 0.245$

$0.245 \times 100 = 24.5\%$

A	24.5%	B	25.6%	C	22.9%	D	23.8%

Question 3:

The price of a stock changes over three consecutive years as follows:

2005	2006	2007
100%	-50%	100%

What is the geometric mean of the stock price change over these three years?

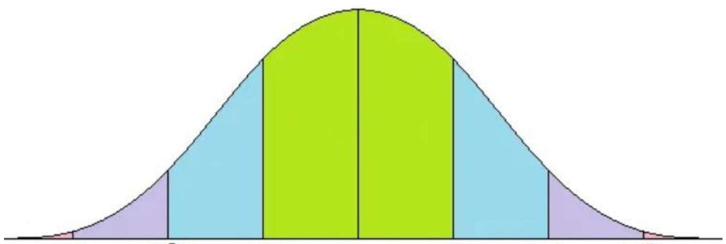
- Year 2005: 100% $\Rightarrow 100\% = 1 \Rightarrow 1 + 1 = 2$
- Year 2006: -50% $\Rightarrow -50\% = -0.50 \Rightarrow 1 - 0.50 = 0.50$
- Year 2007: 100% $\Rightarrow 100\% = 1 \Rightarrow 1 + 1 = 2$

The geometric mean $= \sqrt[3]{2 \times 0.50 \times 2} = 1.26$

A	1.0 times	B	1.26 times	C	1.5 times	D	1.75 times

- **Chebyshev's theorem:**

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$



Question 1:

According to the Chebyshev's theorem, for any random variables X with mean μ and variance σ^2,
what is the lower bound for $P(\mu - 2\sigma < X < \mu + 2\sigma)$?

By Chebyshev's theorem, the lower bound is $1 - \frac{1}{k^2}$
Givens: $P(\mu - 2\sigma < X < \mu + 2\sigma) \Rightarrow k = 2$

The lower bound is $1 - \frac{1}{k^2}$
$= 1 - \frac{1}{2^2} = 1 - \frac{1}{4} = 0.75$

A	4	B	0.50	C	2	D	0.75

Question 2:

Suppose that X is a random variable with mean $\mu = 12$, variance $\sigma^2 = 9$, and unknown probability distribution. Using Chebyshev's theorem, $P(3 < X < 21)$ is at least equal to:

By taking the lower or the upper value	
$\mu - k\sigma = 3$ $12 - 3k = 3$ $-3k = 3 - 12$ $-3k = -9$ $k = 3$	$\mu + k\sigma = 21$ $12 + 3k = 21$ $3k = 21 - 12$ $3k = 9$ $k = 3$

$k = 3 \Rightarrow 1 - \frac{1}{k^2}$
$= 1 - \frac{1}{3^2} = 1 - \frac{1}{9} = \frac{8}{9}$

A	8/9	B	3/4	C	1/4	D	1/16

<u>Question 3:</u>

Suppose that $\mu = 5$ and $\sigma^2 = 4$. Using Chebyshev's Theorem,

1. Find an approximated value of $P(1 < X < 9)$.

By taking the lower or the upper value	
$\mu - k\sigma = 1$	$\mu + k\sigma = 9$
$5 - 2k = 1$	$5 + 2k = 9$
$-2k = 1 - 5$	$2k = 9 - 5$
$-2k = -4$	$2k = 4$
$k = 2$	$k = 2$

$$k = 2 \Rightarrow 1 - \frac{1}{k^2}$$
$$= 1 - \frac{1}{2^2} = 1 - \frac{1}{4} = 0.75$$

2. Find some constants a and b $(a < b)$ such that $P(a < X < b) = 15/16$

To find the constant, a and b, we have to find the value k:

$$1 - \frac{1}{k^2} = \frac{15}{16}$$
$$\frac{1}{k^2} = \frac{1}{16} \Rightarrow k^2 = 16 \Rightarrow k = 4$$

$a = \mu - k\sigma = 5 - 4 \times 2 = -3$ $\quad$ $\mu = 5, \sigma^2 = 4$

$b = \mu + k\sigma = 5 + 4 \times 2 = 13$

$$P(-3 < X < 13) = 15/16$$

<u>Question 4:</u>

According to Chebyshev's theorem, at least what percent of any set of observations will be within 1.8 standard deviations of the mean?

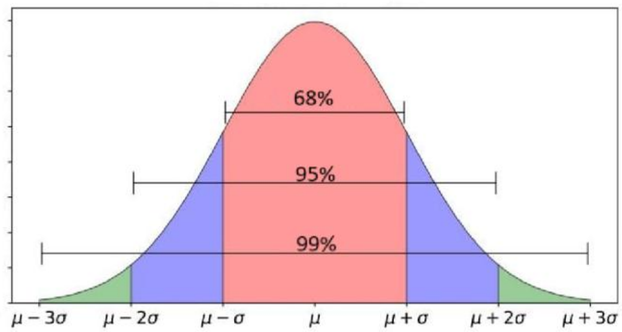
By Chebyshev's theorem, the lower bound is $1 - \frac{1}{k^2}$

Givens: $P(\mu - 1.8\sigma < X < \mu + 1.8\sigma) \Rightarrow k = 1.8$

The lower bound is $1 - \frac{1}{k^2}$
$$= 1 - \frac{1}{1.8^2} = 0.691$$

A	54.7%	B	81.2%	C	69.1%	D	81.2%

The Empirical Rule or Normal Rule provides an approximation. • 1 standard deviation of the mean: about 68% of values. • 2 standard deviations of the mean: about 95% of values. • 3 standard deviations of the mean: about 99% of values



Question 5:

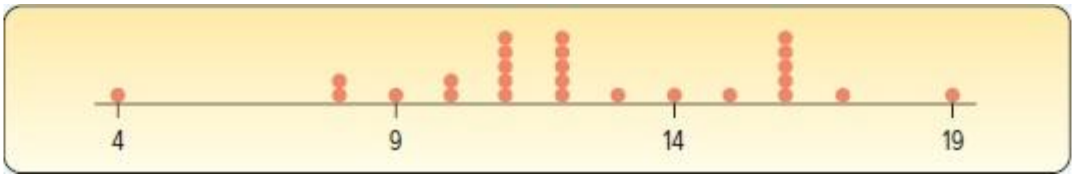
The distribution of the weights of a sample of 1,400 cargo containers is symmetric and bell shaped. According to the Empirical Rule, what percent of the weights will lie between $\bar{X} - 2s$ and $\bar{X} + 2s$

A	68%	B	75%	C	95%	D	68%

Chapter 4:
Describing Data: Displaying and Exploring Data.

<u>Question 1:</u>

The following chart reports the number of cell phones sold at a big-box retail store for the last 26 days.



1. What is the minimum of number of cell phones sold?

A	11,12 and 16	B	19	C	15	D	4

2. What is the maximum of number of cell phones sold?

A	11,12 and 16	B	19	C	15	D	4

3. What is the range of number of cell phones sold?

A	11,12 and 16	B	19	C	15	D	4

4. How would you describe the shape of the distribution:

A	Bell shaped	B	Negative skew	C	Positive skew	D	None

5. Visually, estimate the central location of the distribution.

A	11 or 12	B	12 or 16	C	16	D	12

6. Find the mode of the distribution.

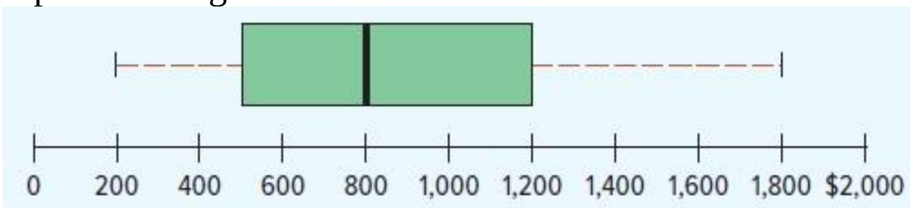
A	11,12 and 16		B	19	C	15	D	4

7. Find the median of the distribution.(≈ 12.38)

A	13	B	12	C	14	D	11

Question 2:

The box plot shows the amount spent for books and supplies per year by students at four-year public colleges.



1. Estimate the median amount spent.

A	1200$	B	500$	C	800$	D	700$

2. Estimate the first quartile for the amount spent.

A	1200$	B	500$	C	800$	D	700$

3. Estimate the third quartile for the amount spent.

A	1200$	B	500$	C	800$	D	700$

4. Estimate the interquartile range for the amount spent.

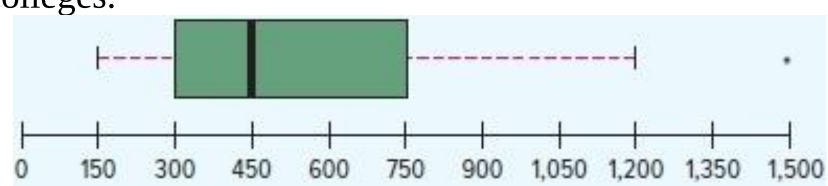
A	1200$	B	500$	C	800$	D	700$

5. Beyond what point is a value considered an outlier?

A	0$	B	1800$	C	800$	D	450$

Question 3:

The box plot shows the undergrad graduate in-state tuition per credit hour at four-year public colleges.



1. Estimate the median.

A	1200	B	1500	C	750	D	450

Question 4:

Listed are the commission earned ($000) last year by 15 sales representatives at Furniture Patch Incorporated.

3.9	5.7	8	11.3	12.7	13.6	14.4	15.2	17	17.4	18.3	22.3	35.4	43.2	78.2

1. What is the minimum?

A	78.2	B	15.2	C	3.9	D	11.3

2. What is the maximum?

A	78.2	B	15.2	C	3.9	D	11.3

3. What is the median.

A	22.3	B	15.2	C	14.4	D	11.3

4. What is the first quartile.

A	22.3	B	15.2	C	14.4	D	11.3

5. What is the third quartile.

A	22.3	B	15.2	C	14.4	D	11.3

Question 5:

For the following data:

46	47	49	49	51	53	54	54	55	55	59

1. What is the median.

A	46	B	53	C	49	D	55

2. What is the first quartile.

A	46	B	53	C	49	D	55

3. What is the third quartile.

A	46	B	53	C	49	D	55

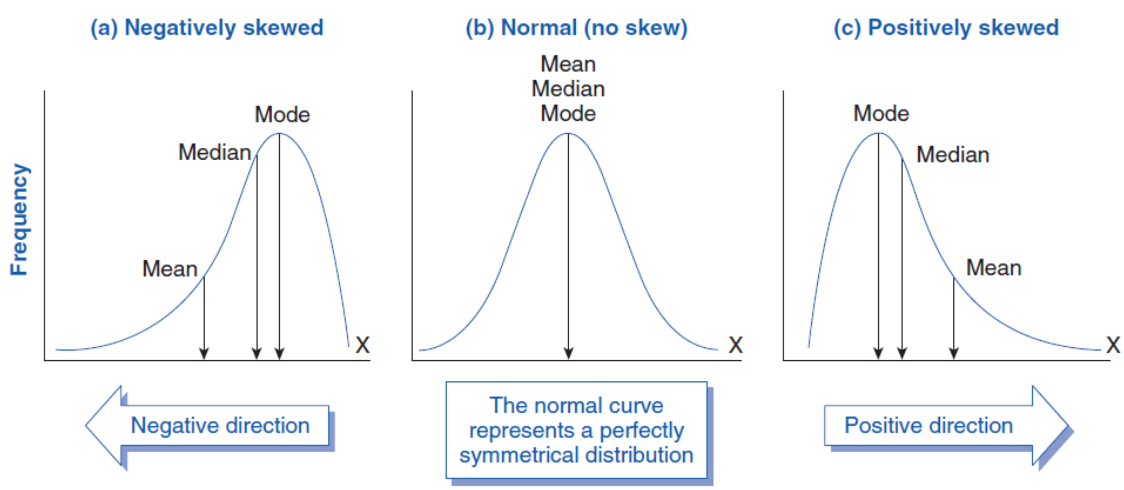
- **<u>Coefficient of Skewness</u>**

$$\text{Pearson: } S_k = \frac{3(\bar{x} - Median)}{s}$$

- Between −3 and 3
- Near −3: considerable negative skewness.
- Near 0: symmetric
- Near 3: considerable positive skewness.

$$\text{Software skewness } S_k = \frac{n}{(n-1)(n-2)}\left[\sum\left(\frac{x-\bar{x}}{s}\right)^3\right]$$

- Focus on the brackets with the mean and standard deviation.
- Called standardizing (discussed later).
- Values are larger than the mean: positive and skewed right.
- Values are less than the mean: negative and skewed left.



Question 1:

Suppose we have data on the annual income (in thousands of riyals) of 10 households in a neighborhood:

Income Data: 30, 35, 45, 50, 55, 60, 70, 85, 100, 150

Find Pearson's skewness:

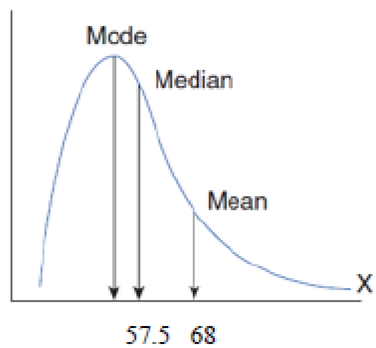
$$\bar{x} = \frac{30+35+45+50+55+60+70+85+100+150}{10} = 68$$

$$\text{Median} = \frac{55+60}{2} = 57.5$$

$$S = \sqrt{\frac{(x-\bar{x})^2}{n-1}} = \sqrt{\frac{(30-68)^2+(35-68)^2+(45-68)^2+\cdots+(150-68)^2}{10-1}} = 35.99383$$

$$S_k = \frac{3(\bar{x}-\text{Median})}{S} = \frac{3(68-57.5)}{35.99383} = 0.88$$

Pearson's Skewness of 0.88 indicates positive skew. This suggests that (Mean = 68 > Median = 57.5) and the distribution has a longer right tail.



Question 2:

Suppose a company conducts a customer satisfaction survey and collects ratings on a scale from 1 to 10 from 15 customers. The ratings data are as follows:

Customer Satisfaction Ratings:8,8,8,9,9,9,9,10,10,10,10,10,10,10,10

Find Pearson's skewness:

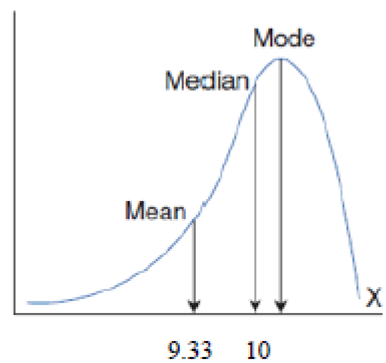
$$\bar{x} = \frac{8+8+8+9+9+9+9+10+10+10+10+10+10+10+10}{15} = 9.33$$

$$\text{Median} = 10$$

$$S = \sqrt{\frac{(x-\bar{x})^2}{n-1}} = \sqrt{\frac{3\times(8-9.33)^2+4\times(9-9.33)^2+8\times(10-9.33)^2}{15-1}} = 0.82$$

$$S_k = \frac{3(\bar{x}-\text{Median})}{S} = \frac{3(9.33-10)}{0.82} = -2.45$$

Pearson's Skewness of -2.45 indicates negative skew. This suggests that (Mean $= 9.33 >$ Median $= 10$) and the distribution has a longer left tail.



- **Correlation Coefficient:**

Measures the direction and strength of the relationship

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{(n-1)s_x s_y} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}}$$

- Ranges of correlation coefficient $(-1, +1)$.
- The closer the coefficient is to −1 or +1, the stronger the relationship.
- If r is close to 0, we can say that there is no relationship.
- Positive indicates a positive relationship.
- Negative indicates a negative relationship.

Question 1:

Suppose we have data on the number of hours studied by 5 students and their corresponding test scores:

Hours Studied (X)	2	3	5	6	8
Test Score (Y)	65	70	75	80	85

Find the correlation coefficient:

n	X	Y	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$
1	2	65	$(2-4.8)=-2.8$	$(65-75)=-10$	$-2.8\times-10=28$	7.84	100
2	3	70	$(3-4.8)=-1.8$	$(70-75)=-5$	$-1.8\times-5=9$	3.24	25
3	5	75	$(5-4.8)=0.2$	$(75-75)=0$	$0.2\times0=0$	0.04	0
4	6	80	1.2	5	6	1.44	25
5	8	85	3.2	10	32	10.24	100
Total	24	375			75	22.8	250

$\bar{x} = \frac{24}{5} = 4.8$, $\bar{y} = \frac{375}{5} = 75$

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} = \frac{75}{\sqrt{22.8\times 250}} = 0.993$$

Question 2:

An economist wants to study the relationship between a country's GDP growth rate and its unemployment rate. The economist collects data from 7 countries:

Country	GDP Growth Rate (%) (X)	Unemployment Rate (%) (Y)
A	3	8
B	5	8
C	2	9
D	4	6
E	1	10
F	6	6
G	0	9

Find the correlation coefficient:

n	X	Y	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$
A	3	8	0	0	0	0	0
B	5	8	2	0	0	4	0
C	2	9	-1	1	-1	1	1
D	4	6	1	-2	-2	1	4
E	1	10	-2	2	-4	4	4
F	6	6	3	-2	-6	9	4
G	0	9	-3	1	-3	9	1
Total	21	56			-16	28	14

$\bar{x} = \frac{21}{7} = 3$, $\bar{y} = \frac{56}{7} = 8$

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} = \frac{-16}{\sqrt{28 \times 14}} = -0.808$$

- **Contingency tables:**

Question 1:

Here is a table showing the number of employed and unemployed workers 20 years or older by gender in Saudi Arabia:

Gender	Number of Workers (000)	
	Employed	Unemployed
Men	70415	4209
Women	61402	3314

1. How many workers were studied?

A	131817	B	139340	C	74624	D	64716

2. What percent of the workers were unemployed?

A	7523	B	131817	C	5.4	D	94.7

3. What percent of men are unemployed?

A	94.4	B	70415	C	4209	D	5.6

4. What percent of women are unemployed?

A	94.9	B	61402	C	3314	D	5.1

Question 2:

A marketing manager wants to analyze customer preferences based on age groups for two types of smartphones: iPhone and Android. The data is collected from a survey of 200 customers, categorized by their age group (Child, Teenager, and Adult) and their preferred smartphone.

Age Group	Smart phone		
	iPhone	Android	Total
Child	20	30	50
Teenager	40	30	70
Adult	30	50	80
Total	90	110	200

1. How many customers prefer Android phones in total?

A	80	B	100	C	110	D	120

2. What is the total number of teenagers surveyed?

A	50	B	60	C	70	D	80

3. How many adults prefer iPhones?

A	20	B	30	C	40	D	50

4. What is the total number of customers who prefer iPhones?

A	80	B	90	C	100	D	110

5. How many children prefer Android phones?

A	20	B	30	C	40	D	50

6. What percentage of the total surveyed customers are adults?

A	30%	B	35%	C	40%	D	45%

7. Which age group has the highest total number of customers?

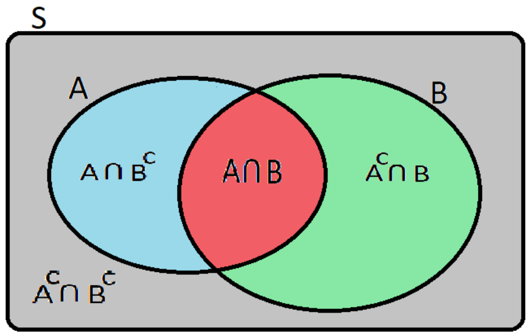
A	Child	B	Teenager	C	Adult	D	None

8. How many customers were surveyed?

A	150	B	200	C	250	D	300

Chapter 5: Probability

Probability



Definitions and Theorems:

- $0 \leq P(A) \leq 1$
- $P(S) = 1$
- $P(\emptyset) = 0$

1- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
2- $P(A \mid B) = P(A \cap B)/P(B)$
3- $P(A \cap B) = P(A) \times P(B)$ (if A & B are independent)
4- $P(A \cap B) = 0$ (if A & B are disjoint)
5- $P(A^c) = 1 - P(A)$; $P(A^c) = P(\bar{A}) = P(A')$

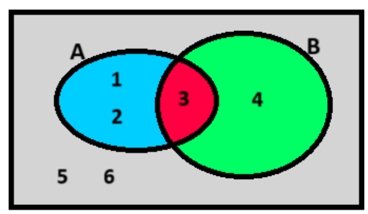
- $P(A \mid B) = P(A)$
- $P(B \mid A) = P(B)$

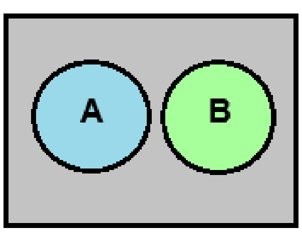
Suppose the sample space $S = \{1,2,3,4,5,6\}$

If $A = \{1,2,3\}$ and $B = \{3,4\}$

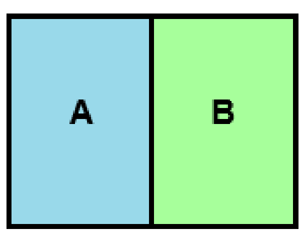
Then, $A \cap B = \{3\}$

$A \cup B = \{1,2,3,4\}$

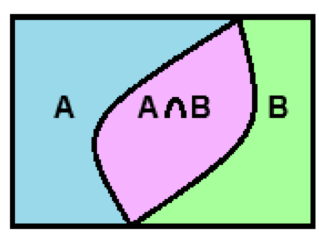




Disjoint
or
Mutually exclusive



Exhaustive
and
Disjoint



Exhaustive

Question 1:

Suppose that we have: $P(A) = 0.4$, $P(B) = 0.5$ and $P(A \cap B) = 0.2$

1. The probability $P(A \cup B)$ is:

$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.4 + 0.5 - 0.2 = 0.7}$$

2. The probability $P(A \cap B^c)$ is:

$$\boxed{P(A \cap B^c) = P(A) - P(A \cap B) = 0.4 - 0.2 = 0.2}$$

3. The probability $P(A^c \cap B)$ is:

$$\boxed{P(A^c \cap B) = P(B) - P(A \cap B) = 0.5 - 0.2 = 0.3}$$

4. The probability $P(A|B)$ is:

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.2}{0.5} = 0.4}$$

5. The events A and B are:

$$\boxed{P(A \cap B) \stackrel{?}{=} P(A) \times P(B) \Longrightarrow 0.2 = 0.4 \times 0.5}$$

A	Disjoint	B	Dependent	C	Equal	D	Independent

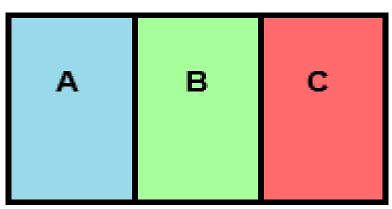
Question 2:

Consider three events A, B and C such that:

$$P(A) = 0.4, P(B) = 0.5 \text{ and } P(C) = 0.1$$

If A, B and C are disjoint events, then they are:

A	Symmetric	B	Exhaustive	C	Not exhaustive	D	None of them



Question 3:

If the events A, B we have: P(A) = 0.2, P(B) = 0.5 and P (A ∩ B) = 0.1, then:

1. The events A, B are:

$$P(A \cap B) \overset{?}{=} P(A) \times P(B) \Longrightarrow 0.1 = 0.2 \times 0.5$$

A	Disjoint	B	Dependent	C	Both are empties	D	Independent

2. The probability of A or B is:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.2 + 0.5 - 0.1 = 0.6$$

3. If P(A) = 0.3, P(B) = 0.4 and that A and B are disjoint, then P(A∪B) =

$$P(A \cup B) = P(A) + P(B) - 0 = 0.3 + 0.4 - 0 = 0.7$$

4. If P(A) = 0.2 and P(B | A) =0.4, then P(A ∩ B) =

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \Longrightarrow 0.4 = \frac{P(A \cap B)}{0.2} \Longrightarrow P(A \cap B) = 0.2 \times 0.4 = 0.08$$

5. Suppose that the probability a patient smoke is 0.20. If the probability that the patient smokes and has a lung cancer is 0.15, then the probability that the patient has a lung cancer given that the patient smokes is

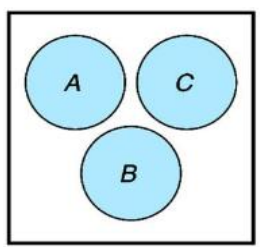
$$P(S) = 0.20 \quad P(S \cap C) = 0.15 \quad P(C|S) = ?$$

$$P(C|S) = \frac{P(C \cap S)}{P(S)} = \frac{0.15}{0.20} = 0.75$$

Question 4:

The probability of three mutually exclusive events A, B and C are given by $1/3$, $1/4$ and $5/12$ then $P(A \cup B \cup C)$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$
$$= \frac{1}{3} + \frac{1}{4} + \frac{5}{12} = 1$$



A	0.57	B	0.43	C	0.58	D	1

Question 5:

Suppose that we have two events A and B such that,

$$P(A) = 0.4, P(B) = 0.5, P(A \cap B) = 0.2.$$

[1] $P(A \cup B)$: $\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.4 + 0.5 - 0.2 = 0.7}$

[2] $P(A^c \cap B)$: $\boxed{P(A^c \cap B) = P(B) - P(A \cap B) = 0.5 - 0.2 = 0.3}$

[3] $P(A^c \cap B^c)$: $\boxed{P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - 0.7 = 0.3}$

[4] $P(A^c)$: $\boxed{P(A^c) = 1 - P(A) = 1 - 0.4 = 0.6}$

[5] $P(A^c \mid B)$: $\boxed{P(A^c \mid B) = \frac{P(A^c \cap B)}{P(B)} = \frac{0.3}{0.5} = 0.6}$

[6] $P(B \mid A)$: $\boxed{P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0.2}{0.4} = 0.5}$

[7] The events A and B are …

$$\boxed{P(A \cap B) \stackrel{?}{=} P(A) \times P(B) \Longrightarrow 0.2 = 0.4 \times 0.5}$$

A	Exhaustive	B	Dependent	C	Equal	D	Independent

Question 6:

Let A and B two events defined on the same sample space.

If $P(A) = 0.7, P(B) = 0.3$

1. If the events A and B are mutually exclusive (disjoint) then, the value of $P(A \cup B)$.

$$P(A \cup B) = P(A) + P(B) - 0 = 0.7 + 0.3 - 0 = 1$$

A	0.21	B	0.52	C	0.79	D	1

2. If the events A and B are independent, then the value of $P(A|B)$.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A) = 0.7$$

A	0.3	B	0.5	C	0.7	D	0.9

3. If the events A and B are independent, then the value of $P(A \cap \bar{B})$.

$$P(A \cap \bar{B}) = P(A)P(\bar{B}) = (0.7)(0.7)$$

A	0.09	B	0.21	C	0.49	D	0.54

4. If the events A and B are independent, then the value of $P(\overline{A \cup B})$.

$$P(\overline{A \cup B}) = 1 - P(A \cup B)$$

$$= 1 - [0.7 + 0.3 - (0.7)(0.3)] = 0.21$$

A	0.21	B	0.39	C	0.49	D	0.54

Question 7:

The following table is for 80 clients at Al Rajhi Bank, classified according to gender and the stock market (سوق الاسهم) in which they invest (Saudi, American and Gulf):

Sex	Stock market		
	Saudi	American	Gulf
Male (M)	25	17	15
Female (F)	11	9	3

1) The probability that a client selected randomly is a male and invest in Saudi stock market is:

A	25/36	B	25/80	C	25/57	D	52/80

2) The probability that a client selected randomly is a female is

A	6/80	B	40/80	C	23/80	D	None

3) A company offers two different products, Product A and Product B. According to market research, 60% of customers purchase Product A, and 50% of customers purchase Product B. It is also found that 30% of customers purchase both products. What is the probability that a customer will purchase either Product A or Product B?

$$P(A) = 0.60 \quad P(B) = 0.50 \quad P(A \cap B) = 0.30 \quad P(A \cup B) = ?$$
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$P(A \cup B) = 0.60 + 0.50 - 0.30 = 0.80$$

Question 8:

Gender	Use stcpay (S)	Use urpay (U)	TOTAL
Male (M)	72	288	360
Female (F)	48	192	240
Total	120	480	600

Consider the information given in the table above. A person is selected randomly

1. The probability that the person found is male and use stcpay is:

$$P(M \cap S) = \frac{72}{600} = 0.12$$

2. The probability that the person found is male or use stcpay is:

$$P(M \cup S) = P(M) + P(S) - P(M \cap S) = \frac{360}{600} + \frac{120}{600} - \frac{72}{600} = \frac{408}{600}$$

3. The probability that the person found is female is:

$$P(F) = \frac{240}{600} = 0.4$$

4. Suppose we know the person found is a male, the probability that he used stcpay, is:

$$P(S|M) = \frac{P(M \cap S)}{P(M)} = \frac{72/600}{360/600} = \frac{72}{360} = 0.2$$

5. The events M and S are:

$$P(M \cap S) = P(M) \times P(S) \Longrightarrow \frac{72}{600} = \frac{360}{600} \times \frac{120}{600}$$
$$\Longrightarrow \frac{72}{600} = \frac{72}{600}$$

A	Mutually exclusive	B	Dependent	C	Equal	D	Independent

Question 9:

A group of QUA107 students, classified according to their grade and their weekly study hours:

Grade	Weekly study hours			Total
	2 hours (T)	3 hours (R)	4 hours (F)	
A	80	35	20	135
B	25	110	45	180
C	15	95	75	185
Total	120	240	140	500

If one of these students is randomly chosen give:

1. The event "(study for 2 hours) and (grade B) ", is defined as.

A	$T \cup B^c$	B	$T \cap B$	C	$T \cup C$	D	$S \cup C$

2. $P(A \cup F) =$

A	0.51	B	0.28	C	0.27	D	0.04

3. $P(B \cap R) =$

A	0.48	B	0.36	C	0.22	D	0.62

4. $P(C^c) =$

A	0.63	B	0.37	C	0.50	D	1

5. $P(B|R) =$

A	0.6111	B	0.2200	C	0.4583	D	0.36

6. $P(F|C) =$

A	0.6111	B	0.2200	C	0.405	D	0.36

Question 10:

In a study of real estate investments, recommendations were taken from 500 real estate consultants from different cities (Riyadh, Jeddah, and Dammam), and the results were as follows:

Recommendations of consultant	City			Total
	Riyadh	Jeddah	Dammam	
Recommend	142	108	135	385
Not recommend	43	42	30	115
Total	185	150	165	500

1. What is the probability that a randomly selected consultant will recommend investing in real estate?

A	0.95	B	0.42	C	0.53	D	0.77

2. Given that the consultant is from Dammam, what is the probability that he will not recommend investing in real estate?

A	0.182	B	0.325	C	0.435	D	0.546

3. What is the probability that a randomly selected consultant will be from Jeddah and recommend investing in real estate?

A	0.115	B	0.523	C	0.216	D	0.756

4. The events {the consultant recommend investing in real estate} and {the consultant not recommend investing in real estate} are:

A	Mutually exclusive	B	Symmetric	C	Equal	D	None of them

5. What is the probability that a consultant chosen at random is not from Riyadh is:

A	0.46	B	0.75	C	0.63	D	0.86

6. Given that the consultant does not recommend investing in real estate, what is the probability that he is from Riyadh?

A	0.115	B	0.643	C	0.374	D	0.756

Question 11:

The following table classifies a sample of individuals according to gender and period (in years) attendance in the college:

Collage attended	Gender		Total
	Male	Female	
None	12	41	53
Two years	14	63	77
Three years	9	49	58
Four years	7	50	57
Total	42	203	245

Suppose we select an individual at random, then:

1. The probability that the individual is male is:

A	0.8286	B	0.1714	C	0.0490	D	0.2857

2. The probability that the individual did not attend college (None) and female is:

A	0.0241	B	0.0490	C	0.1673	D	0.2163

3. The probability that the individual has three year or two-year college attendance is:

A	0.551	B	0.0939	C	0.4571	D	0

4. If we pick an individual at random and found that he had three-year college attendance, the probability that the individual is male is:

A	0.0367	B	0.2143	C	0.1552	D	0.1714

5. The probability that the individual is not a four-year college attendance is:

A	0.7673	B	0.2327	C	0.0286	D	0.1429

6. The probability that the individual is a two-year college attendance or male is:

A	0.0571	B	0.8858	C	0.2571	D	0.4286

7. The events: the individual is a four-year college attendance and male are:

A	Mutually exclusive	B	Independent	C	Dependent	D	None of these

Question 12:

In a study of several companies, the level of risk was measured compared to the number of years of experience of its CEO, and the results were as follows:

CEOs experience	Level of risk			
	Low (L)	Medium (M)	High (H)	
< 5years (B)	25	17	15	57
≥ 5years (B')	11	9	3	23
	36	26	18	80

If a company is selected at random, then the probability that:

1. The CEO have experience less than 5 years or the company has medium risk is equal to

$$P(B \cup M) = P(B) + P(M) - P(B \cap M) = \frac{57}{80} + \frac{26}{80} - \frac{17}{80} = \frac{66}{80} = 0.825$$

A	0.442	B	0.50	C	0.725	D	0.825

2. The company has low risk given that the CEO have experience less than 5 years is equal to

$$P(L \mid B) = \frac{P(L \cap B)}{P(B)} = \frac{25/80}{57/80} = 0.4386$$

A	0.90	B	0.1	C	0.66	D	0.44

Bayes Rule:

If the events $A_1, A_2, \ldots,$ and A_n constitute a partition of the sample space S such that $P(A_i) \neq 0$ for $i = 1, 2, \ldots\ n$, then for any event B:

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B \mid A_i) = \sum_{i=1}^{n} P(A_i \cap B)$$

Note:

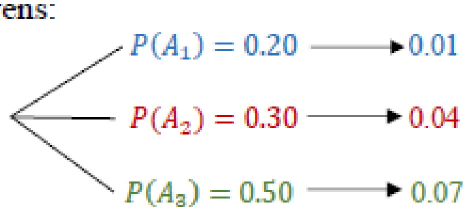
$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(B)}$$

Question 13:

Three machines A_1, A_2, and A_3 make 20%, 30%, and 50%, respectively, of the products. It is known that 1%, 4%, and 7% of the products made by each machine, respectively, are defective. If a finished product is randomly selected

a) What is the probability that it is defective (D)?

Givens:



$$P(D) = P(A_1) \times P(D \mid A_1) + P(A_2) \times P(D \mid A_2) + P(A_3) \times P(D \mid A_3)$$
$$= 0.20 \times 0.01 + 0.30 \times 0.04 + 0.50 \times 0.07 = 0.049$$

b) If it is known that the selected product is defective, what is the probability that it is made by machine A_1?

$$P(A_1 \mid D) = \frac{0.20 \times 0.01}{0.049} = 0.0408$$

c) If it is known that the selected product is defective, what is the probability that it is made by machine A_2?

$$P(A_2 \mid D) = \frac{0.30 \times 0.04}{0.049} = 0.2449$$

d) If it is known that the selected product is defective, what is the probability that it is made by machine A_3?

$$P(A_3 \mid D) = \frac{0.50 \times 0.07}{0.049} = 0.7134$$

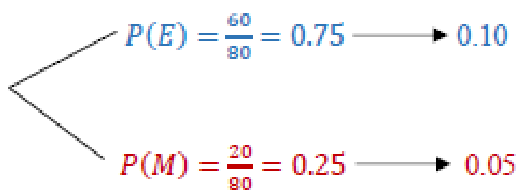
Question 14:

80 students are enrolled in the QUA107 class. 60 students are from the Economics department, and the rest are from the Marketing department. 10% of the Economics department students have taken this course before, and 5% of the Marketing department students have taken this course before. If one student from this class is randomly selected,

a) What is the probability that he has taken this course before?

$P(Economics) = P(E)$
$P(Marketing) = P(M)$
$P(Taking\ this\ course\ bsfore) = P(T)$

Givens:



$$P(T) = P(E) \times P(T \mid E) + P(M) \times P(T \mid M)$$
$$= \ 0.75 \times 0.10 \ + \ 0.25 \times 0.05 \ = 0.0875$$

b) If the selected student has taken this course before, what is the probability that he is from the Economics department?

$$P(E \mid T) = \frac{0.75 \times 0.10}{0.0875} = 0.8571$$

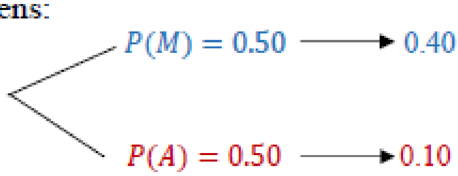
Question 15:

Two brothers, Mohammad and Ahmad own and operate a small restaurant. Mohammad washes 50% of the dishes and Ahmad washes 50% of the dishes. When Mohammad washes a dish, he might break it with probability 0.40. On the other hand, when Ahmad washes a dish, he might break it with probability 0.10.

i) What is the probability that a dish will be broken during washing?

$P(Ahmad) = P(A)$
$P(Mohammad) = P(M)$
$P(Broken\ dish) = P(B)$

Givens:



$$P(T) = P(M) \times P(B \mid M) + P(A) \times P(B \mid A)$$
$$= 0.50 \times 0.40 + 0.50 \times 0.10 = 0.25$$

ii) If a broken dish was found in the washing machine, what is the probability that it was washed by Mohammad?

$$P(M \mid B) = \frac{0.50 \times 0.40}{0.25} = 0.8$$

iii) If a broken dish was found in the washing machine, what is the probability that it was washed by Ahmad?

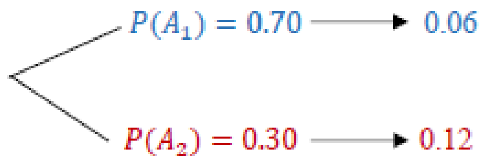
$$P(A \mid B) = \frac{0.50 \times 0.10}{0.25} = 0.2$$

Question 16:

Two machines A_1 and A_2 make 70%, and 30%, respectively, of the products. It is known that 6%, and 12 % of the products made by each machine, respectively, are defective. If a finished product is randomly selected

a) What is the probability that it is defective?

Givens:



$$P(T) = P(A_1) \times P(D \mid A_1) + P(A_2) \times P(D \mid A_2)$$
$$= 0.70 \times 0.06 + 0.30 \times 0.12 = 0.078$$

b) If it is known that the selected product is defective, what is the probability that it is made by machine A_2?

$$P(A_2 \mid D) = \frac{0.30 \times 0.12}{0.078} = 0.4615$$

c) If it is known that the selected product is defective, what is the probability that it is made by machine A_1?

$$P(A_1 \mid D) = \frac{0.70 \times 0.06}{0.078} = 0.5385$$

More Exercises

Question 1:

Givens:

$$\boldsymbol{P(A) = 0.5,\quad P(B) = 0.4,\quad P(C \cap A^c) = 0.6,}$$
$$\boldsymbol{P(C \cap A) = 0.2,\quad P(A \cup B) = 0.9}$$

(a) ***What is the probability of*** $\boldsymbol{P(C)}$**:**

$$P(C) = P(C \cap A^c) + P(C \cap A) = 0.6 + 0.2 = 0.8$$

(b) ***What is the probability of*** $\boldsymbol{P(A \cap B)}$**:**

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$\Rightarrow \quad 0.9 = 0.5 + 0.4 - P(A \cap B)$$
$$P(A \cap B) = 0$$

(c) ***What is the probability of*** $\boldsymbol{P(C \mid A)}$**:**

$$P(C \mid A) = \frac{P(C \cap A)}{P(A)} = \frac{0.2}{0.5} = 0.4$$

(d) ***What is the probability of*** $\boldsymbol{P(B^c \cap A^c)}$:

$$P(B^c \cap A^c) = 1 - P(B \cup A) = 1 - 0.9 = 0.1$$

Question 2:

Givens:

$$\boldsymbol{P(B) = 0.3,\quad P(A \mid B) = 0.4}$$

$\boldsymbol{Then\ find\ P(A \cap B)} = ?$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$
$$\Rightarrow 0.4 = \frac{P(A \cap B)}{0.3}$$
$$\Rightarrow P(A \cap B) = 0.4 \times 0.3 = 0.12$$

Question 3:
Givens:

$$\boldsymbol{P(A) = 0.3,\quad P(B) = 0.4,\quad P(A \cap B \cap C) = 0.03,\quad P(\overline{A \cap B}) = 0.88}$$

(1) Are the event A and b independent?

$$P(A \cap B) = 1 - P(\overline{A \cap B}) = 1 - 0.88 = 0.12$$

$$P(A) \times P(B) = 0.3 \times 0.4 = 0.12$$
$$\Rightarrow P(A \cap B) = P(A) \times P(B)$$
Therefore, A and B are independent.

(2) ***What is the probability of*** $\boldsymbol{P(C \mid A \cap B)}$:

$$P(C \mid A \cap B) = \frac{P(A \cap B \cap C)}{P(A \cap B)} = \frac{0.03}{0.12} = 0.25$$

Question 4:
Givens:
$$\boldsymbol{P(A_1) = 0.4,\quad P(A_1 \cap A_2) = 0.2,\quad P(A_3 \mid A_1 \cap A_2) = 0.75}$$

(1) ***Find the*** $\boldsymbol{P(A_2|A_1)}$:

$$P(A_2|A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)} = \frac{0.2}{0.4} = 0.5$$

(2)***Find the*** $\boldsymbol{P(A_1 \cap A_2 \cap A_3)}$:

$$P(A_3 \mid A_1 \cap A_2) = \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)}$$

$$0.75 = \frac{P(A_1 \cap A_2 \cap A_3)}{0.2}$$

$$P(A_1 \cap A_2 \cap A_3) = 0.75 \times 0.2 = 0.15$$

Exercise 1:

A group of 400 people are classified according to their nationality as (250 Saudi and 150 non-Saudi), and they are classified according to their gender (100 Male and 300 female). The number of Saudi males is 60. Suppose that the experiment is to select a person at random from this group.

1. Summarizing the information in a table:

		Gender		Total
		Male (M)	Female (F)	
Nationality	Saudi (S)	**60**	190	**250**
	Non-Saudi (N)	40	110	**150**
Total		**100**	**300**	**400**

2. The probability that the selected person is Saudi is:
(A) 0.6
(B) 0.15
(C) 0.3158
(D) 0.625 (P(S) = 250/400)

3. The probability that the selected person is female is:
(A) 0.375
(B) 0.75 (P(F) = 300/400)
(C) 0.3667
(D) 0.6333

4. The probability that the selected person is female given that the selected person is Saudi is:
(A) 0.6333
(B) 0.3667
(C) 0.76 (P(F | S) = 190/250)
(D) 0.475

5. The events "S"={Selecting a Saudi} and "F"={Selecting a female} are:
(A) Not independent events (Because: $P(F) \neq P(F \mid S)$)
(B) Complement of each other
(C) Independent events
(D) Disjoint (mutually exclusive) events

Chapter 6:
Discrete Probability Distribution

Random Variables

$$\text{Random Variables} \begin{cases} \text{Discrete Random Variables} \\ \text{Continuous Random Variables} \end{cases}$$

• Probability distribution function (PDF) $f(x) = P(X = x)$	• Cumulative distribution function (CDF) $F(x) = P(X \leq x)$

• Properties of discreate random variable $0 \leq P(x) \leq 1$ $\sum P(x) = 1$	• The expected value (the mean) of discreate random variables $E(X) = \mu = \sum x\, P(x)$ • The variance of discreate random variable $Var(X) = \sigma^2 = \sum (x - \mu)^2\ P(x) = E(X^2) - \mu^2$

Question 1:

Given the following discrete probability distribution:

x	5	6	7	8
f(x)=P(x)	0.35	0.45	0.15	k

Find:

1. The value of k.

$$0.35 + 0.45 + 0.15 + k = 1 \implies \boxed{k = 0.05}$$

x	5	6	7	8
f(x)=P(x)	0.35	0.45	0.15	0.05

2. $P(X > 6) = 0.05 + 0.15 = 0.20$
3. $P(X \geq 6) = 0.45 + 0.15 + 0.05 = 0.65 \quad or \quad (1 - 0.35 = 0.65)$
4. $P(X < 4) = 0$
5. $P(X > 3) = 1$

Question 2:

Which of the following functions can be a probability distribution of a discrete random variable?

(a)

x	P(x)
0	0.6
1	-0.2
2	0.5
3	0.1
✘	

(b)

x	P(x)
0	0.4
1	0.1
2	0.5
3	0.2
✘	

(c)

x	P(x)
0	0.1
1	1.2
2	-0.6
3	0.3
✘	

(d)

x	P(x)
0	0.3
1	0.1
2	0.5
3	0.1
✓	

(e)

x	P(x)
0	0.2
1	0.4
2	0.3
3	0.4
✘	

(f)

x	P(x)
0	0.1
1	0.2
2	0.3
3	0.1
✘	

<u>Question 3</u>:
Which of the following is a probability distribution function:

$a. f(x) = \frac{x+1}{10}$; $x = 0,1,2,3,4$

$b. f(x) = \frac{x-1}{5}$; $x = 0,1,2,3,4$

$c. f(x) = \frac{1}{5}$; $x = 0,1,2,3,4$

$d. f(x) = \frac{5-x^2}{6}$; $x = 0,1,2,3$

$a.$

$$f(x) = \frac{x+1}{10};\ x = 0, 1, 2, 3, 4$$

x	0	1	2	3	4
$f(x)$	$1/10$	$2/10$	$3/10$	$4/10$	$5/10$

$f(x)$ *is not a P.D.F because* $\sum f(x) \neq 1$

$b.$

$$f(x) = \frac{x-1}{5};\ x = 0, 1, 2, 3, 4$$

x	0	1	2	3	4
$f(x)$	$-1/5$				

$f(x)$ *is not a P.D.F because every* $f(x)$ *shoud be* $0 \leq f(x) \leq 1$

$c.$

$$f(x) = \frac{1}{5};\ x = 0, 1, 2, 3, 4$$

x	0	1	2	3	4
$f(x)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$

$f(x)$ *is a P.D.F*

$d.$

$$f(x) = \frac{5-x^2}{6};\ x = 0, 1, 2, 3$$

x	0	1	2	3
$f(x)$				$-4/6$

$f(x)$ *is not a P.D.F because every* $f(x)$ *shoud be* $0 \leq f(x) \leq 1$

Question 4:

Given the following discrete probability distribution:

x	5	6	7	8
f(x)=P(x)	2k	3k	4k	k

Find the value of k.

$$2k + 3k + 4k + k = 1$$

$$10k = 1 \implies \boxed{k = 0.1}$$

x	5	6	7	8
f(x)=P(x)	0.2	0.3	0.4	0.1

Question 5:

Let X be a discrete random variable with probability mass function:
$f(x) = cx$; $x = 1,2,3,4$ What is the value of c?

x	1	2	3	4
P(x)	c	2c	3c	4c

$$c + 2c + 3c + 4c = 1 \implies c = \frac{1}{10}$$

Then probability mass function:

x	1	2	3	4
P(x)	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$

Question 6:

Let X be a discrete random variable with probability function given by:

$$f(x) = c(x^2 + 2) \;\; ; \quad x = 0,1,2,3$$

$f(0) = c(0^2 + 2) = 2c$

$f(1) = c(1^2 + 2) = 3c$

$f(2) = c(2^2 + 2) = 6c$

$f(3) = c(3^2 + 2) = 11c$

x	0	1	2	3
P(x)	2c	3c	6c	11c

$$2c + 3c + 6c + 11c = 1 \qquad c = \frac{1}{22} = 0.04545$$

x	0	1	2	3
P(x)	$\frac{2}{22}$	$\frac{3}{22}$	$\frac{6}{22}$	$\frac{11}{22}$

Question 7:

Given the following discrete probability distribution:

x	5	6	7	8
f(x)=P(x)	0.2	0.4	0.3	0.1

Find:

1. Find the mean of the distribution $\mu = \mu_X = E(X)$.

$$\mu = \sum_{x=5}^{8} x\,P(x)$$
$$= (5)(0.2) + (6)(0.4) + (7)(0.3) + (8)(0.1) = 6.3$$

2. Find the variance of the distribution $\sigma^2 = \sigma_X^2 = Var(X)$.

$$E(X^2) = (5^2 \times 0.2) + (6^2 \times 0.4) + (7^2 \times 0.3) + (8^2 \times 0.1) = 40.5$$

$$\sigma^2 = E(X^2) - \mu^2$$
$$= 40.5 - 6.3^2 = 0.81$$

Or:

$$\sigma^2 = \sum (x - \mu)^2\,P(x)$$
$$= \sum (x - 6.3)^2\,P(x)$$
$$= (5 - 6.3)^2(0.2) + (6 - 6.3)^2(0.4) + (7 - 6.3)^2(0.3) + (8 - 6.3)^2(0.1) = 0.81$$

Question 8:

Given the following discrete distribution:

x	-1	0	1	2	3	4
P(x)	0.15	0.30	M	0.15	0.10	0.10

1. The value of M is equal to

$$M = 1 - (0.15 + 0.30 + 0.15 + 0.10 + 0.10) = 1 - 0.80 = 0.20$$

2. P (X ≤0.5) = $0.15 + 0.30 = 0.45$

3. P(X=0) = 0.30

4. The expected (mean) value E[X] is equal to

$$\mu = (-1 \times 0.15) + (0 \times 0.30) + (1 \times 0.20) + (2 \times 0.15) + (3 \times 0.10) + (4 \times 0.10) = 1.05$$

Question 9:

Assuming that the number of working days (X) required by a Saudi bank to approve a personal loan follows the following probability distribution function:

Number of days	3	4	5	6
Probability	0.4	0.2	0.1	k

(1) The value of k is

$$k = 1 - (0.4 + 0.2 + 0.1) = 1 - 0.7 = 0.3$$

(2) $P(X \leq 0) =$

$$0$$

(3) $P(0 < X \leq 5) =$

$$0.4 + 0.2 + 0.1 = 0.7$$

(4) $P(X \leq 5.5) =$

$$0.4 + 0.2 + 0.1 = 0.7$$

(5) What is the probability that the customer will wait at most 4 days to get the personal loan?

$$0.4 + 0.2 = 0.6$$

(6) What is the average number of days required for approval of a personal loan?

$$\mu = (3 \times 0.4) + (4 \times 0.2) + (5 \times 0.1) + (6 \times 0.3) = 4.3$$

Question 10:

Given the following discrete probability distribution:

x	5	6	7	8
f(x)=P(x)	0.2	0.4	0.3	0.1

1. From the discrete probability distribution of X, find:

a) $P(X \leq 7) = 0.9$

b) $P(X \leq 6.5) = 0.6$

c) $P(X > 6) = 0.4$

d) $P(X > 7) = 0.1$

Discrete Uniform Distribution

$$P(\mathrm{x}) = {}^{1}\!/_{\mathrm{k}} \ ; \quad \mathrm{x} = \mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_\mathrm{k}$$

Q1: X have discrete uniform with parameter $\mathbf{k = 3}$, $\mathbf{x = 0, 1, 2}$

x	0	1	2
P(x)	1/3	1/3	1/3

a. $\mathrm{P(X = 1)} = \frac{1}{3}$

b. $\mu = \left(0 \times \frac{1}{3}\right) + \left(1 \times \frac{1}{3}\right) + \left(2 \times \frac{1}{3}\right) = 1$

c. $\sigma^2 = \sum (\mathrm{x} - \mu)^2 \, \mathrm{P(x)}$

$= \sum (\mathrm{x} - 1)^2 \, \mathrm{P(x)}$

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

Binomial Distribution

$$P(x) = \binom{n}{x} p^x q^{n-x} \;;\quad x = 0,1 \ldots.,n$$
$$*\ \mu = np \qquad *\ \ \sigma^2 = npq$$
$$q = 1 - p$$

Question 1:

Assuming that 25% of Riyadh residents use "stc pay" for their purchases. A sample of 7 people is selected at random. Let X be the number of people in the sample who use "stc pay", follows a binomial distribution then

1) The values of the parameters of the distribution are*:*

A	7 , 0.75	B	7 , 0.25	C	0.25 , 0.75	D	25 , 7

2) The probability that we find exactly one person in the sample use "stc pay", is:

X	0	1	2	3	4	5	6	7
$P(x)$		*						

$$P(X = 1) = \binom{7}{1} (0.25)^1 (0.75)^6 = 0.31146$$

3) The probability that there will be at most one-person use "stc pay", is:

X	0	1	2	3	4	5	6	7
$P(x)$	*	*						

$$P(X \le 1) = \binom{7}{0} (0.25)^0 (0.75)^7 + \binom{7}{1} (0.25)^1 (0.75)^6 = 0.4449$$

4) The probability that we find more than one person use "stc pay", is:

X	0	1	2	3	4	5	6	7
$P(x)$			*	*	*	*	*	*

$$P(X > 1) = 1 - P(X \le 1) = 1 - 0.4449 = 0.5551$$

Question 2:

Assuming that the percentage of investors in date factories in the city of Al-Ahsa is 24%. A sample of 5 people from this city was taken. What is the probability that the number of investors in date factories in the sample will be:

$$p = 0.24 \quad , \quad n = 5$$

1. Zero:

$$P(X = 0) = \binom{5}{0}(0.24)^0(0.76)^5 = 0.2536$$

2. Exactly one

$$P(X = 1) = \binom{5}{1}(0.24)^1(0.76)^4 = 0.4003$$

3. Between one and three, inclusive

$$P(1 \leq X \leq 3) = \binom{5}{1}(0.24)^1(0.76)^4 + \binom{5}{2}(0.24)^2(0.76)^3 + \binom{5}{3}(0.24)^3(0.76)^2 = 0.7330$$

4. Two or fewer (at most two):

$$P(X \leq 2) = \binom{5}{0}(0.24)^0(0.76)^5 + \binom{5}{1}(0.24)^1(0.76)^4 + \binom{5}{2}(0.24)^2(0.76)^3 = 0.9067$$

5. Five:

$$P(X = 5) = \binom{5}{5}(0.24)^5(0.76)^0 = 0.0008$$

6. The mean of the number of investors in date factories is equal to:

$$\mu = np = 5 \times 0.24 = 1.2$$

7. The variance of the number of investors in date factories is:

$$\sigma^2 = npq = 5 \times 0.24 \times 0.76 = 0.912$$

Question 3:

The proportion of students wearing glasses is 35%. Let X the number of students wearing glasses in a random sample of 10 students. Find the following.

1) The standard deviation of X:

$$\boxed{n = 10, \;\; p = 0.35}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{npq} = \sqrt{10 \times 0.35 \times 0.65} = 1.508$$

A	3.542	B	1.508	C	4.568	D	2.275

2) The probability that $P(3 < X < 6)$ is:

X	0	1	2	3	4	5	6	7	8	9	10
$P(X = x)$					*	*					

$$P(3 < X < 6) = \binom{10}{4}(0.35)^4(0.65)^6 + \binom{10}{5}(0.35)^5(0.65)^5 = 0.391$$

A	0.013	B	0.072	C	0.391	D	0.751

3) The probability that X is at most 2 is equal to:

X	0	1	2	3	4	5	6	7	8	9	10
$P(x)$	*	*	*								

$$P(X \le 2) = \binom{10}{0}(0.35)^0(0.65)^{10} + \binom{10}{1}(0.35)^1(0.65)^9 + \binom{10}{2}(0.35)^2(0.65)^8 = 0.2616$$

A	0.752	B	0.995	C	0.854	D	0.262

More Exercises

Exercise 1:

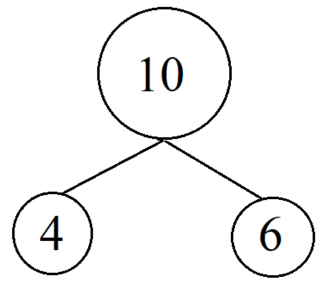
Find:

1. $6! \quad = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$
2. ${}_8C_3 \quad = \frac{8!}{3!\,(8-3)!} = \frac{8!}{3! \times 5!} = \quad 56$
3. ${}_8C_{10} \quad = 0$
4. ${}_8C_{-5} \quad = 0$

Exercise 2:

A box contains 10 cards numbered from 1 to 10. In how many ways can we select 4 cards out of this box?

$$\text{Answer} = {}_{10}C_4 = \frac{10!}{4!\,(10-4)!} = \frac{10!}{4! \times 6!}$$
$$= \frac{10\times9\times8\times7\times6!}{(4\times3\times2\times1)\,6!} = \frac{10\times9\times8\times7}{4\times3\times2\times1}$$
$$= 210$$



Exercise 3:

The manager of a certain bank has recently examined the credit card account balances for the customers of his bank and found that 20% of the customers have excellent records. Suppose that the manager randomly selects a sample of 4 customers.

(A) Define the random variable X as:

X = The number of customers in the sample having excellent records.

Find the probability distribution of X.

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

$n = 4 \quad (Number\ of\ trials)$

$p = \frac{20}{100} = 0.2 \quad (Probability\ of\ success)$

$q = 1 - p = 1 - 0.2 = 0.8 \quad (Probability\ of\ failure)$

$x = 0, 1, 2, 3, 4 \quad (Possible\ values\ of\ X)$

(a) The probability function in a mathematical formula:

$$P(X=x)=\begin{cases}\dfrac{n!}{x!\,(n-x)!}\;p^x\,q^{n-x} & ;\quad x=0,1,2,\dots,n\\ 0 & ;Otherwise\end{cases}$$

$$P(X=x)=\begin{cases}\dfrac{4!}{x!\,(4-x)!}\;(0.2)^x\,(0.8)^{4-x} & ;\quad x=0,1,2,3,4\\ 0 & ;Otherwise\end{cases}$$

(b) The probability function in a table:

x	$P(X=x)$
0	$\frac{4!}{0!\,(4-0)!}\;(0.2)^0\,(0.8)^{4-0}=(1)(0.2)^0(0.8)^4=0.4096$
1	$\frac{4!}{1!\,(4-1)!}\;(0.2)^1\,(0.8)^{4-1}=(4)(0.2)^1(0.8)^3=0.4096$
2	$\frac{4!}{2!\,(4-2)!}\;(0.2)^2\,(0.8)^{4-2}=(6)(0.2)^2\,(0.8)^2=0.1536$
3	$\frac{4!}{3!\,(4-3)!}\;(0.2)^3\,(0.8)^{4-3}=(4)(0.2)^3\,(0.8)^1=0.0256$
4	$\frac{4!}{4!\,(4-4)!}\;(0.2)^4\,(0.8)^{4-4}=(1)(0.2)^4\,(0.8)^0=0.0016$
	Total = 1

x	$P(X=x)$
0	0.4096
1	0.4096
2	0.1536
3	0.0256
4	0.0016

(B) Find:

1. The probability that there will be 3 customers in the sample having excellent records.

$$P(X = 3) = 0.0256$$

2. The probability that there will be no customers in the sample having excellent records.

$$P(X = 0) = 0.4096$$

3. The probability that there will be at least 3 customers in the sample having excellent records.

$$P(X \geq 3) = P(x = 3) + P(X = 4) = 0.0256 + 0.0016$$
$$= 0.0272$$

4. The probability that there will be at most 2 customers in the sample having excellent records.

$$P(X \leq 2) = P(x = 0) + P(X = 1) + P(X = 2)$$
$$= 0.4096 + 0.4096 + 0.1536$$
$$= 0.9728$$

5. The expected number of customers having excellent records in the sample.

$$E(X) = \mu = \mu_X = np = 4 \times 0.2 = 0.8$$

6. The variance of the number of customers having excellent records in the sample.

$$Var(X) = \sigma^2 = \sigma_X^2 = npq = 4 \times 0.2 \times 0.8 = 0.64$$

__Exercise 4__: (Do it at home for yourself)

In a certain hospital, the medical records show that the percentage of lung cancer patients who smoke is 75%. Suppose that a doctor randomly selects a sample of 5 records of lung cancer patients from this hospital.

(A) Define the random variable X as:

X = The number of smokers in the sample.

Find the probability distribution of X.

(B) Find:

1. The probability that there will be 4 smokers in the sample.
2. The probability that there will be no smoker in the sample.
3. The probability that there will be at least 2 smokers in the sample.
4. The probability that there will be at most 3 smokers in the sample.
5. The expected number of smokers in the sample.
6. The variance of the number of smokers in the sample.

Hypergeometric Distribution

$$P(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \quad ; \quad x = 0,1,\ldots.\min(n,k)$$

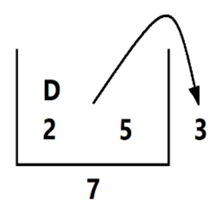
$$* \ \mu = n \times \frac{k}{N} \qquad * \qquad \sigma^2 = n \times \frac{k}{N}\left(1 - \frac{k}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Question 1:
A shipment of 7 television sets contains 2 defective sets. A hotel makes a random purchase of 3 of the sets.

(i) Find the probability distribution function of the random variable X representing the number of defective sets purchased by the hotel.

$$N = 7 \ , \ n = 3 \ , \ k = 2$$

$$f(x) = \frac{\binom{2}{x}\binom{5}{3-x}}{\binom{7}{3}} \quad ; \quad x = 0,1,2$$



(ii) Find the probability that the hotel purchased no defective television sets.

$$P(X = 0) = f(0) = \frac{\binom{2}{0}\binom{5}{3}}{\binom{7}{3}} = 0.29$$

(iii) What is the expected number of defective television sets purchased by the hotel?

$$\mu = n \times \frac{k}{N} = 3 \times \frac{2}{7} = \frac{6}{7}$$

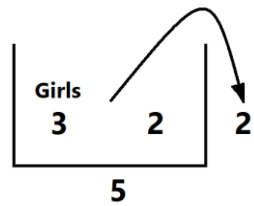
(iv) Find the variance of X.

$$\sigma^2 = n \times \frac{k}{N}\left(1 - \frac{k}{N}\right)\left(\frac{N-n}{N-1}\right) = 3 \times \frac{2}{7}\left(\frac{5}{7}\right)\left(\frac{7-3}{7-1}\right) = 0.41$$

<u>Question 2:</u>
Suppose that a family has 5 children, 3 of them are girls and the rest are boys. A sample of 2 children is selected randomly and without replacement.

$$N = 5 \quad , \quad n = 2 \quad , \quad k = 3$$

$$f(x) = \frac{\binom{3}{x}\binom{2}{2-x}}{\binom{5}{2}} \quad ; \quad x = 0,1,2$$



a. The probability that no girls are selected is

$$P(X = 0) = f(0) = \frac{\binom{3}{0}\binom{2}{2}}{\binom{5}{2}} = 0.1$$

b. The probability that <u>at most one</u> girl are selected is

$$P(X \leq 1) = f(0) + f(1) = \frac{\binom{3}{0}\binom{2}{2}}{\binom{5}{2}} + \frac{\binom{3}{1}\binom{2}{1}}{\binom{5}{2}} = 0.7$$

c. The expected number of girls in the sample is

$$\mu = n \times \frac{k}{N} = 2 \times \frac{3}{5} = \frac{6}{5}$$

d. The variance of the number of girls in the sample is

$$\sigma^2 = n \times \frac{k}{N}\left(1 - \frac{k}{N}\right)\left(\frac{N-n}{N-1}\right) = 2 \times \frac{3}{5}\left(1 - \frac{3}{5}\right)\left(\frac{5-2}{5-1}\right) = 0.36$$

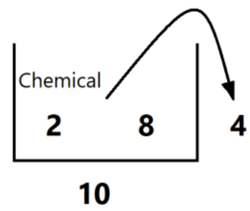
<u>Question 3:</u>

In one of the Saudi oil companies, A random committee of size 4 is selected from 2 chemical engineers and 8 industrial engineers.

(1) Write a formula for the probability distribution function of the random variable X representing the number of chemical engineers in the committee.

$$N = 10 \quad , \quad n = 4 \quad , \quad k = 2$$

$$f(x) = \frac{\binom{2}{x}\binom{8}{4-x}}{\binom{10}{4}} \quad ; \quad x = 0,1,2$$



(2) Find the probability that there will be <u>no chemical engineers</u> in the committee.

$$P(X = 0) = f(0) = \frac{\binom{2}{0}\binom{8}{4}}{\binom{10}{4}} = 0.33$$

(3) Find the probability that there will be <u>at least one chemical</u> engineer in the committee.

$$P(X \geq 1) = 1 - P(X < 1) = 1 - f(0) = 0.67$$

(4) What is the expected number of chemical engineers in the committee?

$$\mu = n \times \frac{k}{N} = 4 \times \frac{2}{10} = 0.8$$

(5) What is the variance of the number of chemical engineers in the committee?

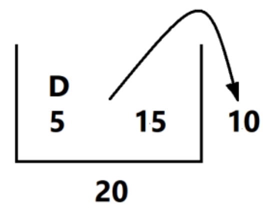
$$\sigma^2 = n \times \frac{k}{N}\left(1 - \frac{k}{N}\right)\left(\frac{N-n}{N-1}\right) = 4 \times \frac{2}{10}\left(1 - \frac{2}{10}\right)\left(\frac{10-4}{10-1}\right) = 0.43$$

Question 4:

A shipment of 20 digital voice recorders contains 5 that are defective. If 10 of them are randomly chosen (without replacement) for inspection, then:

$$N = 20 \quad , \quad n = 10 \quad , \quad k = 5$$

$$f(x) = \frac{\binom{5}{x}\binom{15}{10-x}}{\binom{20}{10}} \quad ; \quad x = 0,1,2,3,4,5$$



(1) The probability that 2 will be defective is:

$$P(X = 2) = f(2) = \frac{\binom{5}{2}\binom{15}{8}}{\binom{20}{10}} = 0.35$$

(2) The probability that at most 1 will be defective is:

$$P(X \leq 1) = f(0) + f(1) = \frac{\binom{5}{0}\binom{15}{10}}{\binom{20}{10}} + \frac{\binom{5}{1}\binom{15}{9}}{\binom{20}{10}} = 0.15$$

(3) The expected number of defective recorders in the sample is:

$$\mu = n \times \frac{k}{N} = 10 \times \frac{5}{20} = \frac{5}{2}$$

(4) The variance of the number of defective recorders in the sample is:

$$\sigma^2 = n \times \frac{k}{N}\left(1 - \frac{k}{N}\right)\left(\frac{N-n}{N-1}\right) = 10 \times \frac{5}{20}\left(\frac{15}{20}\right)\left(\frac{20-10}{20-1}\right) = 0.99$$

Poisson distribution

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

$$\mu = \sigma^2 = \lambda$$

Question 1:

The number of cases received by one of the most famous law firms (شركات المحاماه) in Jeddah follows the Poisson distribution, with an average of 10 cases per day. Then:

1. The probability that 12 cases coming in the next day, is:

$$\lambda_{one\ day} = 10$$

$$P(X = 12) = \frac{e^{-10}\ 10^{12}}{12!} = 0.09478$$

2. The average number of cases in a two days' period is:

$$\lambda_{two\ days} = 20$$

3. The probability that 20 cases coming in next two days is:

$$\lambda_{two\ nights} = 20$$

$$P(X = 20) = \frac{e^{-20}\ 20^{20}}{20!} = 0.0888$$

Question 2:

Given the mean number of serious accidents per year in a large factory is five. If the number of accidents follows a Poisson distribution, then the probability that in the next year there will be:

1. Exactly seven accidents:

$$\lambda_{one\ year} = 5$$

$$P(X = 7) = \frac{e^{-5}\ 5^7}{7!} = 0.1044$$

2. No accidents

$$P(X = 0) = \frac{e^{-5}\ 5^0}{0!} = 0.0067$$

3. one or more accidents

X	0	1	2	3	4	5	6	...
$P(X = x)$		*	*	*	*	*	*	*

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

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

$$= 1 - 0.0067 = 0.9933$$

4. The expected number (mean) of serious accidents in the next two years is equal to

$$\lambda_{two\ years} = 10$$

5. The probability that in the next two years there will be three accidents

$$\lambda_{two\ years} = 10$$

$$P(X = 3) = \frac{e^{-10}\ 10^3}{3!} = 0.0076$$

Question 3:

The number of training programs proposed for Aramco employees follows the Poisson distribution with an average of 5 programs per day, then:

1) The probability that exactly 2 programs will propose in a given day is:

$$\lambda_{one\ day} = 5$$

$$P(X = 2) = \frac{e^{-5}\ 5^{2}}{2!} = 0.0842$$

2) The average number of programs propose in 36 hours is:

$$\lambda_{36\ hours} = \lambda_{1.5\ day} = 1.5 \times 5 = 7.5$$

3) The variance of programs propose in 36 hours is:

$$\lambda_{36\ hours} = \lambda_{1.5\ day} = 1.5 \times 5 = 7.5$$

4) The standard deviation of programs propose in 36 hours is:

$$\sqrt{\lambda_{36\ hours}} = \sqrt{\lambda_{1.5\ day}} = \sqrt{1.5 \times 5} = \sqrt{7.5} = 2.74$$

More Exercise

Exercise 1:

Suppose that in a certain city, the weekly number of infected cases with Corona virus (COVID-19) has a Poisson distribution with an average (mean) of 5 cases per week.

(A) Find:

1. The probability distribution of the weekly number of infected cases (X).

$$P(X=x)=\begin{cases}\dfrac{e^{-\lambda}\,\lambda^{x}}{x!} & ;\quad x=0,1,2,3,\ldots\\ 0 & ; Otherwise\end{cases}$$

$$\lambda=5$$

$$P(X=x)=\begin{cases}\dfrac{e^{-5}\,5^{x}}{x!} & ;\quad x=0,1,2,3,\ldots\\ 0 & ; Otherwise\end{cases}$$

2. The probability that there will be 2 infected cases this week.

$$P(X=2)=\frac{e^{-5}\,5^{2}}{2!}=0.0842$$

3. The probability that there will be 1 infected case this week.

$$P(X=1)=\frac{e^{-5}\,5^{1}}{1!}=0.0337$$

4. The probability that there will be no infected cases this week.

$$P(X=0)=\frac{e^{-5}\,5^{0}}{0!}=0.0067$$

5. The probability that there will be at least 3 infected cases this week.

$$P(X\geq 3)=1-P(X<3)=1-P(X\leq 2)$$
$$=1-[P(X=0)+P(X=1)+P(X=2)]$$
$$=1-[\,0.0067+0.0337+0.0842\,]$$
$$=1-0.1246=0.8754$$

6. The probability that there will be at most 2 infected cases this week.

$$P(X\leq 2)=P(X=0)+P(X=1)+P(X=2)$$
$$=0.0067+0.0337+0.0842$$
$$=0.1246$$

7. The expected number (mean/average) of infected cases this week.

$$E(X) = \mu = \mu_X = \lambda = 5$$

8. The variance of the number of infected cases this week.

$$Var(X) = \sigma^2 = \sigma_X^2 = \lambda = 5$$

(B): Find:

1. The average (mean) of the number infected cases in a day.

$$\lambda = \frac{5}{7} = 0.7143$$

2. The probability distribution of the daily number of infected cases (X).

$$P(X = x) = \begin{cases} \dfrac{e^{-\lambda}\ \lambda^x}{x!} & ; \quad x = 0, 1, 2, 3, \dots \\ 0 & ; Otherwise \end{cases}$$

$$\lambda = \frac{5}{7}$$

$$P(X = x) = \begin{cases} \dfrac{e^{-\frac{5}{7}}\ \left(\frac{5}{7}\right)^x}{x!} & ; \quad x = 0, 1, 2, 3, \dots \\ 0 & ; Otherwise \end{cases}$$

3. The probability that there will be 2 infected cases tomorrow.

$$P(X = 2) = \frac{e^{-\frac{5}{7}}\ \left(\frac{5}{7}\right)^2}{2!} = 0.1249$$

4. The probability that there will be 1 infected case tomorrow.

$$P(X = 1) = \frac{e^{-\frac{5}{7}}\ \left(\frac{5}{7}\right)^1}{1!} = 0.3497$$

5. The probability that there will be no infected cases tomorrow.

$$P(X = 0) = \frac{e^{-\frac{5}{7}}\ \left(\frac{5}{7}\right)^0}{0!} = 0.4895$$

6. The probability that there will be at most 2 infected cases tomorrow.

$$P(X \le 2) = P(X = 0) + P(X = 1) + P(X = 2)$$
$$= 0.4895 + 0.3497 + 0.1249$$
$$= 0.9641$$

7. The probability that there will be at least 2 infected cases tomorrow.

$$P(X \ge 2) = 1 - P(X < 2) = 1 - P(X \le 1)$$
$$= 1 - [P(X = 0) + P(X = 1)]$$
$$= 1 - [0.4895 + 0.3497]$$
$$= 1 - 0.8392 = 0.1608$$

8. The expected number (mean/average) of infected cases tomorrow.

$$E(X) = \mu = \mu_X = \lambda = \frac{5}{7} = 0.7143$$

9. The variance of the number of infected cases tomorrow.

$$Var(X) = \sigma^2 = \sigma_X^2 = \lambda = \frac{5}{7} = 0.7143$$

(C): Assuming that 4 weeks are in a month, find:

1. The average (mean) of the number infected cases per month.

$$E(X) = \mu = \mu_X = \lambda = 5 \times 4 = 20$$

2. The variance of the number of infected cases per month.

$$Var(X) = \sigma^2 = \sigma_X^2 = \lambda = 5 \times 4 = 20$$

Chapter 7:
Continuous Probability Distribution

The Uniform Distribution:

$$f(x) = \frac{1}{b-a} \quad ; \quad a < x < b$$

$$\mu = \frac{a+b}{2} \qquad \sigma^2 = \frac{(b-a)^2}{12}$$

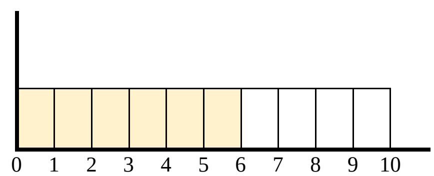
Question 1:

If the random variable X has a uniform distribution on the interval (0,10), then

(1) $P(X < 6)$ equal to:

$$f(x) = \frac{1}{10} \quad , \; 0 < X < 10 \qquad \boxed{a = 0 \quad , \quad b = 10}$$

$$P(X < 6) = \frac{6}{10} = 0.6$$



(2) The mean of X is

$$\mu = \frac{a+b}{2} = \frac{10}{2} = 5$$

(3) The variance of X is

$$\sigma^2 = \frac{(b-a)^2}{12} = \frac{100}{12} = 8.33$$

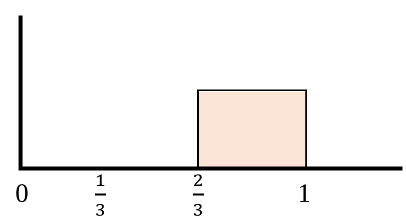
Question 2:

Suppose that the random variable X has the following uniform distribution:

$$f(x) = 3 \quad ; \quad \frac{2}{3} < x < 1$$

1. $P(0.33 < X < 0.5) = 0$
2. $P(X > 1.25) = 0$
3. The variance of X is

$$\sigma^2 = \frac{(b-a)^2}{12} = \frac{\left(1 - {}^2\!/_3\right)^2}{12} = 0.00926$$



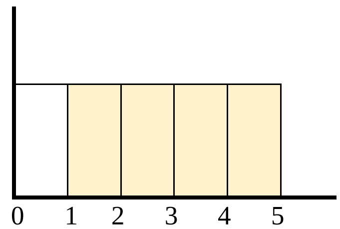
Question 3:

Suppose that the continuous random variable X has the following probability density function (pdf): $f(x) = 0.2 \; ; \; 0 < X < 5$. Then

$$a = 0 \quad , \quad b = 5$$

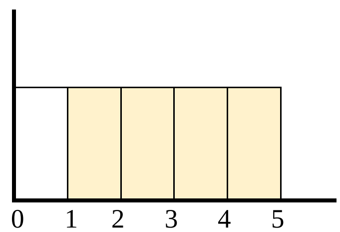
(1): $P(X > 1)$ is equal to:

$$P(X > 1) = \frac{4}{5} = 0.8$$



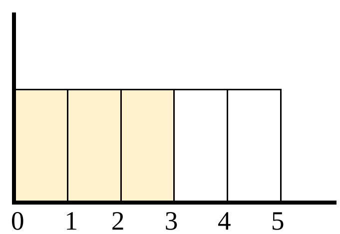
(2): $P(X \geq 1)$ is equal to:

$$P(X \geq 1) = \frac{4}{5} = 0.8$$



(3): $P(X < 3)$ is equal to:

$$P(X < 3) = \frac{3}{5} = 0.6$$



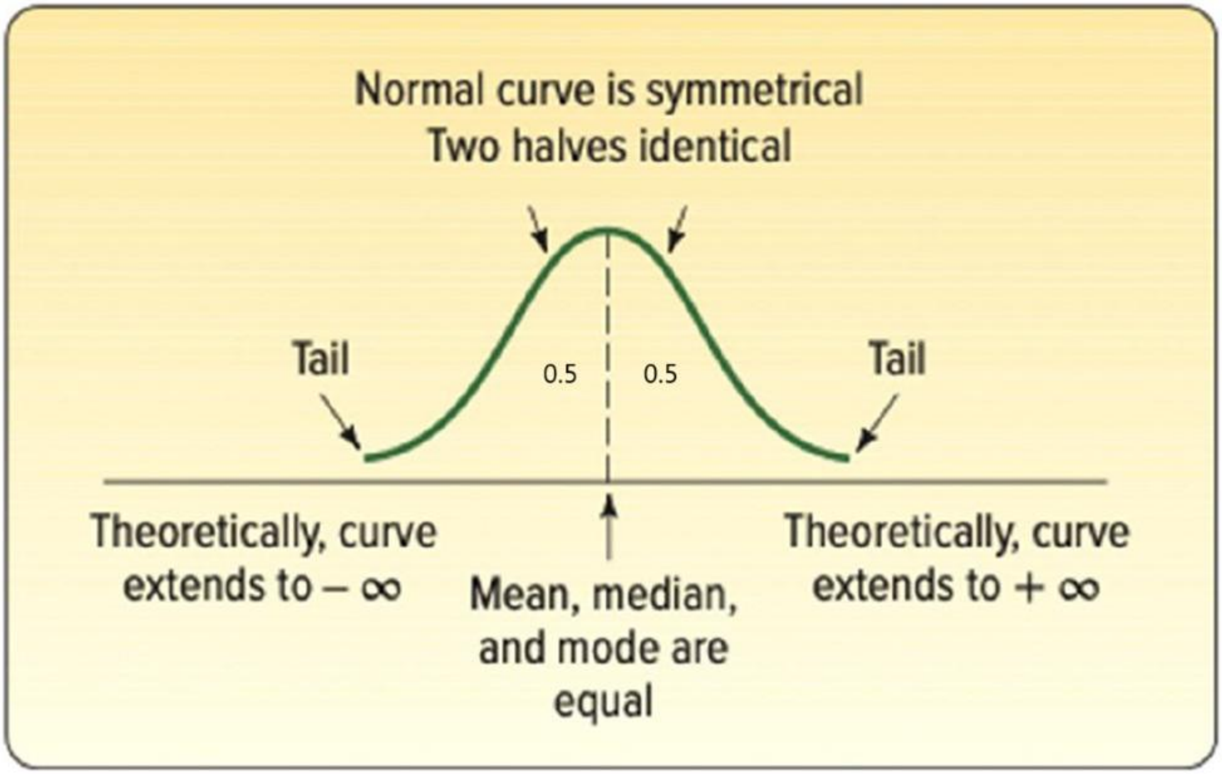
(4): The mean of X is equal to:

$$\mu = \frac{a+b}{2} = \frac{5}{2} = 2.5$$

(5): The variance of X is equal to:

$$\sigma^2 = \frac{(b-a)^2}{12} = \frac{25}{12} = 2.0833$$

Normal Distribution



Question 1:

A normal population has a mean (μ) of 21 and a standard deviation (σ) of 5.

1. Compute the z-value associated with 25.

   $z = \frac{x-\mu}{\sigma} = \frac{25-21}{5} = 0.80$

   The z-value is 0.80

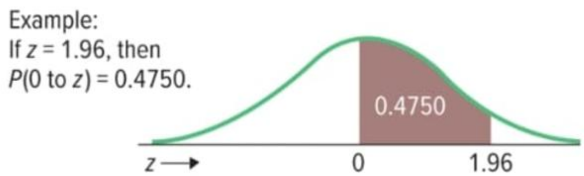
2. What proportion of the population is between 21 and 25?

   The z-value for 25 is $z = \frac{x-\mu}{\sigma} = \frac{25-21}{5} = 0.80$

   The z-value for 21 is $z = \frac{x-\mu}{\sigma} = \frac{21-21}{5} = 0$

   $\mathrm{P}(0 \text{ to } 0.80) = 0.2881$

Using a standard normal distribution table,

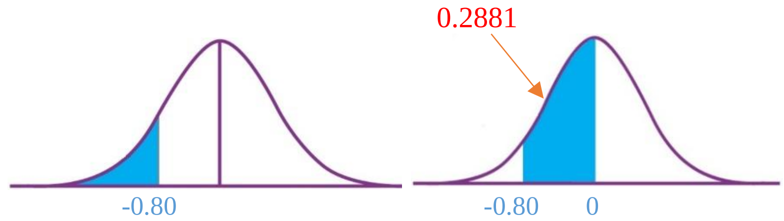


z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621

3. What proportion of the population is less than 17?

The z-value for 17 is $z = \frac{x-\mu}{\sigma} = \frac{17-21}{5} = -0.80$

$P(\text{less than}[-0.80]))$
$= 0.5 - 0.2881 = 0.2119$



Using a standard normal distribution table,

z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621

The proportion of the population less than 17 is 0.2119.

Question 2:

A normal population has a mean (μ) of 20.0 and a standard deviation (σ) of 4.0.

1. Compute the z-value associated with 25.0.

$$z = \frac{x-\mu}{\sigma} = \frac{25-20}{4} = 1.25$$

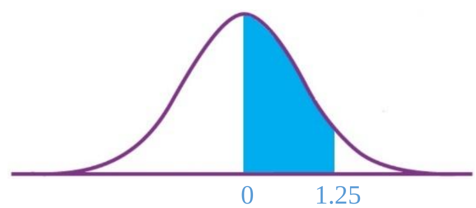
The z-value is 1.25

2. What proportion of the population is between 20.0 and 25.0?

The z-value for 25 is $z = \frac{x-\mu}{\sigma} = \frac{25-20}{4} = 1.25$

The z-value for 20 is $z = \frac{x-\mu}{\sigma} = \frac{20-20}{4} = 0$

$P(0 \text{ to } 1.25) = 0.3944$



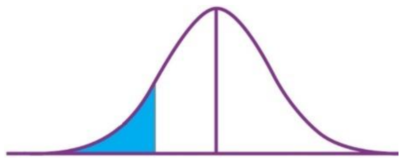
Using a standard normal distribution table,

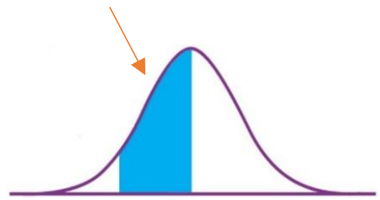
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177

3. What proportion of the population is less than 18.0?

The z-value for 18 is $z = \frac{x-\mu}{\sigma} = \frac{18-20}{4} = -0.50$

$P(\text{less than}[-0.50]))$
$= 0.5 - 0.1915 = 0.3085$





Using a standard normal distribution table,

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852

The proportion of the population less than 18.0 is approximately 0.3085

Question 3:

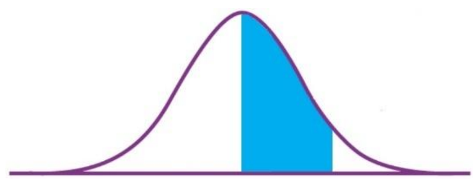
The mean hourly pay of an American Airlines flight attendant is normally distributed with a mean (μ) of \$29.81 per hour and a standard deviation (σ) of \$9.31 per hour.

1. What is the probability that the hourly pay is between the mean and \$35.00 per hour?

The z-value for 35 is $z = \frac{x-\mu}{\sigma} = \frac{35-29.81}{9.31} = 0.56$

The z-value for 29.81 is $z = \frac{x-\mu}{\sigma} = \frac{29.81-29.81}{9.31} = 0$

P(0 to 0.56) = 0.2123



Using the z-table,

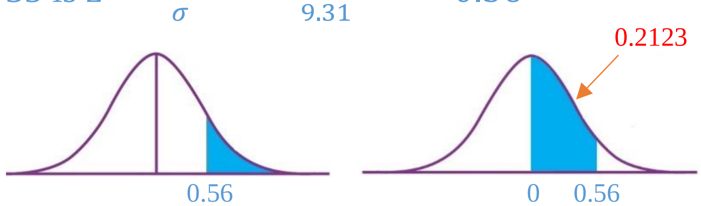
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852

2. What is the probability that the hourly pay is more than \$35.00 per hour?

from \$0 to \$35

The z-value for 35 is $z = \frac{x-\mu}{\sigma} = \frac{35-29.81}{9.31} = 0.56$

$P(\text{more than}[0.56])$
$= 0.5 - 0.2123 = 0.2877$



Using the z-table,

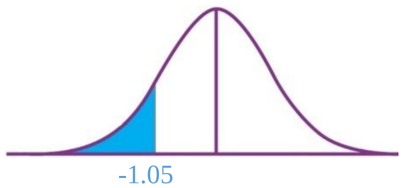
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852

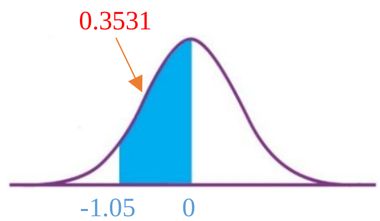
The probability that the hourly pay is more than \$35.00 is 0.2877.

3. What is the probability that the hourly pay is less than $20.00 per hour?

The z-value for 20 is $z = \frac{x-\mu}{\sigma} = \frac{20-29.81}{9.31} = -1.05$

$P(\text{less than}[-1.05]))$
$= 0.5 - 0.3531 = 0.1469$





Using the z-table,

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015

The probability that the hourly pay is less than $20.00 is 0.1469.

Question 4:

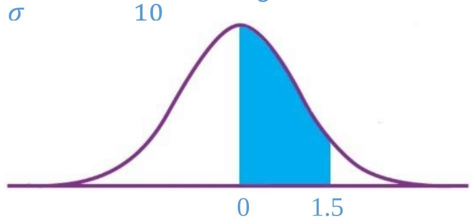
The mean of a normal probability distribution is 400 pounds, and the standard deviation is 10 pounds.

1. What is the area between 415 pounds and the mean of 400 pounds?

The z-value for 415 is $z = \frac{x-\mu}{\sigma} = \frac{415-400}{10} = 1.5$

The z-value for 400 is $z = \frac{x-\mu}{\sigma} = \frac{400-400}{10} = 0$

$P(0 \text{ to } 1.5) = 0.4332$



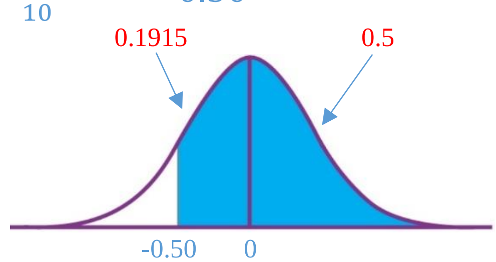
Using a standard normal distribution table,

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545

The proportion between 400 and 415 is 0.4332

2. What is the probability of selecting a value at random and discovering that it has a value of more than 395 pounds?

The z-value for 395 is $z = \frac{x-\mu}{\sigma} = \frac{395-400}{10} = -0.50$



$$P(\text{more than}[-0.50]) = 0.1915 + 0.5 = 0.6915$$

Using a standard normal distribution table,

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852

The probability that the value is more than 395 pounds is 0.6915

3. What is the probability of selecting a value at random and discovering that it has a value of less than 395 pounds?

If the probability that the value is more than 395 pounds is 0.6915

Then,
The probability that the value is less than 395 pounds is

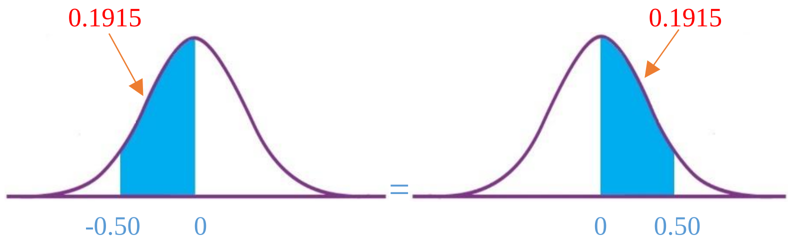
$$1 - 0.6915 = 0.3085$$

4. What is the area between the mean and 395 pounds?

The z-value for 395 is $z = \frac{x-\mu}{\sigma} = \frac{395-400}{10} = -0.50$

The z-value for 400 is $z = \frac{x-\mu}{\sigma} = \frac{400-400}{10} = 0$

$P(-0.50 \text{ to } 0) = 0.1915$

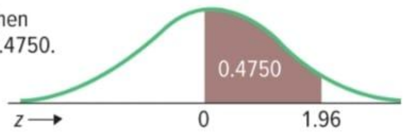


Using a standard normal distribution table,

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852

The probability between the mean and 395 is 0.1915

Example:
If $z = 1.96$, then
$P(0 \text{ to } z) = 0.4750$.



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

The Exponential Distribution:

- Describes times between events in a sequence.
- Events are independent at a constant rate.
- They are non-negative, always positive.
- Positively skewed.

Examples of situations using the exponential distributions.

- The service time for customers at the information desk at Dallas Public Library.
- The time until the next phone call arrives in a customer service center.

- Distribution is described by a single parameter.
- λ (lambda): the rate parameter.
- Decrease λ shape is "less skewed".

$$P(x) = \lambda e^{-\lambda x} \quad ; \quad x > 0$$

Mean: $\mu = \frac{1}{\lambda}$

Variance: $\sigma^2 = \frac{1}{\lambda^2}$

Standard deviation: $\sigma = \frac{1}{\lambda}$

$P(\text{Arrival time} < x) = 1 - e^{-\lambda x}$

Question 1:

Let's consider a bank where customers arrive at the service desk to be served. Suppose the bank has observed that, on average, a customer arrives every 5 minutes. The time between consecutive customer arrivals follow the exponential distribution.

1. What is the rate parameter λ ?

$$\lambda = \frac{1}{\mu} = \frac{1}{5} = 0.2$$

2. What is the variance σ^2 ?

$$\sigma^2 = \frac{1}{\lambda^2} = \frac{1}{0.2^2} = 25$$

3. Find the probability that the next customer will arrive within 3 minutes:

$$P(X \leq 3) = 1 - e^{-\lambda x}$$
$$= 1 - e^{-0.2(3)} = 0.45$$

4. Find the probability that the next customer will arrive in less than 5 minutes:

$$P(X < 5) = 1 - e^{-\lambda x}$$
$$= 1 - e^{-0.2(5)} = 0.63$$

5. Find the probability that the next customer will arrive after more than 2 minutes:

$$P(X > 2) = 1 - P(X \leq 2)$$
$$= 1 - \left(1 - e^{-\lambda x}\right)$$
$$= 1 - 1 + e^{-0.2(2)}$$
$$= e^{-0.2(2)}$$
$$= 0.67$$

6. Find the probability that the next customer will arrive between 2 to 3 minutes

$$P(2 < X < 3)$$
$$= P(X < 3) - P(X < 2)$$
$$= 1 - e^{-0.2(3)} - \left(1 - e^{-0.2(2)}\right)$$
$$= 1 - e^{-0.6} - 1 + e^{-0.4}$$
$$= -e^{-0.6} + e^{-0.4}$$
$$= 0.1215$$

Question 2:

Imagine a machine in a factory that breaks down and requires repair. The time between machine breakdowns is known to follow an exponential distribution with a rate parameter $\lambda = 0.3$ breakdowns per hour.

1. What is the probability that the time until the next breakdown is more than 10 hours:

$$\begin{aligned} P(X > 10) &= 1 - P(X \leq 10) \\ &= 1 - \left(1 - e^{-\lambda x}\right) \\ &= 1 - 1 + e^{-0.3(10)} \\ &= e^{-3} \\ &= 0.0498 \end{aligned}$$

2. Find the mean time between breakdowns:

$$\text{Mean: } \mu = \frac{1}{\lambda} = \frac{1}{0.3} = 3.33$$

3. Find the variance of the time between breakdowns:

$$\text{Variance: } \sigma^2 = \frac{1}{\lambda^2} = \frac{1}{0.3^2} = 11.11$$

Question 3:

A website experiences user logins, and the time between consecutive logins follows an exponential distribution. The average time between logins is 10 mins.

1. Find the rate parameter:

$$\lambda = \frac{1}{\mu} = \frac{1}{10} = 0.1$$

2. Find the probability that the next user login occurs between 5 and 15 mins:

$$\begin{aligned} &P(5 < X < 15) \\ &= P(X < 15) - P(X < 5) \\ &= 1 - e^{-0.1(15)} - \left(1 - e^{-0.1(5)}\right) \\ &= 1 - e^{-0.1(15)} - 1 + e^{-0.1(5)} \\ &= 1 - e^{-1.5} - 1 + e^{-0.5} \\ &= \quad -e^{-1.5} + e^{-0.5} \\ &= 0.3834 \end{aligned}$$