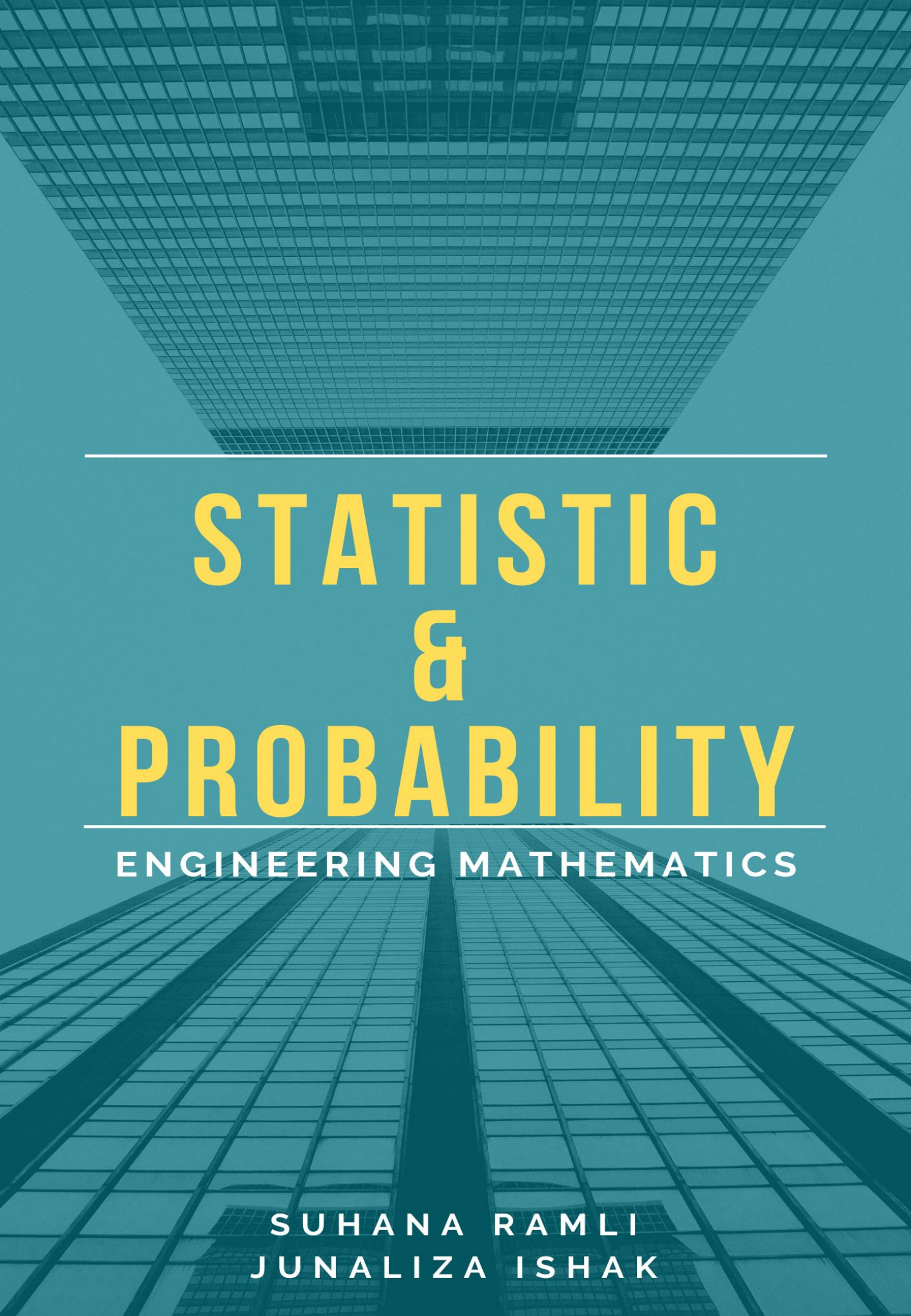


STATISTIC & PROBABILITY

ENGINEERING MATHEMATICS

SUHANA RAMLI
JUNALIZA ISHAK

STATISTIC & PROBABILITY

ENGINEERING MATHEMATICS

PUBLISHED BY :
MATHEMATICS, SCIENCE & COMPUTER DEPARTMENT
POLITEKNIK PORT DICKSON
http : www.polipd.edu.my

Perpustakaan Negara Malaysia
Cataloguing-in-Publication Data

Suhana Ramli, 1979-
STATISTIC & PROBABILITY : ENGINEERING MATHEMATICS / SUHANA RAMLI, JUNALIZA ISHAK
Mode of access: Internet
eISBN 978-967-2897-48-4
1. Statistic
2. Probability
3. Engineering mathematics.
4. Government publications--Malaysia.
5. Electronic books.
I. Junaliza Ishak, 1975-. II. Title.
519.5

Published by :
Politeknik Port Dickson
KM 14 Jalan Pantai, 71050 Si Rusa
Negeri Sembilan

JUN 2022

PREFACE

Grateful to Allah because with His permission, the eBook Statistic & Probability (Engineering Mathematics) was published. This eBook is written by lecturers who have been teaching in Engineering Mathematics for more than 15 years . This eBook can be used by all institutions of higher learning such as Polytechnics and Colleges as well as private and public universities. The purpose for this eBook was written is to make it easier for students to gain knowledge and review the topic of calculus in a simpler and more concise way.

Many examples in various forms of questions are included in this eBook with detailed steps of solution to make it easier for students to quickly understand the method of its solution. In addition, students will also able to improve and strengthen their understanding through the included practice questions. The authors hope that this eBook can benefit all students as well as educators around the world in the field of Engineering Mathematics.

Thank You So Much.

Suhana binti Ramli

Editor

CONTENTS

1.0 Statistical Data

1.1 Define Statistical Terminology

Statistics definition is the study and manipulation of data, including ways to gather, review, analyze, and draw conclusions from data. Its dealing with the collection, analysis, interpretation, and presentation of masses of numerical data.

What is grouped and ungrouped data?

Ungrouped Data

Ungrouped data is the data that first gather from an experiment or study. The data is raw, it's not sorted into categories, classified, or otherwise grouped

Example 1 :

4, 3, 10,7, 8, 7. 3, 20, 15, 4, 4

Example 2 :

Score, x	10	15	20	25
Number of Student	2	7	6	3

Grouped Data

Grouped data is **data** that has been bundled together in categories.

Example :

Score, x	5-8	9-12	13-16	17-20
Number of Student	2	7	6	3

Step to construct a frequency table ungrouped data to grouped data

Step	Task
1	State the highest value and the lowest value
2	Find the number of classes (If not given) $Number\ of\ class, K = 1 + 3.3\ log\ n$ Where n is number of data.
3	Find, $\text{size of class} = \frac{\text{highest value} - \text{lowest value}}{number\ of\ classes}$
4	State the class limit for each class.
5	Determine the frequency of each class

** In some case number
of class can more 1 than '*K*' or less 1 than '*K*'

Example 1

Given below are marks obtained by 20 students in Math out of 25.

21, 23, 19, 17, 12, 15, 15, 17, 17, 19, 23, 23, 21, 23, 25, 25, 21, 19, 19, 19

Construct

a) ungrouped frequency table

b) grouped frequency table

Solution (a)

Marks Obtained	Tally Marks	Frequency
12	I	1
15	II	2
17	III	3
19	IIII	5
21	III	3
23	IIII	4
25	II	2

Solution (b)

21, 23, 19, 17, 12, 15, 15, 17, 17, 19, 23, 23, 21, 23, 25, 25, 21, 19, 19, 19

Step	Task
1	State the highest value and the lowest value The highest value = 25 The lowest value = 12
2	$Number\ of\ class = 1 + 3.3\ log\ n$ = 1 + 3.3 log 20 = 5.29 ≈ 6
3	Size of class $= \frac{\text{highest value} - \text{lowest value}}{number\ of\ classes}$ $= \frac{25 - 12}{6}$ $= \frac{13}{6}$ $= 2.17$ ≈ 3
4	**Construct the table** Size of class

Marks	Frequency
12 – 14	1
15 – 17	5
18 – 20	5
21 – 23	7
24 – 26	2

****Remember !**

In some case number of class can more 1 or less 1 than '*K*'

Example 2

Data shows the weight of 25packets of sugar in kilograms. Construct frequency table with 6 number of class.

1.9	2.9	2.8	1.1	0.5
2.2	2.3	2.3	2.2	1.0
1.6	1.9	0.8	3.2	0.5
2.0	2.7	2.6	2.8	0.8
3.1	2.6	2.9	0.6	3.0

Solution

Step	Task
1	State the highest value and the lowest value Highest value = 3.2 Lowest value = 0.5
2	**Number of classes = 6 (given from question)** $\text{size of class} = \frac{\text{highest value} - \text{lowest value}}{number\ of\ classes}$ $= \frac{3.2-0.5}{6}$ $= 0.45$ $\cong 0.5$

Step	Task
3.	**Construct frequency table**

Weight	Tally	f
0.5 – 0.9	𝍸	5
1.0 – 1.4	//	2
1.5 – 1.9	///	3
2.0 – 2.4	𝍸	5
2.5 – 2.9	𝍸 //	7
3.0 – 3.4	//	3

Example 3

30 AA batteries were tested to determine how long they would last. The results, to the nearest minute were recorded as follows:

423	411	371	409	431	391	400	415	396	419
369	393	377	392	401	405	381	428	372	386
387	394	389	408	363	382	399	422	410	390

Construct **a frequency** table using 363-372 as the first class interval.

Solution

Step	Task
1	**Given 363-372 as the first class interval** **Number of data = 30**
2	**Find the number of classes (not given)** **Number of class, K = 1 + 3.3 log n** = 1 + 3.3 log 30 = 5.87 ≅ 6
3	**Construct frequency table**

Minute	Tally	*f*
363 – 372	////	4
373 – 382	///	3
383 – 392	~~////~~ /	6
393 – 402	~~////~~ //	7
403 – 412	////	4
413 – 422	///	3
423 - 432	///	3

Example 4

Data shows the length of plant stems grown in cm after their seeds were moved to a new location. Prepare a frequency table for this data.

9.55	7.85	8.45	9.65	8.55
8.65	9.55	10.05	8.95	7.25
9.05	7.55	8.25	10.75	9.05
8.25	10.65	7.25	7.85	7.55
7.75	9.45	8.55	9.35	9.95

Solution

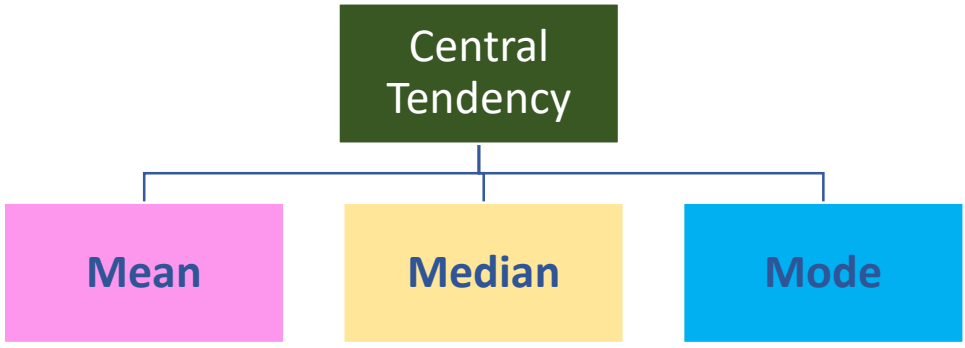
Step	Task
1	State the highest value and the lowest value Highest value = 10.65 Lowest value = 7.25
2	Find the number of classes (not given) **Number of class, K = 1 + 3.3 log n** = 1 + 3.3 log 25 = 5.61 ≅ 6
3	Find, **size of class** $= \frac{\text{highest value} - \text{lowest value}}{number\ of\ classes}$ $= \frac{10.65-7.25}{6}$ = 0.57 ≅ 0.6

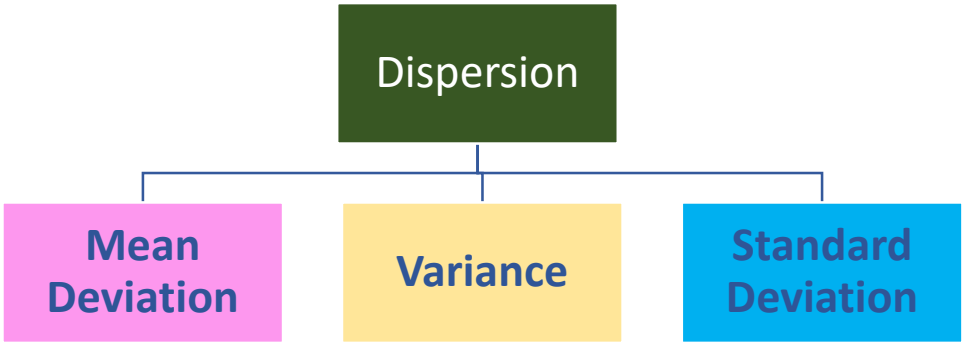
Step	Task
3.	**Construct frequency table**

cm	Tally	*f*
7.25 - 7.84	𝍸	5
7.85 - 8.44	////	4
8.45 – 9.04	𝍸 /	6
9.05 – 9.64	𝍸	5
9.65 – 10.24	///	3
10.25 – 10.84	//	2

1.2 Central Tendency & Dispersion

Central tendency is described by median, mode, and the means (there are different means- geometric and arithmetic). **Dispersion** is the degree to which data is distributed around this central tendency, and is represented by range, deviation, variance and standard deviation.





1.2.1 Definition Mean, Mode And Median

Mean is the average of a set of data

Mode is the number that occurs the most

Median is the middle number in a set of data when the data is arranged in numerical order.

1.2.2 Mean, Median And Mode (ungrouped data)

MEAN	$\bar{x} = \frac{\sum x}{n}$ Where, $\sum x$ = sum of data n is number of data	**With frequency** $\bar{x} = \frac{\sum (fx)}{\sum f}$ Where, $\sum f$= sum of frequency
MEDIAN	Data arranged in ascending or descending **Even data :** Average value of a pair middle data $\text{median} = \frac{\left(\frac{n}{2}\right)^{th} term + \left(\frac{n}{2}+1\right)^{th} term}{2}$ **Odd data :** Value of data that in the middle $\text{median} = \left(\frac{n+1}{2}\right)^{th}$ term **Data with frequency : m =** $\frac{1}{2}\sum f$	
MODE	The mode is value or values that occur the most	

Example 1 (even number)

Find the mean, median and mode of the sequence 19, 12, 9, 5, 8, 5, 25, 8

Solution

Total number of numbers in the sequence(n) = 8

i) Mean $(\vec{x}) = \frac{\sum x}{n}$

$= \frac{19+12+9+5+8+5+25+8}{8}$

$= 11.37$

ii) Arrange the sequence in the ascending order:

5, 5, 8, 8, 9, 12, 19, 25

$$\text{median} = \frac{\left(\frac{n}{2}\right)^{th} term + \left(\frac{n}{2}+1\right)^{th} term}{2}$$

$$= \frac{\left(\frac{8}{2}\right)^{th} term + \left(\frac{8}{2}+1\right)^{th} term}{2}$$

$$= \frac{4^{th}\ term + 5^{th}\ term}{2}$$

$$= \frac{8+9}{2}$$

= 8.5

iii) Mode = 5 and 8

Example 2 (odd number)

Find the mean, median and mode of the sequence 12, 4, 7, 5, 5,9,15

Solution

Total number of numbers in the sequence(n) = 7

i) Mean $(\vec{x}) = \frac{\sum x}{n}$

$= \frac{12+4+7+5+5+9+15}{7}$

$= 8.143$

ii) Arrange the sequence in the ascending order:

4, 5, 5, 7, 9, 12, 15

Median, m $= \left(\frac{n+1}{2}\right)^{th}$ term

$= \left(\frac{7+1}{2}\right)^{th}$ term

= 4^{th} term

= 7

iii) Mode = 5

Example 2 (*data in table*)

Find the mean, median and mode of the marks below.

Score	10	15	20	25	30
Frequency	2	8	5	5	4

Solution

Calculate $f \bullet x$ and cumulative frequency F

Score(x)	Frequency (f)	$f(x)$	F
10	2	20	2
15	8	120	10
20	5	100	15
25	5	125	20
30	4	120	24
TOTAL	24	485	

12^{th}

i) Mean, $\overline{x} = \frac{\sum f(x)}{\sum f}$

$= \frac{485}{24}$

$= 20.208$

iii) Mode = 15

(refer to highest frequency)

ii) Median, $m = \frac{1}{2}\sum f$

$= \frac{1}{2}(24)$

$= 12^{th}$

Take the 12th data (refer to *F*) = 5

1.2.3 Mean, Median And Mode (grouped data)

MEAN	$\bar{x} = \frac{\sum(fx)}{\sum f}$	➢ ∑fx= sum of frequency ➢ x = midpoint ➢ $\sum(fx)$ = sum each fx
MEDIAN	$m = L + \left(\frac{\frac{\sum f}{2} - F_b}{f_m}\right)c$	➢ L : lower boundary of class median ➢ $\sum f$: sum of frequency ➢ F_b : cumulative frequency before class median ➢ f_m : frequency of class median ➢ c-size of class median
MODE	$m_0 = L + \left(\frac{d_1}{d_1 + d_2}\right)c$	➢ L : lower boundary of class mode ➢ d_1 = frequency class mode - frequency BEFORE class mode ➢ d_2 = frequency class mode - frequency AFTER class mode ➢ C : size of class mode

Median of set data also can be determined graphically by using an ogive.

Mode of set data can be determined graphically by using histogram.

Example 1

The following data shows the bonus point earned by customers at Pasarayaku

Bonus	Number of Customers
1 - 5	3
6 - 10	15
11 – 15	8
16 – 20	9
21 – 25	20
26 – 30	18

a) Find mean, median and mode by using formula

b) Find median and mode by using graph

Solution

a) Find mean, median and mode by using formula

Bonus	Midpoint (x)	f	$f(x)$	F	Boundaries
1 - 5	3	3	9	3	0.5 – 5.5
6 – 10	8	15	120	18	5.5 – 10.5
11 – 15	13	8	104	26	10.5 – 15.5
16 – 20	18	9	162	35	15.5 – 20.5
21 – 25	23	20	460	55	20.5 – 25.5
26 – 30	28	18	504	73	25.5 – 30.5
Total		**73**	**1359**		

i) Mean, $(\vec{x}) = \frac{\sum f(x)}{\sum f}$

$= \frac{1359}{73}$

$= 18.62$

ii) Median, $m = \frac{1}{2}(\sum f)$

$= \frac{1}{2}(73)$

$= 36.5^{th}$

take the 36.5^{th} data (refer to *F*) = 21 - 25

Median, $m = L + \left(\frac{\frac{\sum f}{2} - F_b}{f_m}\right)c$

- L : lower boundary of class median
- $\sum f$: sum of frequency
- F_b : cumulative frequency before class median
- f_m : frequency of class median
- c-size of class median

L = 20.5 F_b = 35 f_m = 20 C = 5

$m = 20.5 + \left(\frac{\frac{73}{2} - 35}{20}\right)5$

= 20.88

iii) The group with the highest frequency, which is **21 - 25**

$$\text{Mode, } m_0 = L + \left(\frac{d_1}{d_1+d_2}\right)c$$

- L : lower boundary of class mode
- d_1 = frequency class mode - frequency BEFORE class mode
- d_2 = frequency class mode - frequency AFTER class mode
- C : size of class mode

L = 20.5 d_1 = 20 - 9 = 11 d_2 = 20 - 18 = 2 C = 5

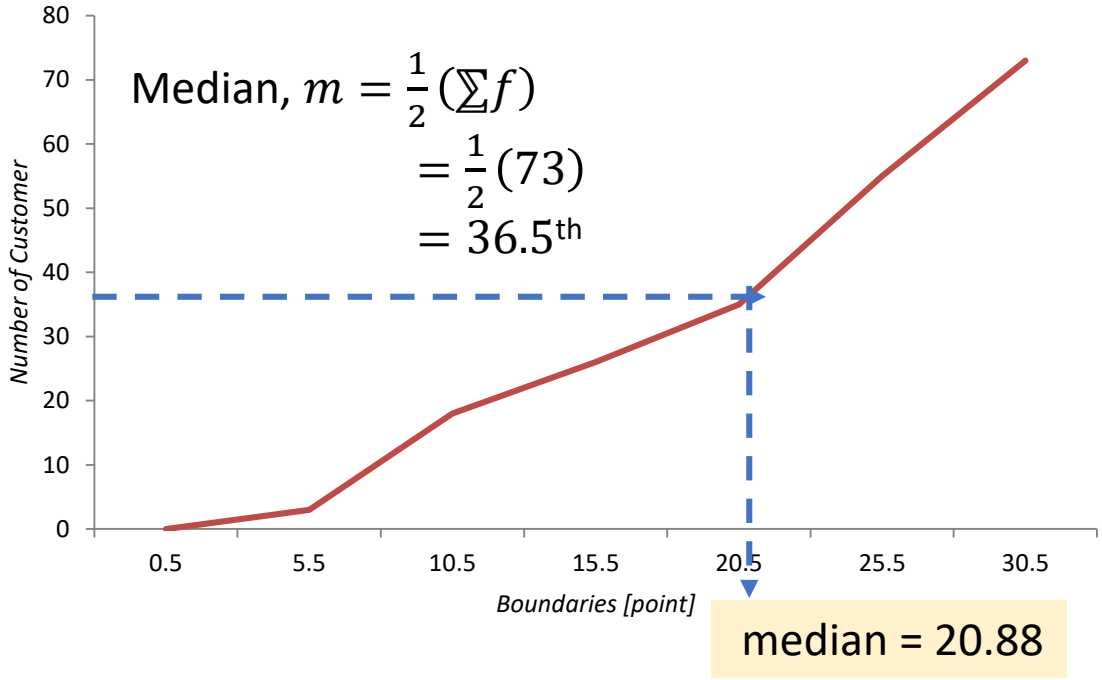
$$\text{Mode, } m_0 = 20.5 + \left(\frac{11}{11+2}\right)5$$

$$= 24.73$$

b) Find median and mode by using graph

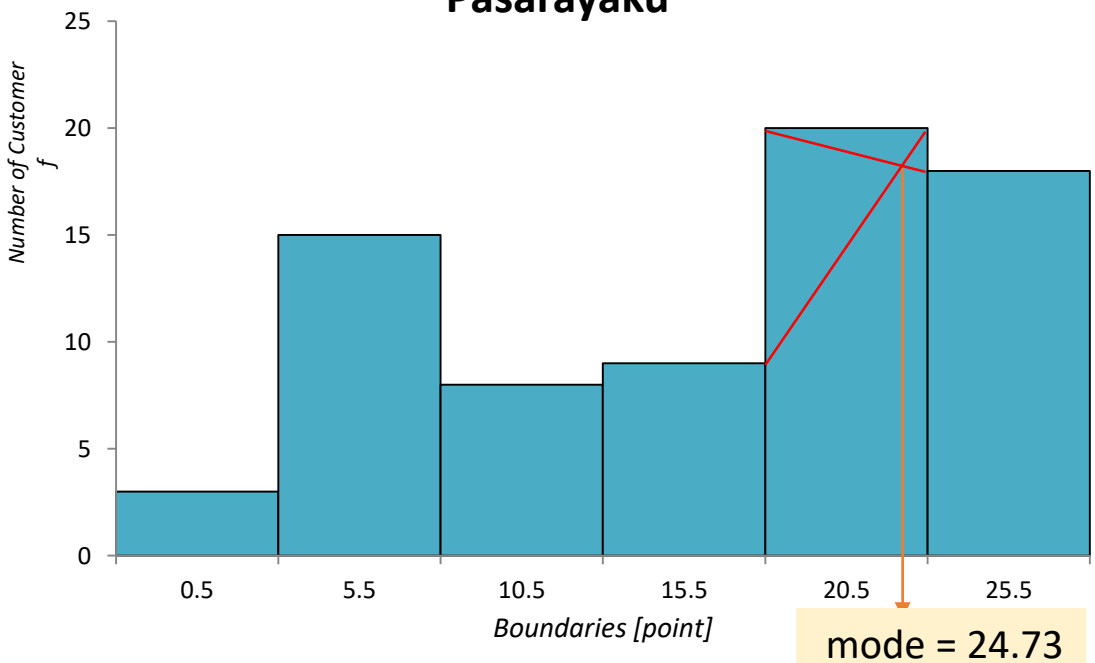
Boundaries	*F*
0.5 – 5.5	3
5.5 – 10.5	18
10.5 – 15.5	26
15.5 – 20.5	35
20.5 – 25.5	55
25.5 – 30.5	73

Bonus Point Earned By Customers At Pasarayaku



median = 20.88

Bonus Point Earned By Customers At Pasarayaku



mode = 24.73

1.2.4 Mean Deviation, Variance And Standard Deviation

Mean deviation is the mean of all the absolute deviations of a set of data.

Variance is the mean squared deviation whereas mean is the average of all values in a given data set. In statistics, the variance is used to determine how well the mean represents an entire set of data.

Standard deviation is the spread of a group of numbers from the mean.

DISPERSION	UNGROUPED DATA	GROUPED DATA
MEAN DEVIATION	$E = \frac{\sum \lvert x - \overline{x} \rvert}{n}$	$E = \frac{\sum f \lvert x - \overline{x} \rvert}{\sum f}$
VARIANCE	$S^2 = \frac{\sum (x - \overline{x})^2}{n}$	$S^2 = \frac{\sum f (x - \overline{x})^2}{\sum f}$
STANDARD DEVIATION	$S = \sqrt{Variance}$	

Example 1

Calculate the mean deviation, variance and standard deviation for the following distribution:

7, 11, 11, 15, 20, 20, 28

Solution

STEP 1

Mean:

$$\bar{x} = \frac{\sum x}{n}$$

$$= \frac{112}{7}$$
$$= 16$$

STEP 2

x	$\lvert x - \bar{x} \rvert$	$(x - \bar{x})^2$
7	9	81
11	5	25
11	5	25
15	1	1
20	4	16
20	4	16
28	12	144
$\sum 112$	$\sum 40$	$\sum 308$

STEP 3

i) Mean deviation

$$E = \frac{\sum |x - \bar{x}|}{n}$$

$$= \frac{40}{7}$$
$$= 5.71$$

ii) Variance

$$S^2 = \frac{\sum (x - \bar{x})^2}{n}$$

$$= \frac{308}{7}$$
$$= 44$$

ii) Standard Deviation

$$S = \sqrt{\text{Variance}}$$
$$= \sqrt{44}$$
$$= 6.63$$

Example 2

Calculate the mean deviation, variance and standard deviation for the following distribution:

x	2	4	6	8	10
f	4	5	3	6	2

Solution

STEP 1

Mean:

$$\overline{x} = \frac{\sum f(x)}{\sum f}$$

$$= \frac{114}{20}$$

$$= 5.7$$

STEP 2

x	f	$f(x)$	$\lvert x - \overline{x} \rvert$	$f\lvert x - \overline{x} \rvert$	$(x - \overline{x})^2$	$(x - \overline{x})^2$
2	4	8	3.7	14.8	13.69	54.76
4	5	20	1.7	8.5	2.89	14.46
6	3	18	0.3	0.9	0.09	0.27
8	6	48	2.3	13.8	5.29	31.74
10	2	20	4.3	8.6	18.49	36.98
	$\sum 20$	$\sum 114$		$\sum 46.6$		$\sum 138.2$

STEP 3

i) Mean deviation

$$E = \frac{\sum f|x - \bar{x}|}{\sum f}$$

$$= \frac{46.6}{20}$$

$$= 2.33$$

ii) Variance

$$S^2 = \frac{\sum f(x-\bar{x})^2}{\sum f}$$

$$= \frac{138.2}{20}$$

$$= 6.91$$

iii) Standard Deviation

$$S = \sqrt{\text{Variance}}$$

$$= \sqrt{6.91}$$

$$= 2.62$$

Example 3

Calculate the mean deviation, variance and standard deviation for the following distribution:

Class	Frequency
37 – 46	19
47 – 56	23
57 – 66	27
67 – 76	28

Solution

STEP 1

Mean :

$$\overline{x} = \frac{\sum f(x)}{\sum f}$$

$$= \frac{5635.5}{97}$$

$$= 58.1$$

STEP 2

class	f	Midpoint	$f(x)$	$\lvert x-\bar{x}\rvert$	$f\lvert x-\bar{x}\rvert$	$(x-\bar{x})^2$	$(x-\bar{x})^2$
37 – 46	19	41.5	788.5	16.6	315.4	2275.56	5235.64
47 – 56	23	51.5	1184.5	6.6	151.8	43.56	1001.88
57 – 66	27	61.5	1660.5	3.4	91.8	11.56	312.12
67 - 76	28	71.5	2002	13.4	375.2	179.56	5027.68
	$\sum 97$		$\sum 5635.5$		$\sum 934.2$		$\sum 11577.32$

STEP 3

i) Mean deviation

$$E = \frac{\sum f|x-\bar{x}|}{\sum f}$$
$$= \frac{934.2}{97}$$
$$= 9.63$$

ii) Variance

$$S^2 = \frac{\sum f(x-\bar{x})^2}{\sum f}$$
$$= \frac{11577.32}{97}$$
$$= 119.35$$

iii) Standard Deviation

$$S = \sqrt{\text{Variance}}$$
$$= \sqrt{119.35}$$
$$= 10.92$$

1.2.5 Quartile, Decile and Percentile

Quartile divide a set of data which are arranged in ascending order into **4 equal parts.**

Decile divide a set of data which are arranged in ascending order into **10 equal parts.**

Percentile divide a set of data which are arranged in ascending order into **100 equal parts.**

Interquartile the value between quartile 1 and quartile 3.

MEASURE	POSITION OF DATA	FORMULA/ HOW TO FIND?
QUARTILES	$Q_k = \frac{k}{4}(n+1)$	**UNGROUPED DATA** Count the position to determine the value
	$Q_k = \frac{k}{4}\sum f$	**GROUPED DATA** $Q_k = L + \left(\frac{\frac{k}{4}N - F}{f_{Q_k}} \times c\right)$
DECILES	$D_k = \frac{k}{10}(n+1)$	**UNGROUPED DATA** Count the position to determine the value
	$D_k = \frac{k}{10}\sum f$	**GROUPED DATA** $D_k = L + \left(\frac{\frac{k}{10}N - F}{f_{D_k}} \times c\right)$

<table>
<tr><th>MEASURE</th><th>POSITION OF DATA</th><th>FORMULA/ HOW TO FIND?</th></tr>
<tr><td rowspan="2">PERCENTILES</td><td>$\boldsymbol{P_k = \frac{k}{100}(n+1)}$</td><td>UNGROUPED DATA

Count the position to determine the value</td></tr>
<tr><td>$P_k = \frac{k}{100}\sum f$</td><td>GROUPED DATA

$\boldsymbol{P_k = L + \left(\frac{\frac{k}{100}N - F}{f_{P_k}} \times c\right)}$</td></tr>
<tr><td>INTERQUARTILE</td><td></td><td>$\boldsymbol{IQR = Q_3 - Q_1}$</td></tr>
<tr><td colspan="3">k = 1, 2 or 3
L = Lower boundary of quartile/ decile/ percentile group
n = number of data
N = Total frequency
F = Cumulative frequency before quartile/ decile/ percentile group
f = frequency of quartile/ decile/ percentile group
c = size of class</td></tr>
</table>

Example 1

Determine Q_1, Q_2, D_8, P_{20} and the interquartile range for the data series

3, 4, 1, 3, 5, 7, 6, 8, 6, 8

Solution

STEP 1 First sort these data in ascending order

1	3	3	4	5	6	6	7	8	8
	P20	Q1		Q2			Q3	D8	

$n = 10$

STEP 2

Quartiles

$$Q_a = \frac{a}{4}(n+1)$$

$$Q_1 = \frac{1}{4}(10+1)$$
$$= 2.75$$
$$\approx 3th$$
$$\therefore Q_1 = 3$$

$$Q_a = \frac{a}{4}(n+1)$$

$$Q_2 = \frac{2}{4}(10+1)$$
$$= 5.5th$$

$$\therefore Q_2 = \frac{5+6}{2}$$
$$= 5.5$$

Deciles

$$D_a = \frac{a}{10}(n+1)$$

$$D_8 = \frac{8}{10}(10+1)$$
$$= 8.8$$
$$\approx 9th$$
$$\therefore D_8 = 8$$

Percentiles

$$P_a = \frac{a}{100}(n+1)$$

$$P_{20} = \frac{20}{100}(10+1)$$
$$= 2.2$$
$$\approx 2th$$

$$\therefore P_{20} = 3$$

Interquartiles

Find Q_3 first

$$Q_3 = \frac{3}{4}(10+1)$$
$$= 8.25$$
$$= 8th$$
$$\therefore Q_3 = 7$$

$$Interquatile = Q_3 - Q_1$$
$$= 7 - 3$$
$$= 4$$

Example 2

Determine Q_3, D_4, P_{20} for the data set below

a) by using formula

b) by using ogive

Class	410 – 419	420 – 429	430 – 439	440 – 449	450 – 459	460 – 469	470 - 479
f	14	20	42	54	45	18	7

Solution

class	f	Boundaries	F	
410 - 419	14	409.5 – 419.5	14	
420 - 429	20	419.5 – 429.5	34	
430 - 439	42	429.5 – 439.5	76	Percentile Group
440 - 449	54	439.5 – 449.5	130	Decile Group
450 - 459	45	449.5 – 459.5	175	Quartile Group
460 - 469	18	459.5 – 469.5	193	
470 - 479	7	469.5 – 479.5	200	

a) using FORMULA

Use formula to locate Q_3, D_4 and P_{20} in the group

$$Q_a = \frac{a}{4}\left(\sum f\right)$$

$$Q_3 = \frac{3}{4}(200)$$

$$= 150^{th}$$

$$D_a = \frac{a}{10}\left(\sum f\right)$$

$$D_4 = \frac{4}{10}(200)$$

$$= 80^{th}$$

$$P_a = \frac{a}{100}\left(\sum f\right)$$

$$P_{20} = \frac{20}{100}(200)$$

$$= 40^{th}$$

Quartile

$$Q_k = L + \left(\frac{\frac{k}{4}N - F}{f_{Q_k}} \times c \right)$$

$$Q_3 = 449.5 + \left(\frac{\frac{3}{4}(200) - 130}{45} \times 10 \right)$$

$$= 449.5 + 4.4$$

$$= 453.9$$

Decile

$$D_k = L + \left(\frac{\frac{k}{10}N - F}{f_{D_k}} \times c \right)$$

$$D_4 = 439.5 + \left(\frac{\frac{4}{10}(200) - 76}{54} \times 10 \right)$$

$$= 439.5 + 0.74$$

$$= 440.2$$

Percentile

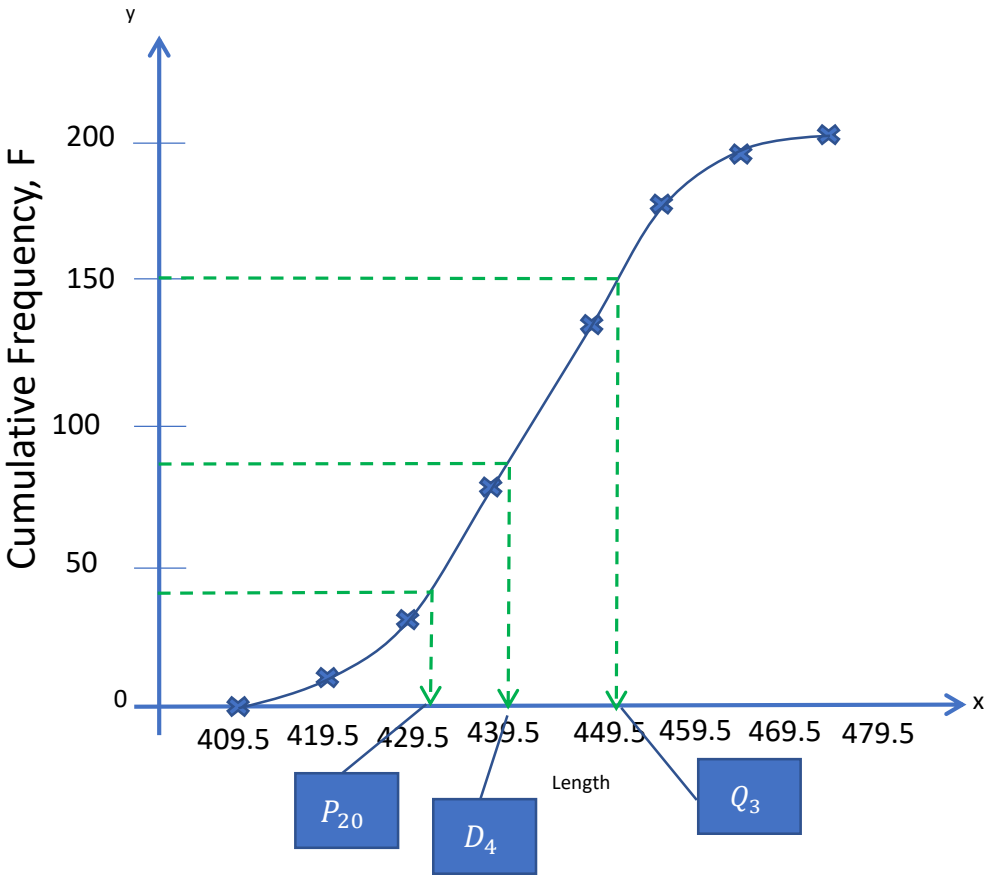
$$P_k = L + \left(\frac{\frac{k}{100}N - F}{f_{P_k}} \times c \right)$$

$$P_{20} = 429.5 + \left(\frac{\frac{20}{100}(200) - 34}{42} \times 10 \right)$$

$$= 429.5 + 1.43$$

$$= 430.93$$

b) using Ogive



$$Q_a = \frac{a}{4}\left(\sum f\right)$$

$$Q_3 = \frac{3}{4}(200)$$

$$= 150^{th}$$

$$\therefore Q_3 = 453.5$$

$$D_a = \frac{a}{10}\left(\sum f\right)$$

$$D_4 = \frac{4}{10}(200)$$

$$= 80^{th}$$

$$\therefore D_4 = 440.5$$

$$P_a = \frac{a}{100}\left(\sum f\right)$$

$$P_{20} = \frac{20}{100}(200)$$

$$= 40^{th}$$

$$\therefore P_{20} = 431$$

2.0 Probability

2.1 Expectation

In probability and statistics, the expectation or expected value, is the **weighted average value of a random variable**.

$$E = P(A) \times Number\ of\ trials$$
$$E = P(A) \times n(S)$$

Example 1

The probability that a computer produced by a factory is defective is $\frac{1}{20}$. If 3000 computers are produced by the factory in a certain period of time, estimate the number of computer that are defective.

Solution

$$P(A) = \frac{1}{20}$$
$$n(S) = 3000$$

$$\therefore E = P(A) \times n(S)$$
$$= \frac{1}{20} \times 3000$$
$$= 150$$

Hence, the number of computer that are defective is 150

Example 2

In a box, there are contain red and blue pens. If a pen is picked random from the box, the probability that is a red pen is $\frac{4}{7}$. If there are 84 red pens in the box, calculate the total number of pens in the box.

Solution

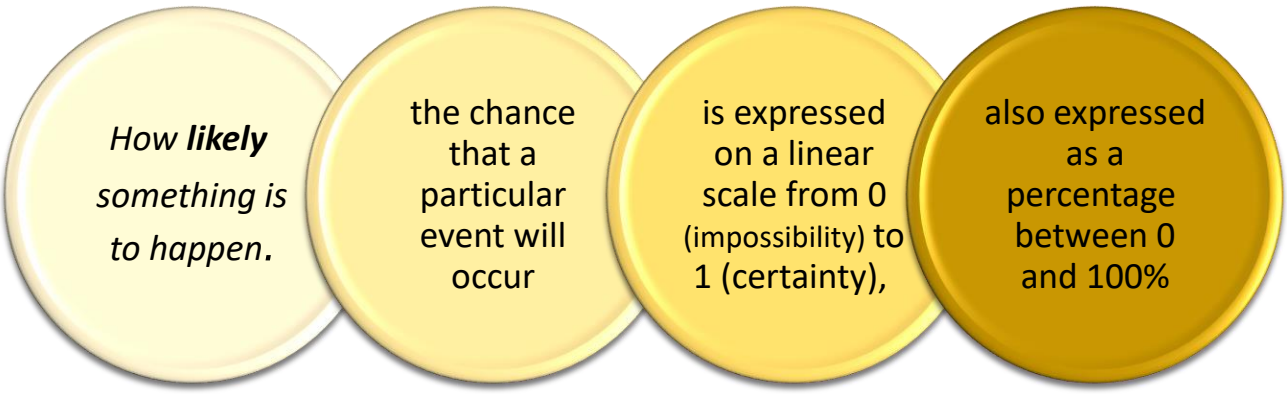
$$P(Red\ pen) = \frac{4}{7}$$
$$Expectation\ red\ pen, E = 84$$

$$\therefore E = P(A) \times n(S)$$
$$84 = \frac{4}{7} \times n(S)$$
$$\mathrm{n}(S) = 147$$

Hence, the total number of pens in the box is 147

2.2 Understand Probability

Probability is the measure of the likelihood that an event will occur in a Random Experiment. Probability is quantified as a number between 0 and 1, where, loosely speaking, 0 indicates impossibility and 1 indicates certainty.

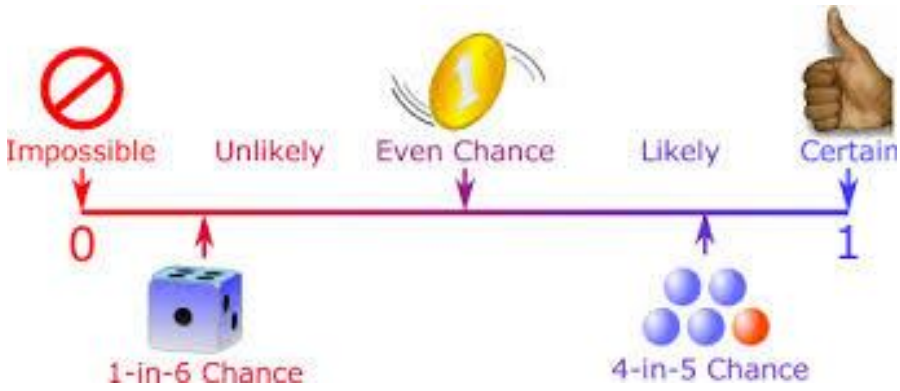


In order to measure probabilities, mathematicians have devised the following formula for finding the probability of an event.

$$\textit{Probability of an event happening} = \frac{\textit{Number of ways it can happen}}{\textit{Total number of outcomes}}$$

$$P(A) = \frac{n(E)}{n(s)}$$

We can show probability on a **Probability Line**



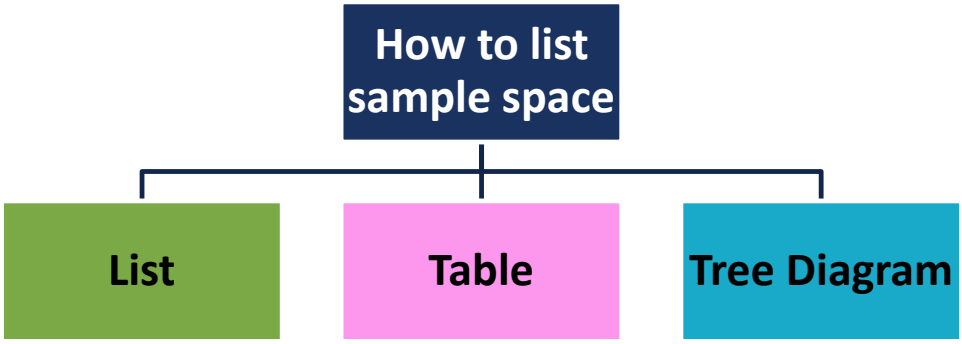
Probability is always between 0 and 1

2.3 Element of Probability

Definition	Examples
An experiment is a situation involving chance or probability that leads to results called outcomes.	- toss a coin once - throwing two dice
An Outcome is a possible result that can be obtained from the experiment	- head, tail - 1, 2, 3, 4, 5, 6
Sample Space is a set of possible outcomes of a random experiment.	- S = { HH,HT,TH, TT } - S = {1, 2, 3, 4, 5, 6 }
Event is a single result of an experiment.	- Getting odd number when throwing a dice. - Getting a Head when tossing a coin

2.3.1 Sample Space

- The sample space is represented using the symbol, "S". A sample space may contain a number of outcomes that depends on the experiment



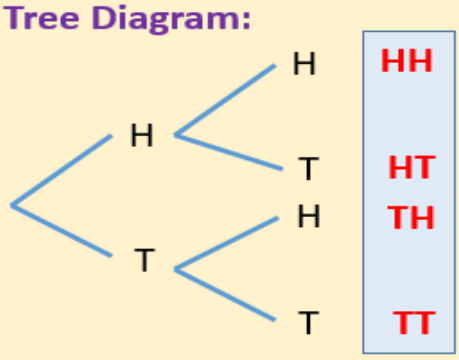
List:

HH HT TH TT

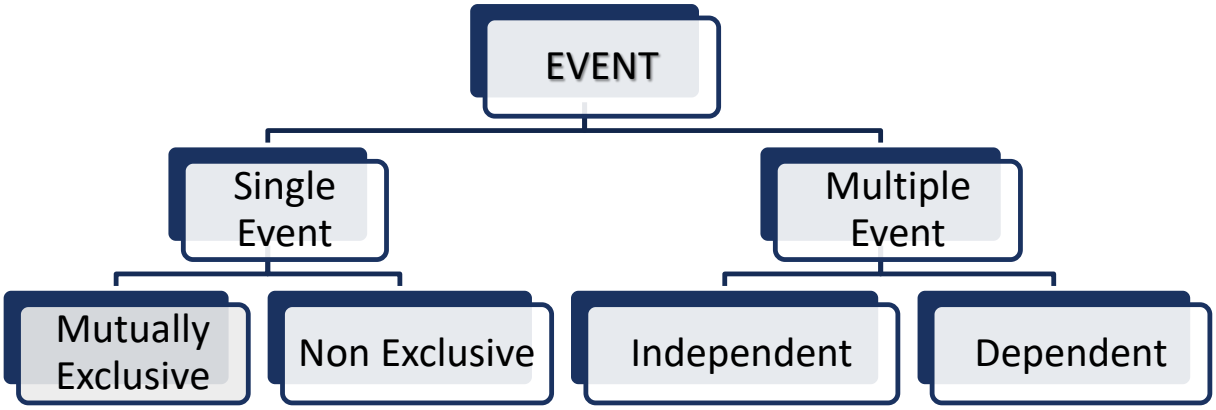
Table:

	H	T
H	HH	HT
T	TH	TT

Tree Diagram:



2.4 Types of Event



2.4.1 Single Event

- **Single-event probability** is used to find the probability for a single event that occurs for an experiment.
- Single event **happen once** and **pick one only**.
- **For example**, consider tossing a coin, we will get single event (either head or tail) as expected result.

How do you find the probability of a single event?

i. Determine a single event with a single outcome
ii. Identify the total number of outcomes that can occur
iii. Divide the number of events by the number of possible outcomes.

2.4.1.1 Mutually Exclusive Event

- An events that **can not happen at the same time**
- **$P(A \cap B) = 0$**
  $P(A \cup B) = P(A) + P(B)$
- **Examples :** Turning left and turning right, Tossing a coin: Heads and Tails

2.4.1.2 Non Exclusive Event

- An events **can happen at the same time**
- **P(A ∪ B) = P(A) + P(B) – P(A and B)**
- **Examples :** even numbers and prime numbers on a die, losing a game and scoring

Example 1

A letter is randomly selected from the word KELANTAN. Find the probability that the selected letter is a

a) Vowel

b) Consonant

Solution

a) sample space; S = { K, E, L, A, N, T, A, N }

n(s) = 8

Event of choosing vowel letter; V = {E, A, A}

n(V) = 3

$P(V) = \frac{3}{8}$

b) Event of choosing consonant letter; C = {K, L, N, T, N}

n(C) = 5

$P(V) = \frac{5}{8}$

Example 2

A card are chosen at random from a pack of 52 playing cards. Find the probability that

a) pick King or Queen

b) pick hearts and Kings

Solution

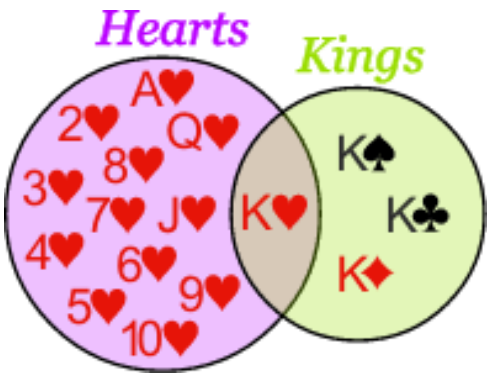
a) The probability of pick a King, $P(\text{King}) = \frac{1}{13}$

The probability of pick a Queen, $P(\text{Queen}) = \frac{1}{13}$

$\therefore P(\text{King or Queen}) = \frac{1}{13} + \frac{1}{13}$

$= \frac{2}{13}$

b) Hearts and Kings



Hearts and Kings together is only the King of Hearts

16 Cards = 13 Hearts + 4 Kings − the 1 extra King of Hearts

$\therefore P(H \cup K) = P(H) + P(K) - P(H \cap K)$

$$= \frac{13}{16} + \frac{4}{16} - \frac{2}{16}$$

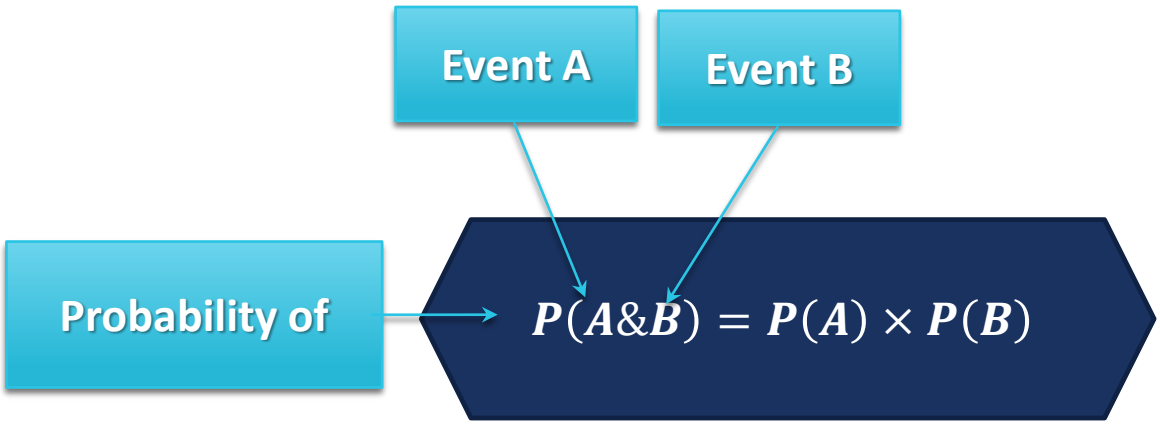
$$= \frac{15}{16}$$

2.4.2 Multiple Event

- The probability of multiple events occurs when we're trying to calculate the probability of observing **two or more events**.
- **Example :** consider tossing a coin twice

2.4.3 Independent Event

- An event that is **not affected by previous events**.
- **Example:** tossing a coin where Heads or tails is not affected by previous tosses and each toss of a coin is a perfect isolated thing.



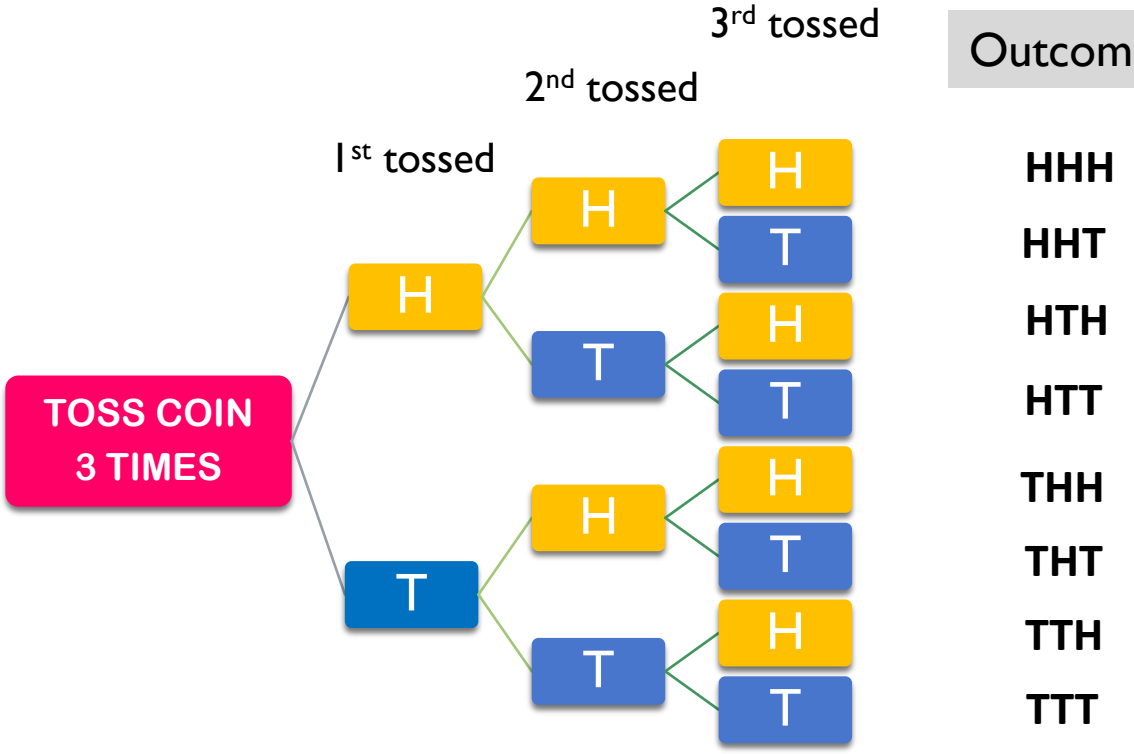
Remember 3 methods to find Sample Space!

- Listing
- Table
- Tree diagram

Example 1

Find the probability of getting a "Head" when tossing a coin three times?

Solution



$$P(Head) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}$$

$$= \frac{1}{8}$$

Example 2

Consider, a die is thrown twice. Calculate the probability of
a) twice gets an odd number
b) the sum of twice is 4

Solution

First, list the sample space

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

a) odd numbers

E = { (1,1), (1,3), (1,5), (3,1), (3,3), (3,5), (5,1), (5,3), (5,5)}

n(S) = 36
n(E) = 9

$$P(odd) = \frac{9}{36}$$
$$= \frac{1}{4}$$

a) the sum of twice is 4

E = { (1,3), (2,2), (3,1)}

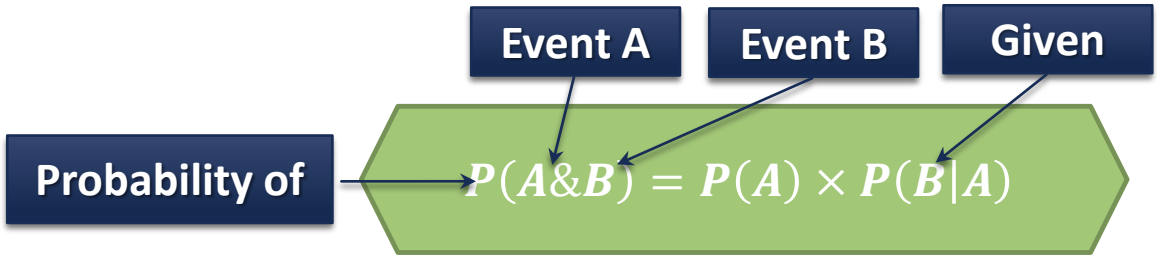
n(S) = 36

n(E) = 3

$$P(odd) = \frac{3}{36}$$
$$= \frac{1}{12}$$

2.4.4 Dependent Event

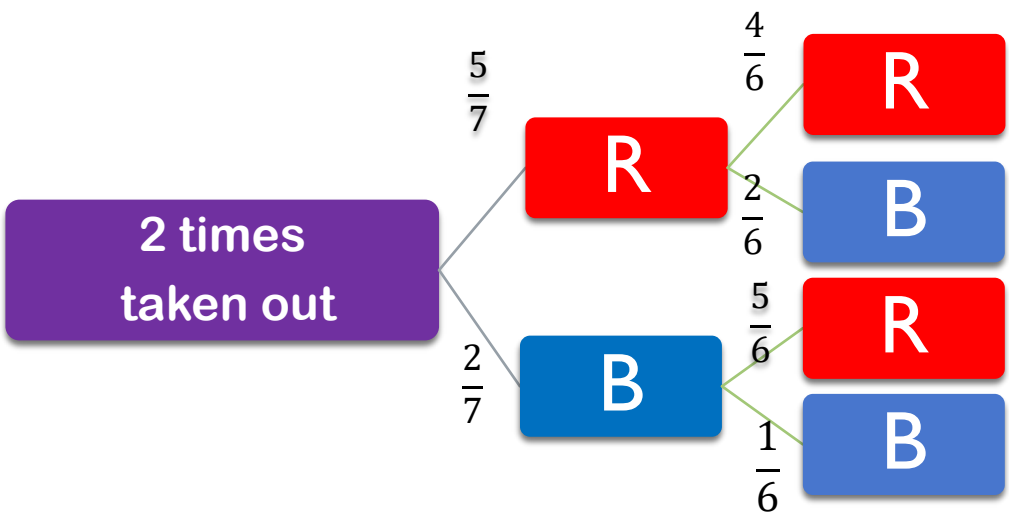
- An event that is **affected by previous events**.
- **Example:** Removing coloured marbles from a bag where each time you remove a marble the chances of drawing out a certain colour will change



Example 1

A bag contain 5 red marbles and two blue marbles. A marble is taken out. Its colour noted and not replaced. A second colour is taken and its colour noted. Find the probability of taking two different colours.

Solution



$$P(2\ different\ colours) = P(R,B) + P(B,R)$$
$$= \frac{5}{7} \times \frac{2}{6} + \frac{2}{7} \times \frac{5}{6}$$
$$= \frac{10}{42} + \frac{10}{42}$$
$$= \frac{10}{21}$$

Example 2

Suppose that we have 50 computers, of which 5 are defective. If 3 computers are chosen at random, find the probability that all are defective.

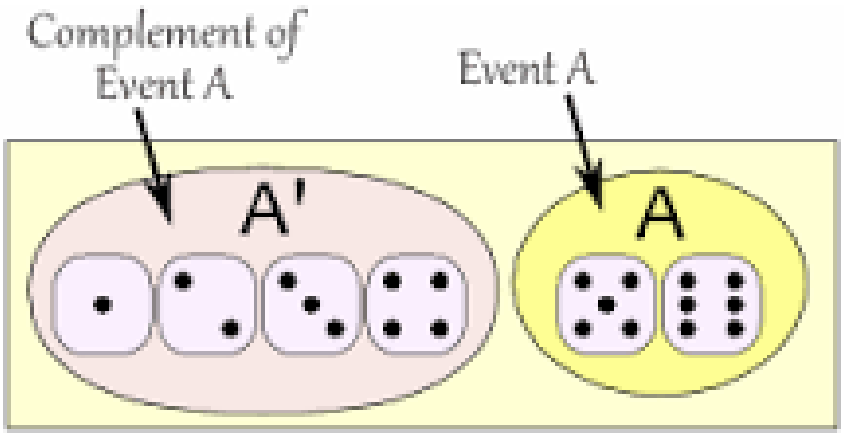
Solution

$$P(defective) = \frac{5}{50}$$

$$P(both\ defective) = \frac{5}{50} \times \frac{4}{49} \times \frac{3}{48}$$
$$= \frac{1}{1960}$$

2.5 Complementary Event

In Probability, the Complement of an event is all outcomes that are NOT the event.



Together the event and its complement make all possible outcomes.

P(A) → Probability of Event A
P(A') → Probability of the complementary of Event A
P(A) + P(A') = I

Example 1

Throw two dice. Find the probability the two scores are **different**?

Solution

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

P(different score) = 1- P(same score)

$$= 1 - \frac{6}{36}$$

$$= \frac{5}{6}$$

Example 2

A lemonade stand has pink and plain lemonade. If the probability of buying pink lemonade is 3/8, what is the probability that the customer does not buy pink lemonade? Of 40 customers, how many do you expect to buy plain lemonade?

Solution

$$P(L) = \frac{3}{8}$$

$$P(L') = 1 - \frac{3}{8}$$
$$= \frac{5}{8}$$

Since the only other option is plain lemonade, P(L') is equal to the probability that the customer buys plain lemonade.

Therefore, of 40 customers we expect:

$\frac{5}{8} \times 40 = 25$ to buy plain lemonade

Example 3

The probability that it will rain in the evening 0.84. What is the probability that it will not rain in the evening?

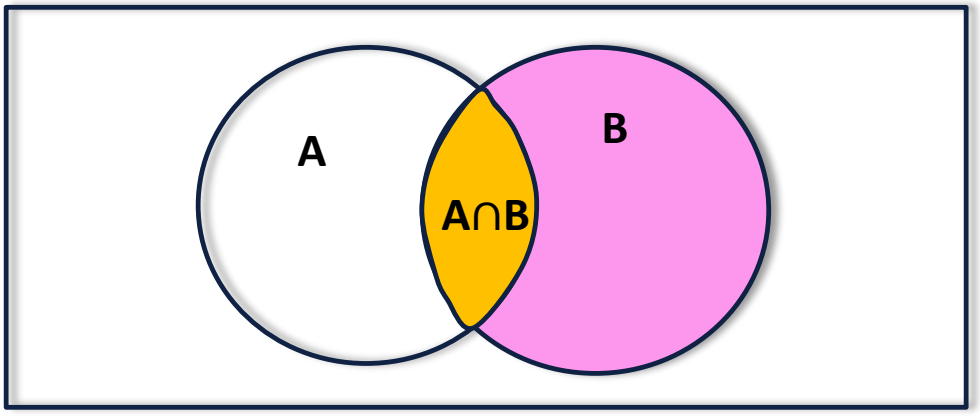
Solution

$$P(not\ rain) = 1\ - P(not\ rain)$$
$$= 1\ - 0.84$$
$$= 0.16$$

2.6 Conditional Probability

Conditional probability: P(A|B) is the probability of event A occurring, given that event B occurs.

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



Example 1

A fair die is rolled, Let A be the event that shows an outcome is an odd number, so A={1, 3, 5}. Also, suppose B the event that shows the outcome is less than or equal to 3, so B= {1, 2, 3}. Then what is the probability of A, $P(A)$, and what is the probability A given B, $P(A|B)$.

Solution

So the solution is for P(A),

total sample space= 6,

Total odd number when rolling dice once= 3

Hence, P(A)= Event A/sample space (S)
= |(1, 3, 5)|/S
= 3/6
= 1/2.

And now, the solution for P(A|B), for calculating conditional probability of A given that B has happened. B has the outcomes {1,2,3} and A has {1, 3, 5}. Here (A⋂B)= {1, 3} that are two numbers.

So, P(A|B)= P(A⋂ B)/ B
= 2 / 3.

Example 2

65% of the 5J students like noodles and 40% like spaghetti and like noodles. What is the probability of those who like spaghetti given that they like noodle?

Solution

Let: A-students that like noodles
B-students that like spaghetti

$$P(A \cap B) = 0.4$$
$$P(A) = 0.65$$

Convert percentage into decimal

Probability of those who like spaghetti given that they like noodle, P(BlA)

$$P(A \cap B) = 0.4$$
$$P(A) = 0.65$$
$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$
$$= \frac{0.4}{0.65}$$
$$= 0.62$$

Example 3

A recent survey asked 100 people if they thought Sardin Cap Ayam is the best sardine. The result of the survey are shown in the table. Find the probabilities;

a) The respondent answered "yes", given that the respondent was a female.

b) The respondent was a male, given that the respondent answered "no"

Gender	YES	NO	Total
Male	32	18	50
Female	8	42	50
Total	40	60	100

Solution

a) The respondent answered "yes", given that the respondent was a female

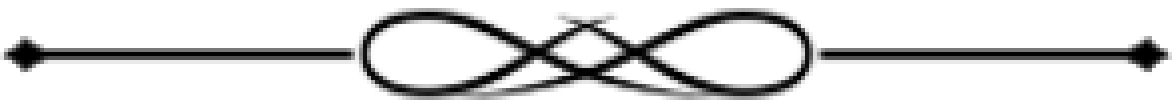
$$P(Y|F) = \frac{P(F \cap Y)}{P(F)}$$

$$= \frac{\frac{8}{100}}{\frac{50}{100}} = \frac{4}{25}$$

b) The respondent was a male, given that the respondent answered “no”

$$P(M|N) = \frac{P(N \cap M)}{P(N)}$$

$$= \frac{\frac{18}{100}}{\frac{60}{100}} = \frac{3}{10}$$



References

Bird, J. (2010). Higher Engineering Mathematics (6th Edition). UK : Newnes

Finney, R.L. & Thomas,G.B. (1993). *Calculus* (2nd Ed.).Boston : Addison – Wesley.

Rahim, C.T. (2005). *Kaedah Berangka : Matematik Untuk Sains dan Kejuruteraan Mengguna Maple*. Johor : Universiti Teknologi Malaysia

Stroud, K. A. (2011). *Advanced Engineering Mathematics* (5th Edition). New York: Industrial Press Inc.

e ISBN 978-967-2897-46-0

www.polipd.edu.my
2022