Question # 1

Randomly survey 30 classmates about the monthly Pocket money, height and understanding level of Probability and statistics course till the time. Record their values.

a) Group these all variables five to Six Classes.
b) Construct an appropriate graph for each case.
c) Are the data's discrete or continuous? How do you know?
d) In complete sentences, describe the shape of each graph.
e) Construct box plots of all data sets and interpret it.
f) Find the suitable measures of central tendency and dispersion for all data sets?
g) Which variable is more consistent?

Collected Data

Index	Weekly Pocket Money (in Rs)	Height (in cm)	Subject Understanding
1	1750	160	95
2	2000	172	80
3	250	165	90
4	500	171	70
5	500	177	80
6	700	167	80
7	3500	167	78
8	1200	170	75
9	3000	170	75
10	500	150	40
11	3000	166	70
12	2500	157	85
13	700	183	60
14	2500	153	74
15	3500	180	89
16	2500	162	56
17	1000	173	70
18	500	165	68
19	750	177	80
20	1000	174	68
21	1700	152	70
22	1000	168	70
23	1750	173	68
24	3000	160	88
25	1000	163	64
26	1500	177	65
27	3000	158	70
28	1000	163	60
29	500	183	60
30	1300	170	72

a) Group variables five to six classes

- Weekly Pocket Money

$X_{min} = 250$ $\qquad$ range $= X_{max} - X_{min}$

$X_{max} = 3500$ $\qquad$ $= \underline{3250}$

$$\text{interval } h = \frac{\text{range } r}{\text{no. of classes}} = \frac{3250}{5}$$

$$h = 650$$

Distribution

Tally	Class	f	Cf
ⅠⅠⅠⅠ̸ ⅠⅠⅠⅠ	250 - 900	9	9
ⅠⅠⅠⅠ̸ ⅠⅠⅠ	900 - 1550	8	17
ⅠⅠⅠⅠ	1550 - 2200	4	21
ⅠⅠⅠ	2200 - 2850	3	24
ⅠⅠⅠⅠ̸ Ⅰ	2850 - 3500	6	30

- Height

$X_{max} = 183$

$X_{min} = 150$

$\text{range} = X_{max} - X_{min}$

$= \underline{33}$

$$\text{interval } h = \frac{33}{5}$$

$$h = 6.6 \Rightarrow 7$$

Distribution

Tally	Class	f	Cf
ⅠⅠⅠⅠ	150 - 157	4	4
ⅠⅠⅠⅠ̸ Ⅰ	157 - 164	6	10
ⅠⅠⅠⅠ̸ ⅠⅠⅠⅠ̸	164 - 171	10	20
ⅠⅠⅠⅠ̸ ⅠⅠ	171 - 178	7	27
ⅠⅠⅠ	178 - 185	3	30

- Subject Understanding

$X_{max} = 95$

$X_{min} = 40$

$\text{range} = X_{max} - X_{min}$

$= 55$

interval $h = \frac{55}{5}$

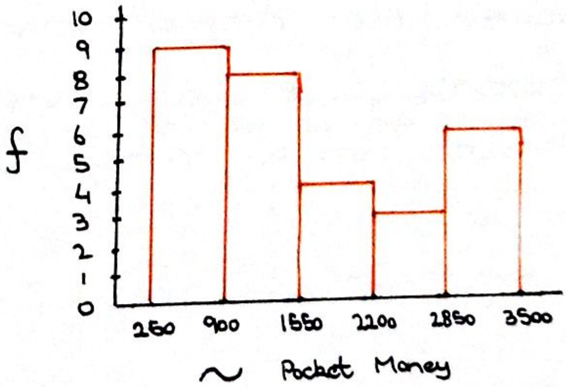
$h = 11$

Distribution

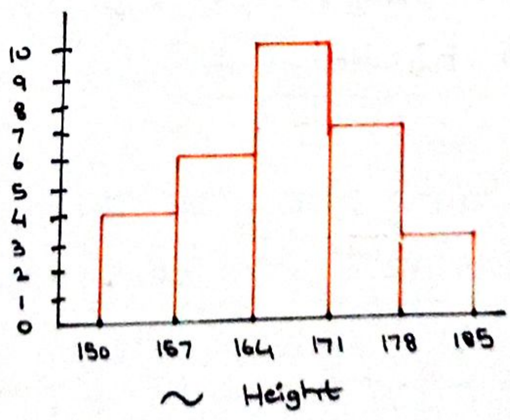
	Class	f	Cf
|	40 - 51	1	1
||||	51 - 62	4	5
卌 卌 ||	62 - 73	12	17
卌 |||	73 - 84	8	25
卌	84 - 95	5	30

b) Construct an appropriate graph

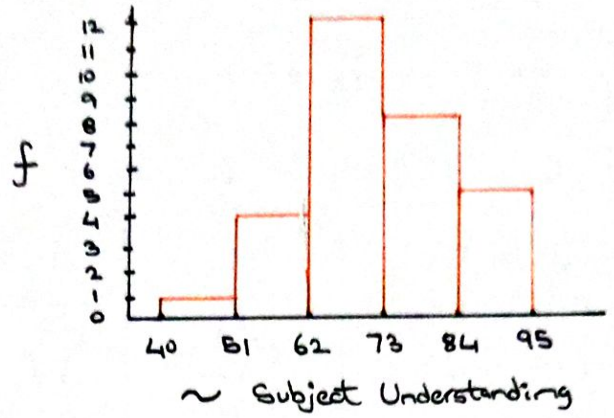
- Weekly Pocket Money (Histogram)



- Height



• Subject Understanding



c) Are data values discrete or continuous?

All data values (of three given variables) are continuous in nature. We can tell this distinction from the nature of data as it is measurable rather than countable. In other words, our data can assume numerical values in a range of numbers.

d) Describe shape of each graph.

• Pocket Money: Since the frequencey increases → decreases → increases again, we can qualify / define our data as a U-shaped histogram.

• Height: Since our data is not too skewed on either side, we can classify it as a bell shaped histogram.

- Subject Understanding: As frequency of data on the higher end is significantly higher than that on the lower end, we can classify it as left-skewed histogram.

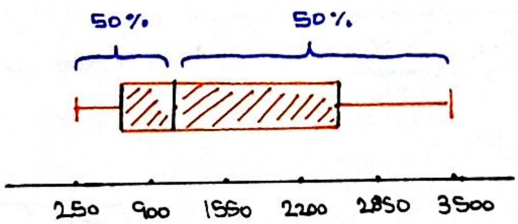
c) Construct box plots

- Pocket Money:

$Q_1 = 30(1/4)^{th}$ value $= 700$

$Q_2 = 30(1/2)^{th}$ value $= 1/2(1200 + 1300) = 1250$

$Q_3 = 30(3/4)^{th}$ value $= 2500$



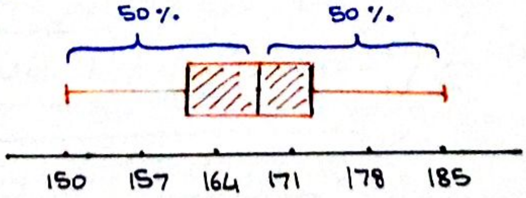
» Most of our data lies on ends; 50% of our values lie in between 250 and 1250.
dense

- Height:

$Q_1 = 160$

$Q_2 = 167.5$

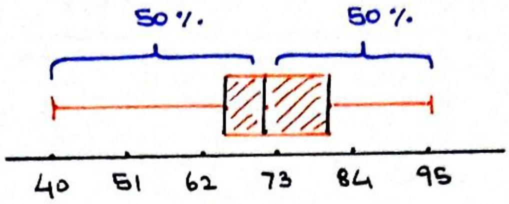
$Q_3 = 173$



» Data is more or less evenly distributed with most lying between 160 and 173.

- Subject Understanding:

$Q_1 = 65$
$Q_2 = 70$
$Q_3 = 80$



>> Data is left skewed (Densely packed on the right of median Q_2.

f) Suitable measure of central tendency and dispersion

→ Mean $= \frac{\Sigma x}{n}$ Median $= \frac{1}{2}\left(n_{th} + (n+1)_{th} \text{ term}\right)$
(since n is even)

→ Standard Deviation $= \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n}}$

- Weekly Pocket Money

$$\text{Mean} = \frac{250 + 500 + \dots + 3500}{30} = 1586.67$$

- However, since the data is U-shaped, we prefer median.

$$\text{Median} = \frac{1}{2}(1200 + 1300) = \boxed{1250}$$

$$\text{Standard Dev.} = \sqrt{\frac{(250 - 1586.6)^2 + \dots + (3500 - 1586.6)^2}{30}}$$

$$\sigma = \boxed{999.57}$$

- Height

$$\text{Mean} = \frac{150 + \ldots\ldots + 183}{30} = \boxed{167.53}$$

$$\text{Standard Dev.} = \sqrt{\frac{(160 - 167.53)^2 + \ldots\ldots + (183 - 167.53)^2}{30}}$$

$$\sigma = \boxed{8.597}$$

- Subject Understanding

$$\text{Mean} = \frac{40 + \ldots\ldots + 95}{30} = \boxed{72.33}$$

$$\text{Standard Dev.} = \sqrt{\frac{(40 - 72.33)^2 + \ldots\ldots + (95 - 72.33)^2}{30}}$$

$$\sigma = \boxed{11.217}$$

g) Which variable is more consistent?

$$\text{Coefficient of variation (C.V)} = \frac{\sigma}{\mu} \times 100\%$$

- Pocket Money (C.V) = 62.99 %
- Height (C.V) = 5.13 %
- Subject Understanding (C.V) = 15.50 %

» "From above calculations, we can see that variable "Height" is most consistent."

HASSAN RIZWAN
335753

Question #2 :

(i) Cumulative frequency curve is sketched below and the axis are as follows.

X - Axis : Time taken (minutes)

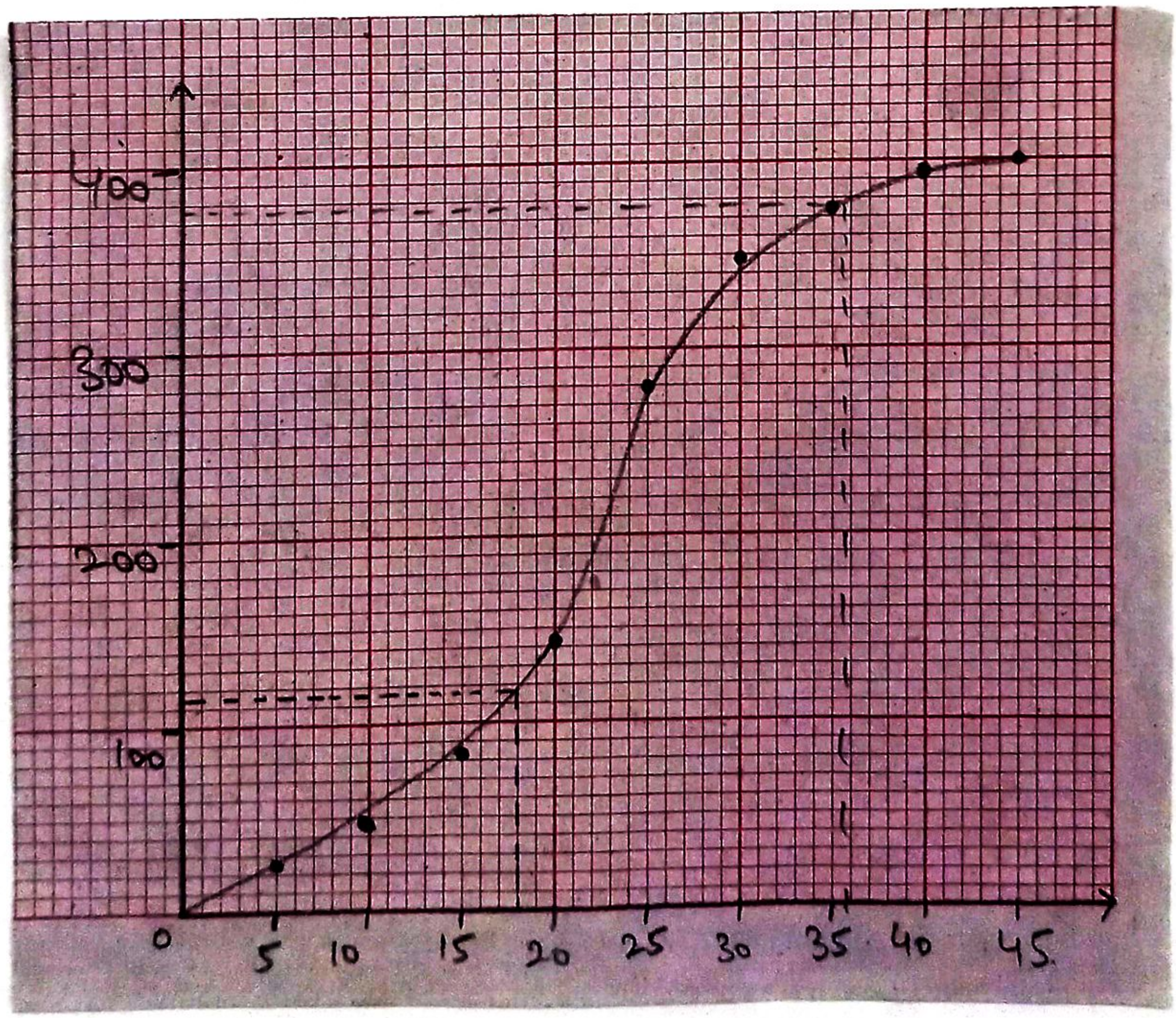
Scale : 10 minutes per solid marker line

Y-axis : Cumulative frequency

Scale : 100 per solid marker line.

(a)

Estimate: $\boxed{\text{115 People}}$ took less than 18 minutes



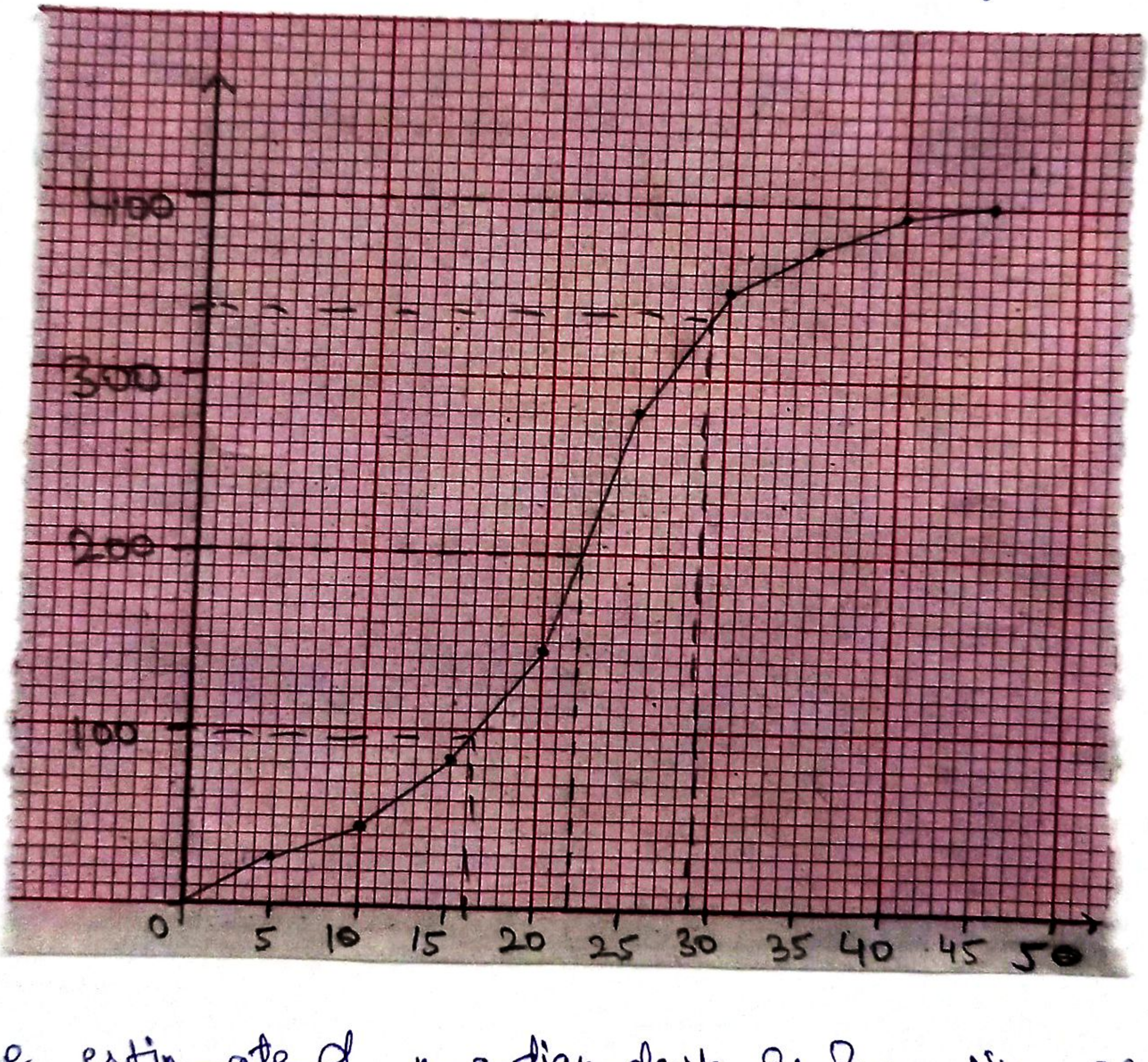
(b) 6% of 400 = 24

400 − 24 = 376.

Time taken for 376 freq. = 36.

$\boxed{\text{36 minutes or longer}}$

(c) Ogive: (Same scale, X-axis and Y-axis values as the cumulative frequency curve).



The estimated median, decile & percentile are as follows:

Median = 22
10th decile = 16
98th percentile = 29

Topic: Assignment: 2

Question: 2
According to the given data:

Classes	F	x_i	Fx_i	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	$(x_i-\bar{x})^3$
15-19	291	17	4947	-22.83	521.208	-11899.20
20-24	766	22	16852	-17.83	317.909	-5668.31
25-29	1287	27	34749	-12.83	164.608	-2111.93
30-34	2333	32	74656	-7.83	61.308	-480.041
35-39	4245	37	157065	-2.83	8.008	-22.66
40-44	6582	42	276444	2.17	4.709	10.218
45-49	3995	47	187765	7.17	51.409	368.60
50-54	1999	52	103948	12.17	148.109	1802.48
Σ	21498		856426			

Now calculating $\bar{x}$, we observe:

$$\bar{x} = \frac{\Sigma f x}{\Sigma f}$$

$$= \frac{856426}{21498}$$

$$\bar{x} = 39.83$$

Now using this in the above table

Now also;

$$\Sigma f_i(x_i-\bar{x})^2 = 1816515.332$$

$$\Sigma f_i(x_i-\bar{x})^3 = -6595831.49$$

Now calculating m_2 & m_3 by using formulas we have:

$$m_2 = \frac{\Sigma f_i (x_i - \bar{x})^2}{\Sigma f_i}$$

$$= \frac{1316515.332}{21498}$$

$$= 61.2389$$

Similarly;

$$m_3 = \frac{\Sigma f_i (x_i - \bar{x})^3}{\Sigma f}$$

$$= \frac{-6595831.49}{21498}$$

$$= -306.811$$

Thus;

$$b_1 = m_3 / m_2^{3/2}$$

By putting the values:

$$b_1 = \frac{-306.811}{(61.2389)^{3/2}}$$

$$= -0.6402$$

Result:

The result is a negative number, thus

Kurtosis:

Kurtosis characterizes the relative peakedness or flatness of a distribution compared to normal distribution.

The formula to find Kurtos is given as:

$$b_2 = \frac{m_4}{(m_2)^2} \rightarrow (i)$$

We already have found m_2.

Now finding m_4, We have:

$$\Sigma f (x_i - \bar{x})^4 = 254938274$$

$$m_4 = \frac{\Sigma f_i (x_i - \bar{x})^4}{\Sigma f_i}$$

By putting the values;

$$m_4 = \frac{234938274}{21498}$$

$$= 11858.697$$

Thus, by putting the values in ①

$$b_2 = \frac{m_4}{(m_2)^2}$$

$$b_2 = \frac{11858.697}{(61.2389)^2}$$

$$b_2 = 3.162$$

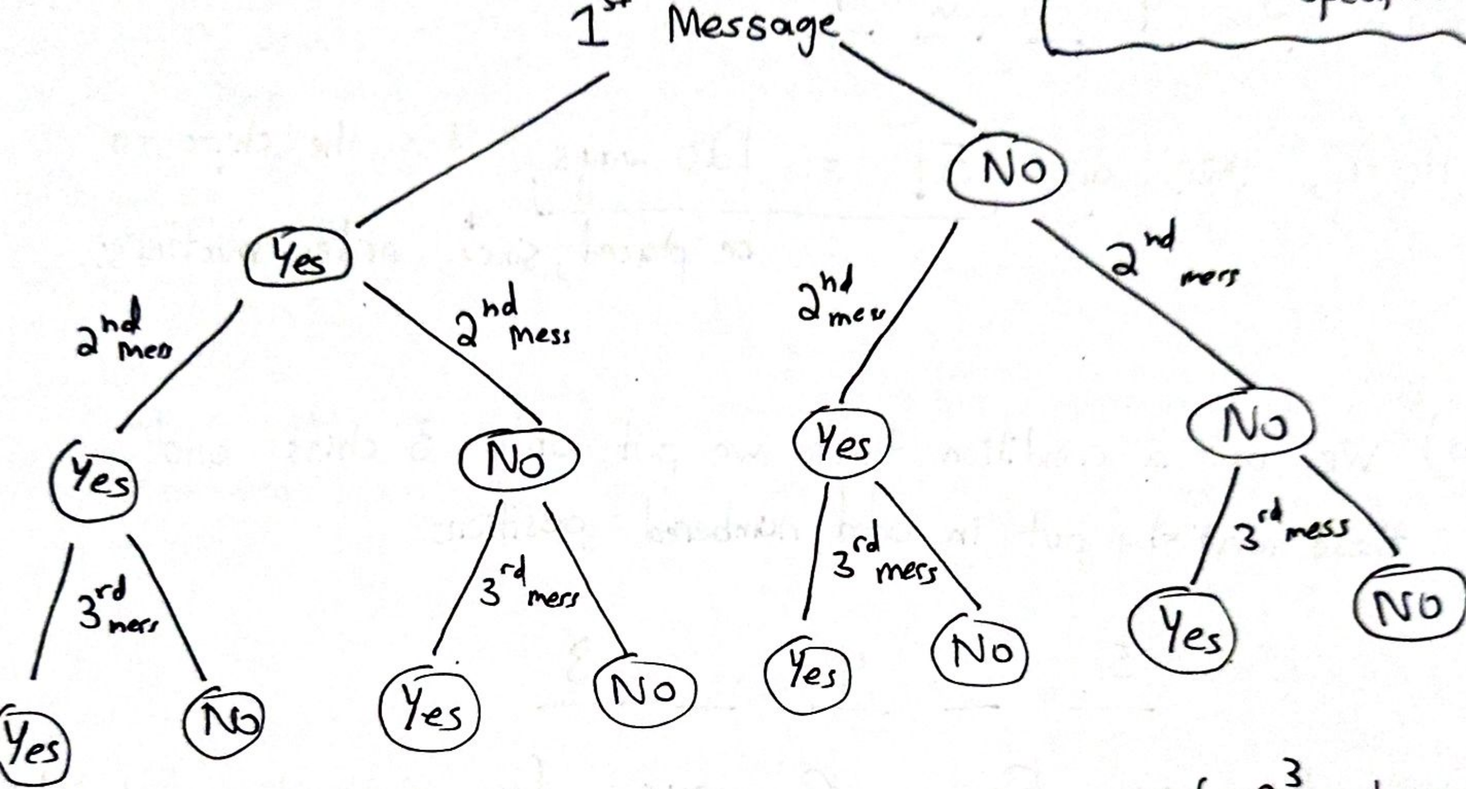
Result:

The result is approximately equal to 3 thus, Question 3;

1)

Key: Yes: recieved within specified time

No: not recieved within specified time



Hence we can see the sample space, which consists of 2^3 outcomes

(YYY) (NYY)
(YYN) (NYN)
(YNY) (NNY)
(YNN) (NNN)

8 outcomes.

2) There are 5 distinct memory chips.

a) When we consider there to be five positions in the microcontroller;

$\underline{5}\ \underline{4}\ \underline{3}\ \underline{2}\ \underline{1}$

Hence, there are $\underbrace{5! = 120 \text{ ways}}$, for the chips to be placed, since order matters.

b) We put a condition that we put only 3 chips and those must be put in odd numbered positions

$\underline{5}\ \underline{\quad}\ \underline{4}\ \underline{\quad}\ \underline{3}$

Hence there are $\underbrace{{}^{5}P_{3} = 60 \text{ ways}}$ for chips to be placed

Q3

3) An Engineering Association consist of 5 civil engineers and 5 mechanical engineers.

a) In how many ways can a committee of 3 civil engineers & 2 mechanical engineers be appointed ?

In choosing 3 civil engineers out of 5, there will be no repetition and order does not matter. Similar will be the case for 2 mechanical engineers.
Thus using combination.

i) $^5C_3 = 10$ ii) $^5C_2 = 10$

Since we are finding possible ways between Civil & (and) Mechanical, we will multiply the above combinations.

$$^5C_3 \times {^5C_2} = 10 \times 10$$
$$\boxed{= 100} \text{ Ans.}$$

A committee of 3 civil and 2 mechanical can be appointed in 100 ways.

b) If 2 civil engineers disagree with each other & refuse to be on the same committee together, how many different ways can a committee of 3 civil engineers & 2 mechanical be appointed.

Here we can have 3 types of sample spaces for civil engineers.

1) We ignore one of the 2 civil engineers disagreeing i.e $^4C_3$
2) Similarly we can ignore the other i.e $^4C_3$
3) We ignore both of the disagreeing engineers .. i.e $^3C_3$

Since these are or conditions, we add them.

$$2 \times {^4C_3} + {^3C_3} = 9$$

The combinations for the 2 mechanical engineers will reamam the same :- ie.

$$^5C_2 = 10$$

Thus total ways $= 9 \times 10$
$= 90$

The total ways in which a committee of 3 civil & 2 mechanical engineers, where 2 civil engineers does not agree with one another are 90.

4) Suppose that 3 balls are placed at random into 3 cells, where more than 1 ball is can be placed in a cell. What are the total possibilities that all cells are occupied? List all outcomes of sample space.

→ Let the cells be L_1, L_2 & L_3.

→ Assuming all balls are identical.

→ Sample space :-

L_1	L_2	L_3
3	0	0
2	1	0
2	0	1
1	1	1
1	0	2
1	2	0
0	2	1
0	1	2
0	3	0
0	0	3

These are the 10 total sample spaces.

In these there is only 1 possibility where all cells are occupied.

→ Assuming the balls are not identical.
Let they be R G B.

→ Sample space :-

L_1	L_2	L_3
RGB	0	0
RG	B	0
RG	0	B
RB	G	0
RB	0	G
BG	R	0
BG	0	R
R	BG	0
0	BG	R
G	BR	0
0	BR	G
B	RG	0
0	RG	B
B	0	RG
0	B	RG
G	0	RB
0	G	RB
R	0	BG
0	R	BG
R	G	B
R	B	G
B	R	G
B	G	R
G	B	R
G	R	B
0	GBR	0
0	0	RGB

Total Possibilities = $3 \times 3 \times 3 = 27$

Possibilities where all cells are occupied = 6

Question No 5:-

Sample size = 4 men.

a) no man gets his own watch back.

Let A define the probability that one man gets his watch back

$$P(A) = \frac{1}{4}.$$

The probability that he doesn't get his watch back

$$P(\bar{A}) = 1 - \frac{1}{4} = \frac{3}{4}.$$

For all men, if no one gets their watch back. (let it be denoted by B)

$$P(B) = \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4}$$

$$= \frac{81}{256}$$

$$P(B) = 0.316$$

b) Exactly one man gets his own watch back.

The probability of one man getting his watch back as calculated in part (a) is.

$$P(A) = \frac{1}{4}.$$

For exactly one man, the other men do not get their watch back.

Let P(B) denote required probability.

$$P(B) = \frac{1}{4} \times \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4}$$

$$= \frac{27}{256}$$

$$= 0.105$$

c) At least three men get their own watch back?

The required probability can be found out as

P(X) = 1 − (Probability of one man getting his watch back) B
− (Probability of two men getting their watch) C
− (Probability of no man getting his watch) A

$P(A) = \frac{81}{256}$.

$P(B) = \frac{27}{256}$

$P(C) = \frac{1}{4} \times \frac{1}{4} \times \frac{3}{4} \times \frac{3}{4} = \frac{9}{256}$

$P(X) = 1 - \frac{81}{256} - \frac{27}{256} - \frac{9}{256}$

$P(X) = \frac{139}{256} = 0.54$.

Question 6:-

Sample size = 5 balls.

Boxes = 3.

Balls are different but all boxes are identical.

To fill all boxes, all boxes must have atleast one ball. The ways to arrange the other two balls would be:-

$\underbrace{(2,2,1)}_{B}$ or $\underbrace{(3,1,1)}_{A}$.

$= \underbrace{{}^5C_3 + {}^2C_1}_{A} + \underbrace{{}^5C_2 + {}^3C_2}_{B}$

Total Combinations = 25

Question - 7

Password contains 5 letters and they cannot repeat.

Sample size = 26.

Total combinations = without replacement

$^{26}C_1\ ^{25}C_1\ ^{24}C_1\ ^{23}C_1\ ^{22}C_1$

| 26 | 25 | 24 | 23 | 22 |

$= 26 \times 25 \times 24 \times 23 \times 22$

Event (A) = 7893600

* Combinations in which the password consists of
- no vowels.

Subtracting 5 vowels from the sample space.

Sample size = 26 - 5 = 21

| 21 | 20 | 19 | 18 | 17 |

Combinations = $21 \times 20 \times 19 \times 18 \times 17$

Event (B) = 2441880

Probability = $\frac{\text{Event B}}{\text{Event A}} = \frac{2441880}{7893600}$

Probability of no vowels = 0.309

- Probability that the first letter is m.

| m | 25 | 24 | 23 | 22 |

combinations = $1 \times 25 \times 24 \times 23 \times 22$

= 303600

Probability of first letter being m:-

$$= \frac{303600}{7893600}$$

$$= \frac{1}{26} = 0.03.$$