UTM
UNIVERSITI TEKNOLOGI MALAYSIA

ASSIGNMENT 2

COURSE NAME : SECI2143 PROBABILITY & STATISTICAL DATA ANALYSIS
SECTION : 01
TITLE : CHAPTER 3 & CHAPTER 4
LECTURER'S NAME : DR SEAH CHOON SEN
GROUP MEMBERS : -

NAME	MATRIC NUMBER
ARULMURUGAN A/L SUBRAMANIYAM	SX210291ECJHS01

Ref: Date:

Question 1

a) Mean

Class Interval	Frequency, fi	Midpoint, xi	fixi
$150 \leq x < 160$	12	155	1860
$160 \leq x < 170$	20	165	3300
$170 \leq x < 180$	5	175	875
$180 \leq x < 190$	3	185	555
Total	40		6590

$$\bar{x} = \frac{\sum_{i=1}^{n} fixi}{n} = \frac{6590}{40}$$

Mean = 164.75

b) Median

Class Interval	Frequency	Class Boundaries	Cumulative Frequency
$150 \leq x \leq 160$	12	149.5 - 159.5	12
$160 \leq x \leq 170$	20	159.5 - 169.5	32
$170 \leq x < 180$	5	169.5 - 179.5	37
$180 \leq x < 190$	3	179.5 - 189.5	40
Total	40		

$$\text{Median} = \frac{L + \frac{N}{2} - cfp}{fmed}(w)$$

$N \div 2 = 40 \div 2 = 20$, ∴ Median class = $160 \leq x < 170$

$L = 159.5$

$N = 40$

$cfp = 12$

Ref: Date:

$w = 169.5 - 159.5 = 10$

$f_{med} = 20$

Thus, Median $= 159.5 + \dfrac{\frac{40}{2} - 12}{20}(10)$

$= 159.5 + (0.4)(10)$

$= 163.5$ ✱

c) Mode.

$$\text{Mode} = l + h \times \frac{(f_1 - f_0)}{(2f_1 - f_0 - f_2)}$$

Most occurance frequency = 20 (From table)
Thus Modal class $= 160 \leq x < 170$.

$l = 160 - 0.5 = 159.5$
$h = 10$
$f_1 = 20$
$f_0 = 12$
$f_2 = 5$

$$\text{Mode} = 159.5 + 10 \times \frac{(20-12)}{(2(20) - 12 - 5)}$$

$$= 159.5 + \frac{8}{(40 - 12 - 5)} \times (10)$$

$= 159.5 + (3.478)$
$= 162.98$ ✱✱

Ref: Date:

d) Model class = $160 \leq x < 170$ //

Question 2

a) Mean, $\mu = \frac{\sum_{i=1}^{N} x_i}{N}$

$N = 10$

$$\frac{85+90+75+88+92+80+85+82+90+85}{10}$$

$$= \frac{852}{10}$$

Mean = 85.2 //

Median

→ Re-arrange list,

75, 80, 82, 85, 85 | 85, 88, 90, 90, 92

Middle Number = 85, 85

Median = Mean of two middle numbers

$$= \frac{85+85}{2}$$

Median = 85 //

Ref: Date:

Mode = Most occurance number

Thus, the mode of participant scores = 85 ✱

b. The mean is 85.2, the median is 85, while the mode is also 85. Since the mean, median and mode are all very close in value (all around 85), it means that the participants scores are symmetrically distributed and there are no significant outliers or extreme values.

For this score lists, the mean is the most appropriate statistic to represent the overall summary of the participation scores because the distribution appears to be balanced, no large gaps between all three value measured and it accurately reflects the central tendency when there is no skewness.

c)

ii) Mean, $\mu = \frac{\sum_{i=1}^{N} x_i}{N}$

$N = 10$

$$\text{mean} = \frac{55 + 65 + 65 + 70 + 85 + 95 + 95 + 95 + 100 + 100}{10}$$

$$= \frac{825}{10}$$

Mean = 82.5 ✱

Ref: Date:

Median

55, 65, 65, 70, 85, 95, 95, 95, 100, 100

Middle number = 85, 95

Median = Mean of two middle numbers

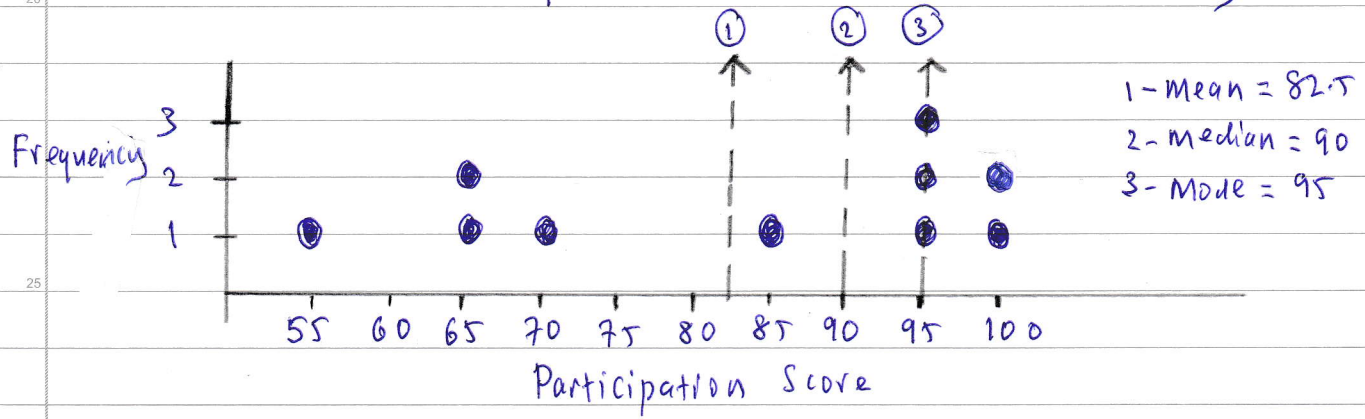
$$= \frac{85 + 95}{2}$$

Median = 90 ✱

Mode = Most occurance number.

The mode of the participation scores = 95 ✱✱

ii) Dot Plot of Participation Scores with Central Tendency



iii) Compared with original data where scores were consistent and closed to 85 points, the corrected data shows more varied scores, with some doing really well ~~(95-100)~~ (95-100) and others not as much (55 - 65). This shows inconsistent participations scores and slightly right skewed. due to Mean < Median < Mode, caused by few low scores pull the mean down.

Ref: Date:

Question 3

a)

i) Range = $X_h - X_1$ (X_h – maximum)
(X_1 – minimum)

Thus from table, $X_h = 40$, $X_1 = 25$

Range = $40 - 25 = 15$

ii) variance $\sigma^2 = \dfrac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}$

use population formula due to data for whole year (Jan - Dec)

Find Mean Value,

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

$N = 12$

$$\mu = \frac{25 + 30 + 28 + 35 + 32 + 27 + 40 + 38 + 33 + 36 + 31 + 29}{12}$$

$$= \frac{384}{12}$$

Mean = 32.0

Ref: Date:

Month	Sales (x_i)	$(x_i - \mu)^2$
January	25	$(25-32)^2 = 49$
February	30	4
March	28	16
April	35	9
May	32	0
June	27	25
July	40	64
August	38	36
September	33	1
October	36	16
November	31	1
December	29	9
Total		230

$$\text{Variance} : \sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}, \quad N = 12$$

$$= \frac{230}{12}$$

$$= 19.17 \text{ ✱}$$

ii) Standard deviation

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}}$$

$$= \sqrt{19.17}$$

$$= 4.38 \text{ ✱}$$

Ref: Date:

b) The range ~~Also~~ of 15 means there is a fair difference between the highest and lowest monthly sales. The standard deviation of 4.38 shows that most monthly sales are usually within about 4.38 units of the average. The variance of 19.17 also shows that the sales numbers do change, but not by huge amount. Overall, the monthly sales change from month to month, but not too much. Some months are better than others, but the sales are mostly steady throughout the year.

c) Understanding the changes in monthly sales can helps the bussiness owner to plan better. For example they can plan their inventory in advance such as ordering stocks more for busy months ~~so~~ like July (40) and August (38) ~~and~~ while order less for slower months. The bussiness owner also can plan staff, promotion, production baled on their sales figures differences. They need to focus on marketing for low sales months like January and June to help keep sales steady all year.

Ref: Date:

Question 4

a) Mean = 50

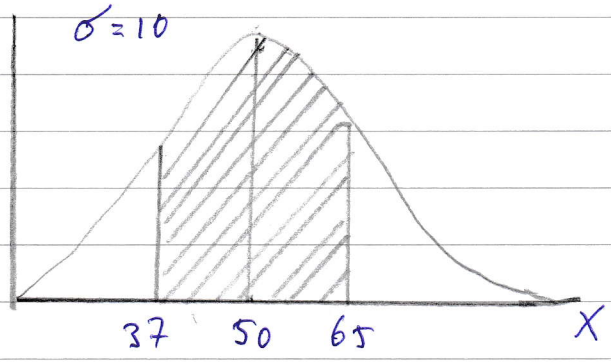
$x > 50$ (Improvement)

That means 50% increased in productivity. ✱

b)

$P(37 \leq x \leq 65)$

mean = 50, standard deviation = 10.



$$z = \frac{x-\mu}{\sigma} = \frac{37-50}{10} = -1.3$$

$$z = \frac{x-\mu}{\sigma} = \frac{65-50}{10} = 1.5$$

$P(z = -1.3) = 0.5 - 0.09680 = 0.4032$

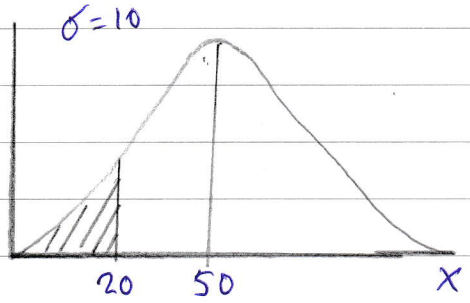
$P(z = 1.5) = 0.93319 - 0.5 = 0.43319$

$\therefore = 0.4032 + 0.43319$

$= 0.83639$ ✱

Ref: Date:

c) $P(20 \leq x)$



$$z = \frac{x-\mu}{\sigma} = \frac{20-50}{10} = -3$$

$P(z=-3) = 0.00135$

Employee needs team building = 1000×0.00135
$= 1.3 \approx 2$ person.

Budget for the team building cost = RM200 x 2
= RM400 ✱

d) Top productivity employees = 5%

→ 5% = 0.05
= 1 - 0.05
= 0.95 ⟶ From z score, $z = 1.65$

$$z = \frac{x-\mu}{\sigma}$$

$$1.65 = \frac{x-50}{10}$$

$1.65(10) = x - 50$

$16.5 = x - 50$

$x = 16.50 + 50$

$X = 66.50$ ✱

Ref: Date:

Question 5

a) X = The number of correct answer a student gets on the 6 multiple choice questions.

b) $n = 6$

$p = \frac{1}{4} = 0.25$ (probility of correct answer)

Table Coultruction with Binomial Distribution.

$$b(x; n, p) = \binom{n}{x} (p)^x q^{x-n}.$$

$q = 1 - 0.25 = 0.75$

X (Correct Answer)	Probability $P(X = x)$	Calculation
0	$\binom{6}{0}(0.25)^0(0.75)^{6-0} = 0.178$	$1 \times 1 \times 0.178$
1	$\binom{6}{1}(0.25)^1(0.75)^{6-1} = 0.356$	$6 \times 0.25 \times 0.316$
2	$\binom{6}{2}(0.25)^2(0.75)^{6-2} = 0.296$	$15 \times 0.0625 \times 0.316$
3	$\binom{6}{3}(0.25)^3(0.75)^{6-3} = 0.132$	$20 \times 0.015625 \times 0.422$
4	$\binom{6}{4}(0.25)^4(0.75)^{6-4} = 0.033$	$15 \times 0.003906 \times 0.562$
5	$\binom{6}{5}(0.25)^5(0.75)^{6-5} = 0.004$	$6 \times 0.000976 \times 0.75$
6	$\binom{6}{6}(0.25)^6(0.75)^{6-6} = 0.0002$	1×0.000244

Ref: Date:

c)

Mean : $\mu = n \times p$

$n = 6, \quad p = 0.25$

$\mu = 6 \times 0.25 = 1.5$ ✱

d) At least three correct answers :

$P(x \geq 3) = P(X=3) + P(x=4) + P(X=5) + P(X=6)$

From table in question 5b,

$P(X=3) = 0.132$
$P(X=4) = 0.033$
$P(X=5) = 0.004$
$P(X=6) = 0.0002$

$P(x \geq 3) = 0.132 + 0.033 + 0.004 + 0.0002$
$= 0.1692$ ✱

e) Probability for correct answer = 0.25
Probability for incorrect answer = 0.75.

$g(x;p) = (1-p)^{x-1} p$

$X = 4$

$g(4;0.75) = (1-0.75)^{4-1} (0.75)$

$= 0.01171$ ✱✱

Ref: Date:

Question 6

a) Negative Binomial distribution

$b^*(x; k, p)$

$b^*(x; 4, 0.7)$

b) Standard Deviation :

$$\sigma = \sqrt{\frac{r(1-p)}{p^2}}$$

$$= \sqrt{\frac{4(1-0.7)}{(0.7)^2}}$$

$$= \sqrt{\frac{4(0.30)}{0.49}}$$

$$= \sqrt{\frac{1.2}{0.49}}$$

$$= \sqrt{2.44898}$$

$$= 1.56 \text{ ✱}$$

Ref: Date:

c) $X = 6$

$p = 0.7$

$k = 4$

$q = 1 - 0.7 = 0.3$

$$b^*(x; k, p) = \binom{x-1}{k-1} p^k q^{x-k}$$

$$b^*(6; 4, 0.7) = \binom{6-1}{4-1} (0.7)^4 (0.3)^{6-4}$$

$$= \binom{5}{3} (0.7)^4 (0.3)^2$$

$$= 10 (0.2401) (0.09)$$

$$= 0.2161 \text{ ✱}$$

d) $x = 7, n = 12, p = 0.7, q = 0.3$

$$\text{Binomial} = b(x; n, p) = \binom{n}{x} (p)^x (q)^{n-x}$$

$$b(7; 12, 0.7) = \binom{12}{7} (0.7)^7 (0.3)^{12-7}$$

$$= \binom{12}{7} (0.7)^7 (0.3)^5$$

$$= 792 (0.082) (0.00243)$$

$$= 0.158 \text{ ✱}$$