ENGINEERING STATISTIC

3rd LECTURE: NUMERICAL DESCRIPTIVE MEASURES

The numerical descriptive measures are :

1- Measures of central tendency (mean, median, mode, mid range)

2- Measures of dispersion (range, variance, standard deviation)

3- Measures of position (percentiles, deciles, quartiles) .

***a- Mean*:** is the most frequently used measure of central tendency, it is denoted by x- for sample data and μ for population.

** For un grouped data, the mean is obtained by :

$$Mean\ for\ population\ data = \frac{Sum\ of\ all\ valuse}{Number\ of\ values} \rightarrow \mu = \frac{\sum x}{N}$$

$$Mean\ for\ sample\ data = \frac{Sum\ of\ all\ valuse}{Number\ of\ values} \rightarrow x^{-} = \frac{\sum x}{n}$$

Exp (1): Following are the ages of all eight employees of small company: 53 32 61 27 39 44 49 57

1- find the mean age of these employees.

2- Calculate the mean age for a sample of three employees(ages are : 32, 39, 57) from this company.

Solution :

1. $\mu = \frac{\sum x}{N} \rightarrow \frac{362}{8} = 45.25\ years$

2. $x^{-} = \frac{\sum x}{n} \rightarrow \frac{32+39+57}{3} = 42.67\ years$

*** For grouped data, the mean is obtained by :*

$$Mean\ for\ population\ data: \quad \mu = \frac{\sum f\ .x_m}{N}$$

$$Mean\ for\ sample\ data: \quad x^{-} = \frac{\sum f.x_{m\cdot}}{n}$$

Where x_m is the midpoint and f is the frequency of a class.

Exp.(2): Calculate the mean for the frequency distribution table below:

i	Class limit	Freq. (f_i)
1	10-12	4
2	13-15	12
3	16-18	20
4	19-21	14

Solution :

i	Class limit	Freq. (f_i)	x_m	$X_m.f$
1	10-12	4	11	44
2	13-15	12	14	168
3	16-18	20	17	340
4	19-21	14	20	280
		n= 50		∑= 832

$$x^- = \frac{\sum f.x_m}{n} \quad , \quad x^- = \frac{832}{50} = 16.64$$

***b- Median*:** is another important measure of central tendency, it is the value of the middle term in a data set that has been ranked in increasing

- Rank the data set in increasing order

- Find the position of the middle term which obtained as : (n+1)/2

** For un grouped data, the median is obtained by :

$$\text{Median} = \text{value of the } \left(\frac{n+1}{2}\right)\text{th term in a ranked data set}$$

** If the data set represents a population, replace n with N.

Exp (3): Find the median for the Following data: 10 5 19 8 3

Solution :

1- range data : 3 5 8 10 19

2- The position of the middle term : (n+1)/2 → (5+1) /2= 3 , the median is 8.

Exp (4): Find the median for the Following data: 11 15 23 9 14 17

Solution : 9 11 14 15 17 23

(n+1)/2= (6+1)/2= 3.5 , the position of the median is between third and forth value.

Median is (15+14)/2 = 14.5

*** For un grouped data, the median is obtained by :*

1- Construct the cumulative frequency distribution.
2- find the class that contain the median. Class Median is the first class with the value of cumulative frequency equal at least n/2.
3- Find the median by using the following formula:

$$\text{Median (M)} = L_m + \frac{(n/2) - F}{f_m} \times \Delta$$

L_m :The lower class boundary of the median class.
n: Total number of data.
F: The cumulative frequency of the class before the median class.
f_m: The frequency of the median class. , Δ: the class width.

Exp.(5): Calculate the median for the frequency distribution of the following data which represented the average rain precipitation depth in 20 gauge station during one year. Data are record to the nearest mm.

i	Class limit	Freq. (f_i)
1	10-14	2
2	15-19	3
3	20-24	4
4	25-29	5
5	30-34	3
6	35-39	2
7	40-44	1

Solution : $n = \sum fi = 20$

$n/2 = 20/2 = 10$

median class (24.5-29.5)

i	Class limit	Class boundaries	Freq. (fi)	Cumulative Frequency
1	10-14	9.5-14.5	2	2
2	15-19	14.5- 19.5	3	5
3	20-24	19.5- 24.5	4	9
4	25-29	24.5- 29.5	5	14
5	30-34	29.5- 34.5	3	17
6	35-39	34.5- 39.5	2	19
7	40-44	39.5- 44.5	1	20

$$\text{Median (M)} = L_m + \frac{(n/2) - F}{f_m} \times \Delta$$

$$M = 24.5 + \frac{10 - 9}{5} \times 5 = 25.5$$

b- Mode: is the value that occurs with highest frequency in a data set.

Exp: Find the mode for these data sets.

a- 1, 1, 2, 3, 4, 4, 4, 5, 6, 6. Mode =4 (unimodal)

b- 2, 3, 6, 8, 9 No mode.

c- 1, 2, 3, 3, 3, 4, 4, 4, 5, 6. Mode = 3,4 (bimodal)

*** For un grouped data, the mode is obtained by :*

$$\text{Mode} = L + \frac{f_m - f_{m-1}}{(f_m - f_{m-1}) + (f_m - f_{m+1})} \times \Delta$$

L : The lower boundary of the modal class

f_{m-1} : The frequency of the class before the modal class

f_m : The frequency of the modal class

f_{m+1} :The frequency of the class after the modal class.

Δ: class width

Exp.(6): Find the mode for data in table.

i	Class limit	Freq. (f_i)
1	10-14	2
2	15-19	3
3	20-24	4
4	25-29	5
5	30-34	3
6	35-39	2
7	40-44	1

Solution : The mode class is (25-29)

$$\text{Mode} = L + \frac{f_m - f_{m-1}}{(f_m - f_{m-1}) + (f_m - f_{m+1})} \times \Delta$$

$$\text{Mode} = 24.5 + \frac{1}{1 + 2} \times 5 = 26.17$$

Relation among the Mean, Median, and Mode

The value of mean, median, and mode can give some idea about the shape of a frequency curve.

1- Symmetric histogram and distribution curve : The value of mean, median, and mode are identical, and they lie at the center of distribution fig.(1).

2- The histogram and distribution curve are skewed to the right : The value of mean is largest, mode are smallest, and the value of median lies between them fig.(2).

3- The histogram and distribution curve are skewed to the left: The value of mean is smallest, mode are largest, and the value of median lies between them fig.(3).

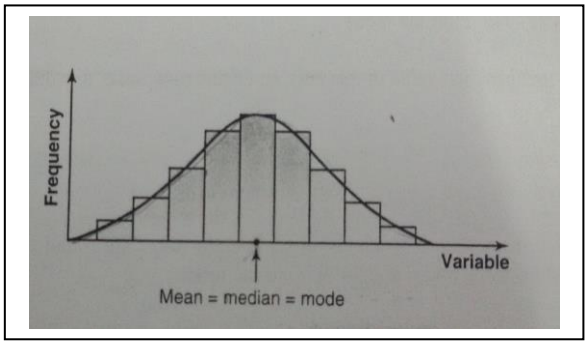


Fig. (1)

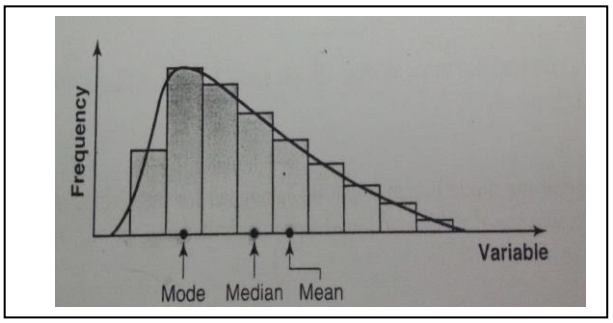


Fig. (2)

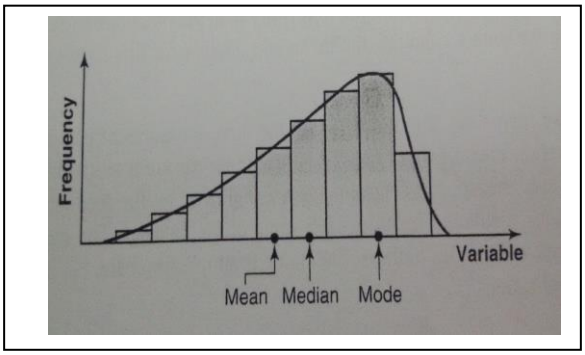


Fig. (3)

**** Find the relation among the mean, mode, and median for this frequency distribution table:

i	Class limit	Class boundaries	freq. (f_i)	x_m	x_m.fi	cumulative freq. (F)
1	10-12	9.5-12.5	4	11	44	4
2	13-15	12.5-15.5	12	14	168	16
3	16-18	15.5-18.5	20	17	340	36
4	19-21	18.5-21.5	14	20	280	50
			n= 50		∑= 832	

$$x^- = \frac{\sum f.x_m}{n} \quad , \quad x^- = \frac{832}{50} = 16.64$$

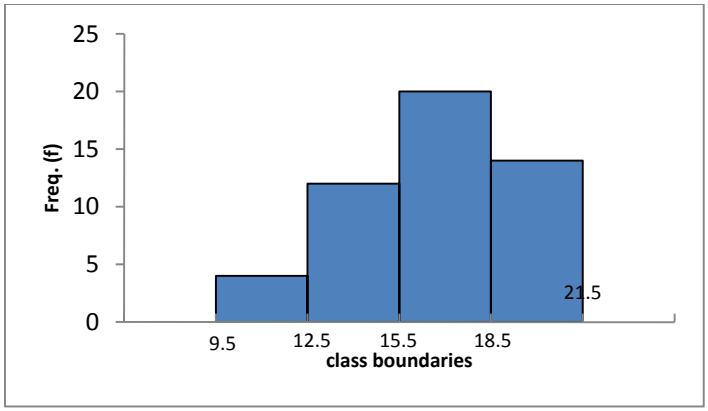
n/2= 50/2=25 , the median class is (15.5-18.5)

median(m) = L_m+ ((n/2-F)/f_m)*Δ , median = 15.5 +((25-16)/20)*3 = 16.85

the mode class is (15.5-18.5).

mode = L+ ((f_m – f_{m-1}) / (f_m-f_{m-1}) + (f_m-f_{m+1}))* Δ, mode= 15.5 + (8/(8+6))*3 = 17.21

The histogram and distribution curve are skewed to the left



d- Mid range: is the mean of the largest and the smallest values in a data set.

Exp: Find the mid range of data : 10, 4, 8, 6, 2, 12, 20

Sol:

2, 4, 6, 8, 10, 12, 20

Mid range = (smallest value + largest value)/2 = (2+20) / 2= 11