UNIT : 5

Date __/__/__ Statistical Technique (Saathi) (14m)

Measures of central technique
Measures of central Tendency
Arithmatic mean
The arithmatic mean $\bar{x}$ of n observation is given by $x_1, x_2, x_3 ---- x_n$

$$\bar{x} = \frac{x_1 + x_2 + x_3 + ------ + x_n}{n}$$

$$\bar{x} = \frac{\Sigma x_i}{n}$$

The arthmatic mean of n observation with frequencies $f_1, f_2, ------ f_n$ is given by

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + ------ + x_n f_n}{f_1 + f_2 + ------ + f_n}$$

$$\bar{x} = \frac{\Sigma x_i f_i}{\Sigma f_i}$$

NOTE : If X is a random variable then arthmatic mean E(X) which is the measure of central tendency

Median
Median of the distrubution is the value of variable which divides it into two equal parts. If the no. of observation is odd then median is the middle value of after the values have been arrange in ascending or desscending order of magnitude

If the no. of observation is even then their are two middle even value and middle is obtaine by taking arthmatic mean

Date __/__/__

of middle term.

In case of continuous freq distribution the class corresponding than N/2 is causet the median class and the value of the median is obtained by the formula.

$$\text{median} = l + \left[\frac{\left(\frac{N}{2} - c\right)}{f} \times h\right]$$

where, $l \rightarrow$ is the lower limit of median class

$f \rightarrow$ frequency of median

$h \rightarrow$ The magnitude of the median class (interval) Difference betn class interval

$c \rightarrow$ commitective frequency or class presiding previous class

$N \rightarrow \Sigma f$

Mode : In case of discreate frequency distribution mode is the value of x corresponding to maximum frequency

The reln betn mean, mode & meadian is given by

mode = 3 median − 2 mean

Date __/__/__

If the group data is given by then the mode is given by

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

where,

l = lower limit of modal class

f_1 = frequency of modal class

f_0 = freq before the modal class

f_2 = freq after the modal class

h = difference betn class interval.

Q Find the arithmatic mean of the following frequency distribution. also find mode and median.

Ans:

x	1	2	3	4	5	6	7
f	5	9	12	17	14	10	6

→

x	f	xf
1	5	5
2	9	18
3	12	36
4	17	68
5	14	70
6	10	60
7	6	42

$\Sigma xf = 299$

Arithmatic mean $(\bar{x}) = \frac{\Sigma x_i f_i}{\Sigma f_i} = \frac{299}{73}$

$= 4.09$

Page No.

Date __/__/__

mode = 4

mode = 3 median − 2 mean

4 = 3 median − 2 × 4.09

median = $\frac{4 - 2\times4.09}{3}$

median = 4.06

Q2. obtained the median for the following frequency distribution

x	1	2	3	4	5	6	7	8	9
f	8	10	11	16	20	25	15	9	6

Sol^n:-

x	f	Cummulative freq
1	8	8
2	10	8+10 = 18
3	11	18+11 = 29
4	16	29+16 = 45
5	20	45+20 = 65
6	25	65+25 = 90
7	15	90+15 = 105
8	9	105+9 = 114
9	6	114+6 = 120 → N
	$\Sigma f = 120$	

Here N → 120, ∴ $\frac{N}{2} = 60$

∴ the commulative freq just greater than median class 60 is greater than 65 and corresponding value of 65 for is 5.

∴ the median is equal to 5.

Date __/__/__

Q3 find the value of median from the following data

Class interval	0-10	10-20	20-30	30-40	40-50	50-60
freq	30	40	50	48	24	62

60-70	70-80
142	104

→

class interval	mid value (x)	freq (f)	commulative freq
0-10	5	30	30
10-20	15	40	70
20-30	25	50	120
30-40	35	48	168
40-50	45	24	192
50-60	55	f_0 = 62	254 → median class
60-70	65	142 = f_1	396 — modal class
70-80	75	f_2 = 104	500 → N

Hence, N = $\frac{500}{2}$ = 250th term

Commulative freq^n greater than 250 is 254

Hence the class interval is 50-60, so, therefore median = $l + \left[\frac{\left(\frac{N}{2} - c\right)}{f} \times h\right]$

$= 50 + \left[\frac{(250 - 192)}{62} \times 10\right]$

$= 59.35$

Mode

Page No.

Date __/__/__

class limit 60-70
diff (h)

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

$$= 60 + \left[\left(\frac{142 - 62}{2 \times (142) - 62 - 104}\right) \times 10\right]$$

$$= 66.70$$

Q find median & mode for the following distn

class interval	0-100	100-200	200-300	300-400
f	5	7	12	18

	400-500	500-600	600-700
freqn	16	10	5

Q. find mode of the following data.

Age	0-6	6-12	12-18	18-24	24-30
freqn	6	11	25	35	18

30-36	36-40
12	5

Age (x)	freq (f)
0-6	6
6-12	11
12-18	25
18-24	35
24-30	18
30-36	12
36-40	5

Page No.

Date __/__/__

$l = 18$, $f_1 = 35$, $f_0 = 25$, $f_2 = 18$, $h = 6$

$$\text{Mode} = l + \left[\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right] \times h$$

$$= 18 + \left[\frac{35 - 25}{2(35) - 25 - 18}\right] \times 6$$

$$= 20.22$$

Average deviation or Mean deviation

It is mean of absolute value ~~no.s~~ of the deviation of a given set of numbers of their arithmatic mean $\bar{x}$.

$x_1, x_2 \ldots x_n$ be the set of no with freqn $f_1, f_2 \ldots f_n$ respectively

Therefore mean deviation for the given data is.

$$\text{Mean deviation} = \frac{\sum f_i |x_i - \bar{x}|}{\sum f_i}$$

Standard deviation (σ)

Std. deviation is defined as the square root of mean of the square of the deviation arithmatic mean

$$S.D = \sigma = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}}$$

Page No.

Q Calculate mean and std deviation for the following data.

Size	6	7	8	9	10	11	12
freq	3	6	9	13	8	5	4

→ $\bar{x} = \frac{\Sigma x}{n} = \frac{63}{7} = 9$

x	f	$x-\bar{x}$	$\lvert x-\bar{x}\rvert$	$f(x-\bar{x})$	$(x-\bar{x})^2$	$f\times(x-\bar{x})^2$
6	3	-3	3	+9	9	27
7	6	-2	2	+12	4	24
8	9	-1	1	+9	1	9
9	13	0	0	0	0	0
10	8	1	1	8	1	8
11	5	2	2	10	4	20
12	4	3	3	12	9	36

$\Sigma f = 48$ $\quad \Sigma f(x-\bar{x}) = 60$ $\quad f\times(x-\bar{x})^2 = 124$

Mean deviation $= \frac{\Sigma f_i \lvert x_i - \bar{x}\rvert}{\Sigma f_i}$

$= \frac{60}{48} = 1.25$

Std. deviation $= \sigma = \sqrt{\frac{\Sigma f_i (x_i-\bar{x})^2}{\Sigma f_i}}$

$= \sqrt{\frac{124}{48}}$

$= 1.60$

~~From the following distbn compute mean & std deviation~~

Q From the following distbn compute mean & std deviation of 200 students and also find the variable.

→

Class of student	60-62	63-65	66-68	69-71	72-74
	5	18	42	27	8

→ $\bar{x} = \frac{\Sigma x}{n} = 67$

x	No. of freq	$(x-\bar{x})$	$\lvert x-\bar{x}\rvert$	$f(x-\bar{x})$	$(x-\bar{x})^2$	$f.(x-\bar{x})^2$
61	5	-6	6	30	36	180
64	18	-3	3	54	9	162
67	42	0	0	0	0	0
70	27	3	3	81	9	243
73	8	6	6	48	36	280

$\Sigma f = 100$ $\quad \Sigma f\lvert x-\bar{x}\rvert = 213$ $\quad \Sigma f(x-\bar{x})^2 = 973$

mean deviation $= \frac{\Sigma f_i \lvert x-\bar{x}\rvert}{\Sigma f_i} = \frac{213}{100} = 2.13$

Std. deviation $= \sqrt{\frac{\Sigma f_i (x_i-\bar{x})^2}{\Sigma f_i}} = \sqrt{\frac{973}{100}} = 2.95$

$Var(x) = 8.706$

UNIT :5

Date __/__/__ Statistical Technique

Saathi

* Multiple Co-relation
When more than two variable are studied then the problem of multiple correlation is

1) Partial co-reln
2) Total co-reln

1) Partial co-relation
This is used when we identify more than two variables but consider only two variable assuming other variable are constant.

Let x_1, x_2, x_3 be the variables and r_{ij} be the correln betn the variable x_i and x_j. When the partial correln coefficient betn x_1 and x_2 considering x_3 as constant then it is return as

$$r_{12.3} = \frac{r_{12} - r_{13} r_{23}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{23}^2}}$$

When x_1 and x_3 are partially correlated and considering x_2 constant

$$r_{13.2} = \frac{r_{13} - r_{12} \cdot r_{32}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{32}^2}}$$

When x_2 and x_3 are partially correlated considering x_1 constant

$$r_{23.1} = \frac{r_{23} - r_{21} \cdot r_{31}}{\sqrt{1-r_{21}^2}\sqrt{1-r_{31}^2}}$$

Date __/__/__

Saathi

Total Co-reln
It is used when we study all the existing variable.

Let r_{ij} be the co-reln coeffent betn the variables x_i and x_j then formula for multiple co-reln coefficient for two independent variable is.

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1-r_{23}^2}}$$

Q1. Given co-reln coefficient of order $r_{12} = 0.97$, which $r_{13} = 0.46$, $r_{23} = 0.53$. Calculate multiply correln coefficient by taking first variable as dependent and 2nd and 3rd variable as independent.

$$R_{1.23} = \sqrt{\frac{(0.97)^2 + (0.46)^2 - 2(0.97)(0.46)(0.53)}{1-(0.53)^2}}$$

$$= 0.97$$

Q 2nd is dependent and 1st and 3rd variable are indepent in multiple co-reln.

→

$$R_{2.13} = \sqrt{\frac{r_{12}^2 + r_{23}^2 - r_{31} \cdot r_{23} r_{13}}{1-r_{13}^2}}$$

Page No.

Q Let x_1, x_2, x_3 are 3 variables Calculated from mean with $n=10$, $\Sigma x_1^2 = 90$, $\Sigma x_2^2 = 160$, $\Sigma x_3^2 = 40$, $\Sigma x_1 \cdot x_2 = 60$, $\Sigma x_1 \cdot x_3 = 40$, $\Sigma x_2 \cdot x_3 = 60$, values given. Calculate multiple corrdn coefficient corrdn.

→

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{23}^2}}$$

Calculate, r_{12}, r_{13}, r_{23} by Karlperson's corrdn coefficient.

$$r_{12} = \frac{\Sigma x_1 x_2}{\sqrt{\Sigma x_1^2 \cdot \Sigma x_2^2}}$$

$$= \frac{60}{\sqrt{90 \cdot 160}} = 0.5$$

$$r_{13} = \frac{\Sigma x_1 \cdot x_3}{\sqrt{\Sigma x_1^2 \cdot \Sigma x_3^2}} = \frac{40}{\sqrt{90 \cdot 40}}$$

$$= 0.66$$

$$r_{23} = \frac{\Sigma x_2 \cdot x_3}{\sqrt{\Sigma x_2^2 \cdot \Sigma x_3^2}} = \frac{60}{\sqrt{40 \cdot 160}} = 0.75$$

$$r_{1.23} = \sqrt{\frac{(0.5)^2 + (0.66)^2 - 2(0.5 \times 0.66 \times 0.75)}{1 - (0.75)^2}}$$

$$= 0.60$$

Q In a distribution of 3 variables it is obser that $r_{12} = 0.70$, $r_{13} = 0.61$, $r_{23} = 0.40$. Calculate the value of $r_{12.3}$, $r_{13.2}$ and $r_{23.1}$

→

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1 - r_{13}^2}\sqrt{1 - r_{23}^2}}$$

$$= \frac{0.70 - 0.61 \times 0.40}{\sqrt{1 - (0.61)^2}\sqrt{1 - (0.40)^2}}$$

$$= 0.627$$

$$r_{13.2} = \frac{r_{13} - r_{12} \cdot r_{32}}{\sqrt{1 - r_{12}^2} \cdot \sqrt{1 - r_{32}^2}}$$

$$= \frac{0.61 - 0.70 \times (0.40)}{\sqrt{1 - (0.70)^2} \times \sqrt{1 - (0.40)^2}}$$

$$= 0.50$$

$$r_{23.1} = \frac{r_{23} - r_{21} \cdot r_{31}}{\sqrt{1 - r_{21}^2} \cdot \sqrt{1 - r_{31}^2}}$$

$$= \frac{0.40 - 0.70 - 0.61}{\sqrt{1 - 0.70^2} \cdot \sqrt{1 - 0.61^2}}$$

$$= 0.04$$

Date __/__/__

$r_{12.3}$

Q Let weigh $= x_1$, hight $= x_2$ and sugar level x_3 of an group of individuals are correlated so follow $r_{12} = 0.75$, $r_{13} = 0.76$ $r_{23} = 0.61$ calculate the partial corelⁿ betⁿ weight x_1 and hight x_2 this regarding the effect of sugar level x_3.

Q the ~~$r_{12.3}$~~ following zero order corelⁿ cofficient corelⁿ $r_{12} = 0.97$, $r_{13} = 0.46$, $r_{23} = 0.53$. Calculate the partial corelⁿ betⁿ first 3rd variable keeping the effect 2nd variable constant

$r_{13.2}$

* Regression

Let x & y be two variables if the scatterred diagram shows some relationship be in two variables then the dot of statered dig. will be concentrated near the curv this curve is called curve of Regression.

Multiple Regression

Multiple Regression is a Bet of technique that discribed the relation between two or more independed variable and one independent variable.

Regression eqⁿ of three variables

Let three linear relationship betⁿ the dependent variable x_3 and two independent variable

Page No.

Date __/__/__

x_1 & x_2 such that

$$x_3 = a + bx_1 + cx_2 \longrightarrow (1)$$

x_1 is dependent variable & x_2 & x_3 are independent variables.

$$x_1 = a + bx_2 + cx_3$$

x_2 is dependent variable & x_1 & x_3 are independent variable

$$x_2 = a + bx_1 + cx_3$$

Ex: Find the multiple line regression of x_1 on x_2 & x_3 from the data relation to three variables given below

x_1	4	6	7	9	13	16
x_2	15	12	8	6	4	9
x_3	20	24	20	14	10	4

Solⁿ →

Regression eqⁿ x_1 on x_2 & x_3

$$x_1 = a + bx_2 + cx_3 \longrightarrow (1)$$

Normal eqⁿ of the

1st taking Σ on both side

$$\Sigma x_1 = na + b\Sigma x_2 + c\Sigma x_3 \longrightarrow (a)$$

$$\Sigma x_1 x_2 = a\Sigma x_2 + b\Sigma x_2^2 + c\Sigma x_2 x_3 \longrightarrow (b)$$

$$\Sigma x_1 x_3 = a\Sigma x_3 + b\Sigma x_2 x_3 + c\Sigma x_3^2 \longrightarrow (c)$$

Page No.

Date ___/___/___

Here $n = 6$

x_1	x_2	x_3	$x_1 x_2$	$x_2 x_3$	x_2^2	$x_1 x_3$
4	15	20	60	300	225	80
6	12	24	72	288	144	144
7	8	20	56	160	64	140
9	6	14	54	84	36	126
13	4	10	52	40	16	130
15	3	4	45	12	9	60
54	48	92	339	884	494	680
Σx_1	Σx_2	Σx_3	$\Sigma x_1 x_3$	$\Sigma x_2 x_3$	Σx_2^2	$\Sigma x_1 x_3$

Substitute all above value in eqn A, b, c

(a) = $54 = 6a + 48b + 92c$

(b) = $339 = 48a + 494b + 884b$

(c) = $680 = 92a + 884b + 1688c$

$a = 17.100$, $b = -0.4520$, $c = -0.2925$

so therefore x_1 on x_2 and x_3 is

(1) $\Rightarrow x_1 = 17.100 - 0.4520 x_3 - 0.2925 x_3$