



University of Chittagong
Department of computer science
&
Engineering

Assignment instead of the 2nd Tutorial

Session : [2019-2020](#)

Semester : 1st

Course Title : Basic Statistics

Course Code : STA 151

Submitted To :

Md. Shahidul Hoque
Assistant Professor
Department of Statistics
University of chittagong

Submitted By :

Shadekur Rahman Tamgid
Id-[20701064](#)
Batch : [2019-20](#)

Date of Submission : December 5, 2020

'Ans to the ques No. 1 - a.'

01.

Solution:

Frequency distribution table for calculation -

BCL(mg/dl)	Mid points x	No. of patients f	fx	C.f
149.5 - 169.5	159.5	4	638	4
169.5 - 189.5	179.5	11	1974.5	15
189.5 - 209.5	199.5	15	2992.5	30
209.5 - 229.5	219.5	25	5487.5	55
229.5 - 249.5	239.5	13	3113.5	68
249.5 - 269.5	259.5	7	1816.5	75
269.5 - 289.5	279.5	3	838.5	78
289.5 - 309.5	299.5	2	599	80
		$N = 80$	$\Sigma fx = 17460$	

Now,

$$\text{Mean, } \bar{x} = \frac{\Sigma fx}{N}$$

$$= \frac{17460}{80}$$

$$= 218.25$$

Comment: the average blood cholesterol level is 218.25 mg/dl

$$\begin{aligned}\text{Median} &= \text{Size of } \frac{N}{2} \text{th observation} \\ &= \frac{80}{2} = 40\text{th observation,}\end{aligned}$$

Hence, median lies in the class 209.5 - 229.5

$$\begin{aligned}\therefore \text{Median} &= L + \frac{N/2 - p.c.f}{f} \times c \\ &= 209.5 + \frac{40 - 30}{25} \times 20 \\ &= 209.5 + 8 \\ &= 217.5\end{aligned}$$

Comment: the median blood cholesterol level of the patients is 217.5 mg/dl,

mode:

Since, the maximum frequency is 25, which is in the class 209.5 - 229.5. therefore, 209.5 - 229.5 is the modal class.

$$\begin{aligned}\text{Now, mode} &= L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times c, \text{ where } \Delta_1 = 25 - 15 = 10 \\ &\quad \Delta_2 = 25 - 13 = 12 \\ &= 209.5 + \frac{10}{10 + 12} \times 20\end{aligned}$$

$$= 209.5 + 9.091$$

$$= 218.59$$

Comment: Hence, modal blood cholesterol level is 218.59mg/dl

1st Quartiles:

$$Q_1 = \text{Size of } \frac{N}{4} \text{th observation}$$

$$= \frac{80}{4} = 20 \text{th observation}$$

Hence, Q_1 lies in the class 189.5 - 209.5

$$\therefore Q_1 = L + \frac{N/4 - p.c.b}{h} \times c$$

$$= 189.5 + \frac{20 - 15}{15} \times 20$$

$$= 189.5 + 6.67$$

$$= 196.17$$

comments: 25 per cent of the patients blood cholesterol level is 196.17 mg/dl or less.

3rd Quartiles:

09.

$$Q_3 = \frac{3N}{4} \text{th observation,}$$
$$= 60 \text{th observation,}$$

Hence, Q_3 lies in the class 229.5 - 249.5

$$\therefore Q_3 = L + \frac{\frac{3N}{4} - p.c.b}{b} \times c$$
$$= 229.5 + \frac{60 - 55}{13} \times 20$$
$$= 229.5 + 7.69$$
$$= 237.19$$

Comments: 75 per cent of the patients blood cholesterol level is 237.19 mg/dl or less.

4th Deciles:

$$D_4 = \text{Size of } \frac{4N}{10} \text{th observation,}$$
$$= 32 \text{th observation,}$$

Hence, D_4 lies in the class 209.5 - 229.5

$$\therefore D_4 = L + \frac{\frac{4N}{10} - p.c.b}{b} \times c$$

$$= 209.5 + \frac{32 - 30}{25} \times 20$$

$$= 209.5 + 1.6$$

$$= 211.1$$

comment: 40 percent of the patients blood cholesterol level is 211.1 mg/dl or less.

65th percentiles:

$$P_{65} = \text{size of } \frac{65N}{100} \text{ th observation}$$

$$= 52\text{th observation,}$$

Hence,

P_{65} lies in the class 209.5 - 229.5

$$\therefore P_{65} = L + \frac{\frac{65N}{100} - p.c.f}{h} \times c$$

$$= 209.5 + \frac{52 - 30}{25} \times 20$$

$$= 209.5 + 17.6$$

$$= 227.1$$

Comment: 65 percent of the patients blood cholesterol level is 227.1 mg/dl or less

'Ans to the ques No.1-b'

06.

Solution: frequency distribution table for calculation -

BCL(mg/dl)	Mid point X	No of patient f	$X - 219.5/20$ d	fd	fd ²
149.5-169.5	159.5	4	-3	-12	36
169.5-189.5	179.5	11	-2	-22	44
189.5-209.5	199.5	15	-1	-15	15
209.5-229.5	219.5	25	0	0	0
229.5-249.5	239.5	13	+1	13	13
249.5-269.5	259.5	7	+2	14	28
269.5-289.5	279.5	3	+3	9	27
289.5-309.5	299.5	2	+4	8	32
N = 80			$\Sigma fd = -5$ $\Sigma fd^2 = 195$		

Standard deviation:

$$\begin{aligned}
 \sigma &= \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times c \\
 &= \sqrt{\frac{195}{80} - \left(\frac{-5}{80}\right)^2} \times c \\
 &= \sqrt{2.4375 - 0.00391} \times 20 \\
 &= 31.2 \text{ Ans.}
 \end{aligned}$$

Variance:

we know that,

$$\text{Variance} = (\text{Standard deviation})^2$$

$$= (\sigma)^2$$

$$= (31.2)^2$$

$$= 973.44$$

\therefore variance is 973.44

Coefficient of variation:

we know that,

$$C.V = \frac{\sigma}{\bar{X}} \times 100$$

Now,

$$\bar{X} = A + \frac{\sum fd}{N} \times c$$

$$= 219.5 + \frac{-5}{80} \times 20$$

$$= 219.5 + (-1.25)$$

$$= 219.5 - 1.25$$

$$= 218.25$$

$$\therefore C.V = \frac{31.2}{218.25} \times 100$$

$$= 14.296$$

\therefore coefficient of variation is 14.296 %.

'Ans to the ques No. 1-c'

Solution: From the solution of 1-a,

we get,

$$\text{median} = 217.5$$

$$\text{mode} = 218.59$$

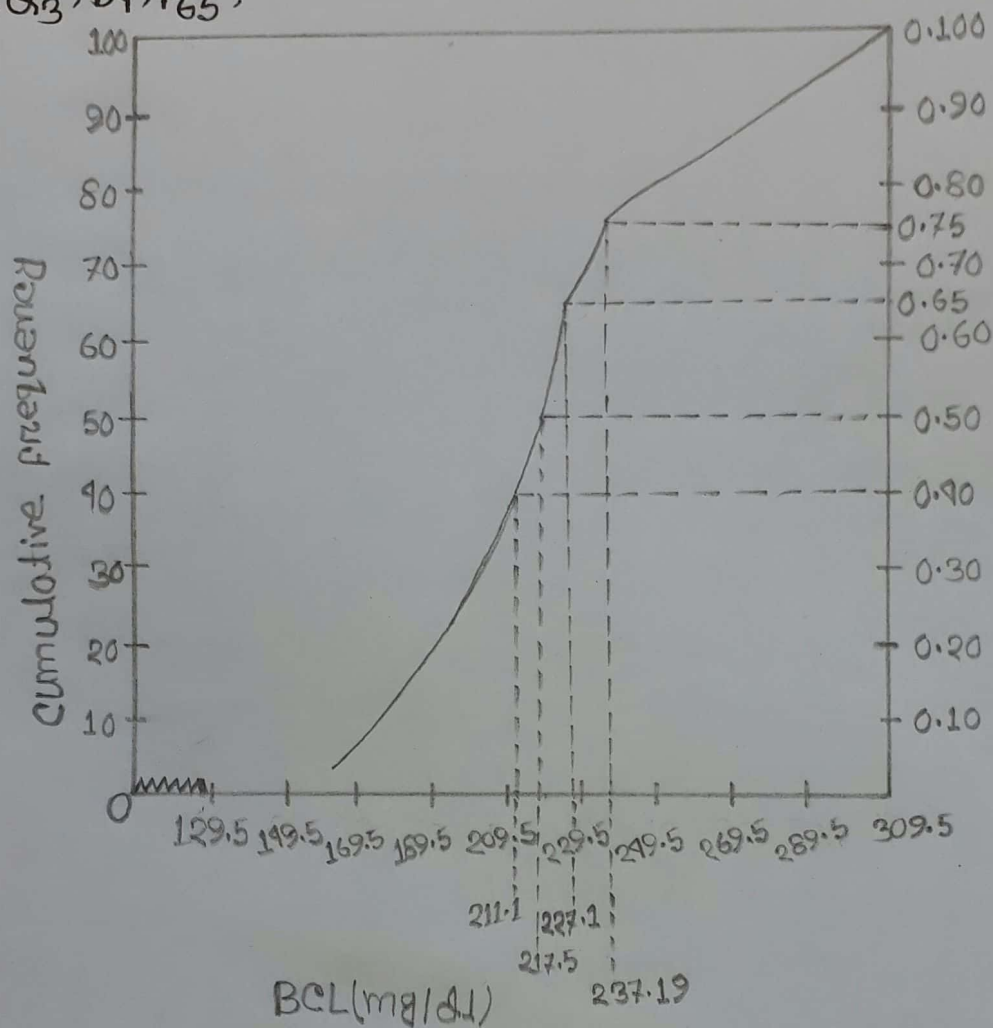
$$Q_3 = 237.19$$

$$D_4 = 211.1$$

$$P_{65} = 227.1$$

Now, using the frequency distribution table of 1-a,

Let's draw the ogive by the 'less than' method - for median, Q_3 , D_4 , P_{65} .



To determine different quartiles, horizontal lines (broken) are drawn from the cumulative frequency values. For example, if we want to determine the value of median, a horizontal line can be drawn from the cumulative frequency value of 0.50 to the less than curve and then extending a vertical line to the

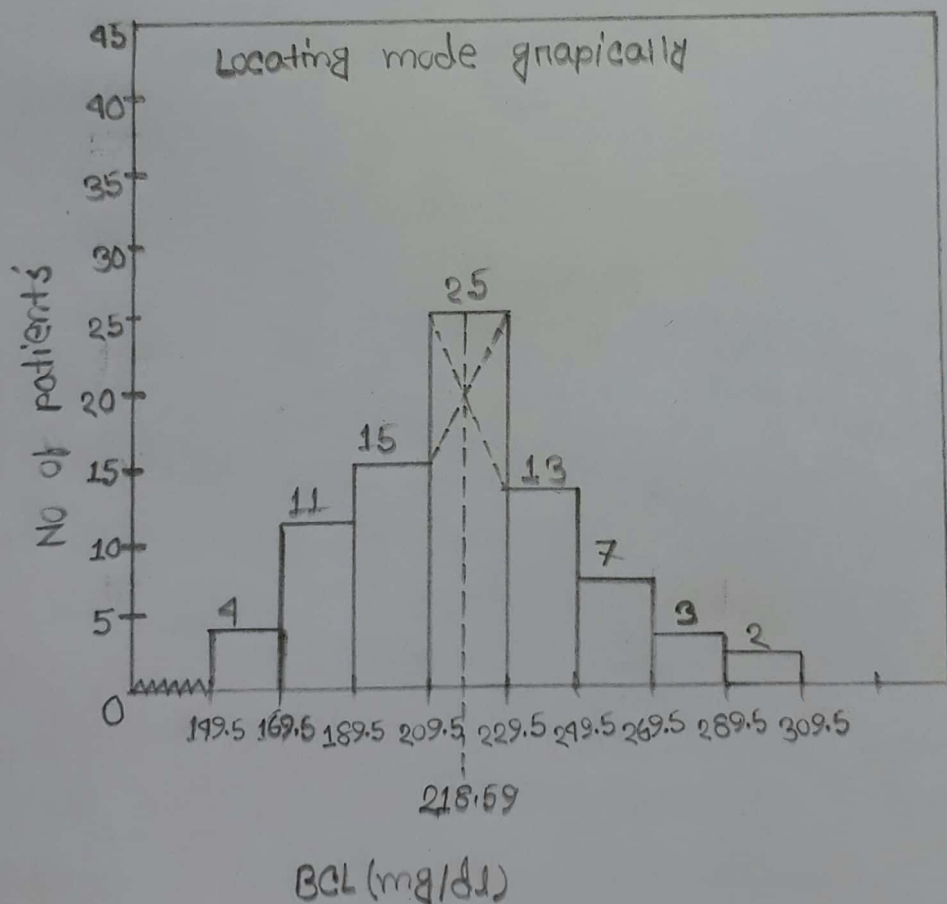
horizontal axis. In a similar manner other values can be determined as shown in the graph. 10.

Therefore, $Q_2 = 217.5$, $Q_3 = 237.19$, $D_4 = 211.1$, and $P_{65} = 227.1$.

This may be noted down here that these graphical values are same as obtained by the formulae.

Now,

Lets draw the histogram for finding the modal value.



Solution: Table for calculation.

product	cost price X	Selling price Y	XY	X ²	Y ²
A	10	11.5	115	100	132.25
B	15	18.0	270	225	324
C	14	18.5	259	196	342.25
D	20	20.9	418	400	436.81
E	31	33.2	1029.2	961	1102.24
F	31	39.0	1326	1156	1521
G	57	64.2	3659.4	3249	4121.64
H	65	74	4810	4225	5476
n=8	$\Sigma X = 246$	$\Sigma Y = 279.3$	$\Sigma XY = 11886.6$	$\Sigma X^2 = 10512$	$\Sigma Y^2 = 13456.19$

Now, Let cost price be denoted by x , which is independent variable. And selling price be denoted by y , which is dependent variable.

So, Regression line of y on x be,

$$y = a + bx$$

Here,

$$\begin{aligned}
 b &= \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} \\
 &= \frac{11886.6 - \frac{246 \times 279.3}{8}}{10512 - \frac{(246)^2}{8}} \\
 &= \frac{11886.6 - 8588.475}{10512 - 7564.5} \\
 &= \frac{3298.125}{2947.5} \\
 &= 1.119
 \end{aligned}$$

and,

$$\hat{a} = \bar{y} - b\bar{x}$$

where,

$$\bar{y} = \frac{\sum y}{n} = \frac{279.3}{8} = 34.9$$

$$\text{and, } \bar{x} = \frac{\sum x}{n} = \frac{246}{8} = 30.75$$

$$\begin{aligned}
 \text{So, } \hat{a} &= 34.9 - 1.119 \times 30.75 \\
 &= 0.491
 \end{aligned}$$

\therefore Regression line of y on x ,
 $y = 0.491 + 1.119x$

Comment: This is a simple or linear regression

Solution: from the solution of a,
we get,

the regression line of y on x ,

$$y = 0.491 + 1.119x$$

Now,

If, the cost price for a product of company
is take 50 (hundred), then the selling price is -

$$\begin{aligned} y &= 0.491 + 1.119 \times 50 \\ &= 56.441 \text{ (hundred)} \end{aligned}$$

So,

the profit incurred by the company -

$$\begin{aligned} \text{profit} &= \text{selling price} - \text{cost price} \\ &= 56.441 - 50 \text{ (hundred)} \\ &= 6.441 \text{ (hundred)} \end{aligned}$$

\therefore the profit incurred by the company is 6.441 (hundred)

Solution: using the table of 2-a,

first of all we have to find regression coefficient.

Now,

Let the cost price be denoted by x . and selling price be denoted by y .

So, regression coefficient of y on x -

$$b_{yx} = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

$$= \frac{11886.6 - 8588.475}{10512 - 7564.5}$$

$$= 1.119$$

then,

regression coefficient of x on y -

$$b_{xy} = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum y^2 - \frac{(\sum y)^2}{n}}$$

$$= \frac{11886.6 - 8588.475}{13456.19 - 9751.06}$$

$$= \frac{3298.125}{3705.13}$$

1.5

$$= 0.89$$

Now,

correlation coefficient,

$$r = \sqrt{b_{xy} \times b_{yx}}$$

$$= \sqrt{0.89 \times 1.119}$$

$$= \sqrt{0.9959}$$

$$= 0.998$$

\therefore correlation coefficient is 0.998

comment: There is a strong positive correlation between cost price and selling price of the company