"BIOSTATISTICS"

24.1-17

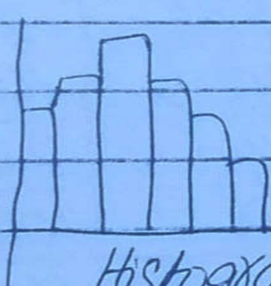
It

Population ⇒ "Total area under discussion"

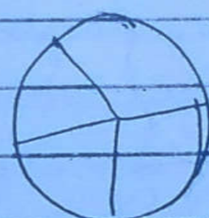
* Populations sy value → ~~Statistics~~ Parameter
* Samples sy value → ~~Parameter~~ Statistics

Qualitative data.

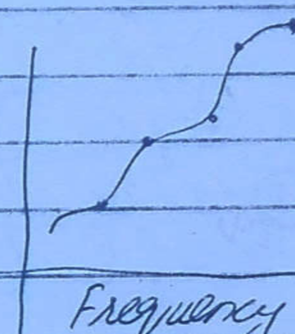
The data which can be measure in the form of quality.



Histogram



Pi- chart



Frequency polygon.

Measures of dispersion.

Spread of values from origin.

Dispersion of constant value is always '0' zero.

(i) Absolute dispersion → same unit answer
(ii) Relative dispersion → unit less answer
→ also known as coefficient of variation.

⇒ Standard deviation is the most important variation of dispersion.

Lecture #3

31-1-17

$$C.V = \frac{S.d}{\bar{x}} \times 100$$

Range = largest value − Smallest value.

Range → Difference b/w two extreme values ↳ low, highest

Group Data values.

mid values (x)	Classes	f	fx	$fx^2 = fx \times x$
smallest value 2.5	0 − 5	3	7.5	18.75
7.5	5 − 10	5	37.5	281.25
12.5	10 − 15	9	112.5	1406.25
17.5	15 − 20	8	140	2450
largest 22.5	20 − 25	2	45	1012.5
		$\Sigma f = 27$	$\Sigma fx = 342.5$	$\Sigma fx^2 = 5168.75$

Range = 22.5 − 2.5

① (a) variance. $S^2 = \frac{\Sigma f x^2}{\Sigma f} - \left(\frac{\Sigma fx}{\Sigma f}\right)^2$ ✓

① (b) $S^2 = \frac{\Sigma f (x - \bar{x})^2}{\Sigma f}$

② Standard Deviation. $S.D = \sqrt{\frac{\Sigma fx^2}{\Sigma f} - \left(\frac{\Sigma fx}{\Sigma f}\right)^2}$

③ $C.V = \frac{S.d}{\bar{x}} \times 100$

$$\boxed{\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{342.5}{27} = 12.685}$$

① $$S^2 = \frac{\Sigma fx^2}{\Sigma f} - \left(\frac{\Sigma fx}{\Sigma f}\right)^2$$

$$S^2 = \frac{5168.75}{27} - \left(\frac{342.5}{27}\right)^2$$

$$S^2 = 191.43 - (12.68)^2$$

$$S^2 = 191.43 - 160.84$$

$$\boxed{S^2 = 30.59 = 30.6}$$

② $$S.D = \sqrt{\frac{\Sigma fx^2}{\Sigma f} - \left(\frac{\Sigma fx}{\Sigma f}\right)^2}$$

$$S.D = \sqrt{30.6}$$

$$\boxed{S.D = 5.53}$$

③ $$C.V = \frac{S.d}{\bar{x}} \times 100$$

$$C.V = \frac{5.53}{12.68} \times 100$$

$$\boxed{C.V = 43.61}$$

(i) **Mean** $= \frac{\Sigma fx}{\Sigma f}$ $\quad \boxed{\bar{x}} = \frac{\Sigma fx}{\Sigma f}$

(ii) **Median** $L + \frac{(n/2) - B}{G} \times W$

(iii) **Mode** $= L + \frac{f_m - f_{m-1}}{(f_m - f_{m-1}) + (f_m - f_{m+1})} \times w$

Mean deviation (Mean, Median)

Mean Deviation from mean

$$M.D = \frac{\Sigma f\,|x-\bar{x}|}{\Sigma f} \quad \text{(Grouped)}$$

$$M.D = \frac{\Sigma\,|x-\bar{x}|}{n} \quad \text{(Ungroup)}$$

Mean deviation from Median

$$M.D = \frac{\Sigma f\,|x-\tilde{x}|}{\Sigma f} \quad \text{(Group)}$$

$$M.D = \frac{\Sigma\,|x-\tilde{x}|}{n} \quad \text{(Ungroup)}$$

Group Data

Classes	f	x	fx	$x-\bar{x}$	$f\|x-\bar{x}\|$
0-5	2	2.5	5	12.7	25.4
5-10	3	7.5	22.5	7.7	23.1
10-15	5	12.5	62.5	2.7	13.5
15-20	9	17.5	157.5	2.3	20.9
20-25	4	22.5	90	7.3	29.2
25-30	1	27.5	27.5	12.3	12.3
	$\Sigma f = 24 = n$		$\Sigma fx = 365$		$\Sigma f\|x-\bar{x}\| = 124.2$

$$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{365}{24}$$

$$\boxed{\bar{x} = 15.20}$$

$$|x-\bar{x}| = |x-15.20|$$

$x \quad \bar{x} \quad 5$

these bars change the -ve sign into +ve

$|x-\bar{x}| = |2.5 - 15.20| = |-12.7| = 12.7$

$|x-\bar{x}| = |7.5 - 15.20| = 7.7$

$|\tilde{x}-\bar{x}| = |12.5 - 15.20| = 2.7$

$|x-\bar{x}| = |17.5 - 15.20| = 2.3$

$|x-\bar{x}| = |22.5 - 15.20| = 7.3$

$|x-\bar{x}| = |27.5 - 15.20| = 12.3$

$|x-\tilde{x}| = |2.5 - 31.1|$
$= |7.5 - 31.1|$
$= |12.5 - 31.1|$
$= |17.5 - 31.1|$
$= |22.5 - 31.1|$
$= |27.5 - 31.1|$

mean deviation from mean

$$M.D = \frac{\Sigma f\,|x-\bar{x}|}{\Sigma f}$$

$$M.D = \frac{124.2}{24}$$

$$\boxed{M.D = 5.17}$$

$\tilde{x} = 31.1$

Classes	x	f	C.f	$x-\tilde{x}$	$f\|x-\tilde{x}\|$
0 - 5	2.5	2	2	28.6	57.2
5 - 10	7.5	3	5	23.6	70.8
10 - 15	12.5	5	$\boxed{10}$ C	18.6	93
15 - 20	17.5	9	19	13.6	122.4
20 - 25	22.5	4	23	8.6	34.4
25 - 30	27.5	1	24	3.6	3.6
		$\Sigma f = 24$			

lower class, upper class

model class

$\Sigma f = n$

$\frac{n}{2} = \frac{24}{2} = 12$

$\boxed{\Sigma f|x-\tilde{x}| = 381.4}$

$$\tilde{x} = l + \frac{h}{f}\left[\frac{n}{2} - C\right]$$

C = is the previous camulative frequency of the model class

l = lower class value of model class

f = frequency of model class

h = hyphon is the difference b/w two or upper classes boundries.

$$\tilde{x} = l + \frac{h}{f}\left[\frac{n}{2} - C\right]$$

$$\tilde{x} = 15 + \frac{5}{9}\left[\frac{24}{2} - 10\right]$$

$$\tilde{x} = 15 + 0.55\left[12 - 10\right]$$

$$\tilde{x} = 15 + 0.55\,(+2)$$

$$\boxed{\tilde{x} = 31.1}$$

tion
m
de

$$M.D = \frac{\sum f\,|x - \bar{x}|}{\sum f}$$

$$M.D = \frac{381.4}{24}$$

$$\boxed{M.D = 15.89}$$

Model Class ⇒ The class which have hig[h] frequency.

Lecture # 4

1-2-017

Skewness:
Kurtosis

Mean; median and mode

Moments:

(i) About mean :-
(ii) About any value :-
(iii) About zero :-

Person. $S_k = \dfrac{\text{Mean} - \text{Mode}}{\text{S.d}}$

Bowelley's: $S_k = \dfrac{Q_3 + Q_1 - 2\tilde{x}}{Q_3 - Q_1}$

Q_1 Q_3
First and third quartile.

$$Q_1 = l + \frac{h}{f}\left[\frac{n}{4} - c\right]$$

$$Q_3 = l + \frac{h}{f}\left[\frac{3n}{4} - C\right]$$

$$\tilde{x} = l + \frac{h}{f}\left[\frac{n}{2} - C\right]$$

$$Q_2 = l + \frac{h}{f}\left[\frac{2n}{4} - C\right]$$

Quantiles.
(i) median
(ii) Quartiles
(iii) Deciles
(iv) Percentiles

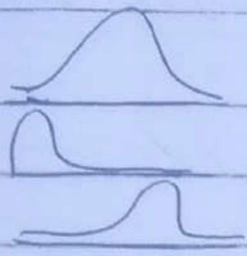
Quantiles.
value which divide an array data into some equal parts.

Median:
median is the ~~data~~ value which divide an array data into 2 equal parts.

Quartiles.
which divide array data into 4 equal parts

Deciles.
which divide - - into 10 equal parts

Percentile. - 100 equal parts



$S_K = \overset{\text{value}}{0}$ → Symmetrical distribution

$S_K = +ve$ — +vely Skewed distribution

$S_K = -ve$ — -vely skewed distribution

compares to zero ← $\beta_1 = \dfrac{(m_3)^2}{(m_2)^3}$, $\beta_2 = \dfrac{m_4}{(m_2)^2}$ → compares to 3

m_1, m_2, m_3, m_4 → are moments of mean.

Moments About Mean:

$$m_r = \frac{\Sigma(x-\bar{x})^r}{n} \quad , \quad m_r = \frac{\Sigma f(x-\bar{x})^r}{\Sigma f}$$

$$r = 1, 2, 3, 4, \ldots$$

Moments about any value

$$m'_r = \frac{\Sigma(x-A)^2}{n} \quad , \quad m'_r = \frac{\Sigma f(x-A)^2}{\Sigma f}$$

Moments about zero

$$m'_r = \frac{\Sigma(x-0)^2}{n} = \frac{\Sigma x^2}{n} \quad , \quad m'_r = \frac{\Sigma f(x-0)^2}{\Sigma f} = \frac{\Sigma f x^2}{\Sigma f}$$

⇒ First moment about mean is always zero.

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

$$\bar{X} = \frac{34}{5} = \boxed{6.8}$$

171 29
562.018

+10.4
-10.4

9

Ungroup Data: → where no frequency present.

x	$(x-\bar{x})$	$(x-\bar{x})^2$	$(x-\bar{x})^3$	$(x-\bar{x})^4$
2	-4.8	$(-4.8)^2$ ~~4.8~~ = 23.04	-110.59	530.84
3	-3.8	$(-3.8)^2$ = 14.44	-54.87	208.51
5	-1.8	$(-1.8)^2$ = 3.24	-5.83	10.49
9	2.2	$(2.2)^2$ = 4.84	10.65	23.42
15	8.2	$(8.2)^2$ = 67.24	551.368	4521.21
$\Sigma x = 34$	= 0	= 112.8	= 390.728	= 5294.4

$$m_r = \frac{\Sigma (x-\bar{x})^r}{n}$$

$$m'_1 = \frac{\Sigma (x-\bar{x})^1}{n}$$

$$\boxed{m_1 = \frac{0}{5} = 0}$$

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

$$m_2 = \frac{112.8}{5}$$

$$\boxed{m_2 = 22.56}$$

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

$$m_4 = \frac{5294.4}{5}$$

$$\boxed{m_4 = 1058.84}$$

$$m_3 = \frac{\Sigma (x-\bar{x})^3}{n}$$

$$m_3 = \frac{(390.72)}{5}$$

$$\boxed{m_3 = 78.14}$$

$$\beta_1 = \frac{(m_3)^2}{(m_2)^3}$$

$$\beta_1 = \frac{(78.14)^2}{(22.56)^3}$$

$$\boxed{\beta_1 = \frac{6105.8}{11481.9} = 0.53}$$

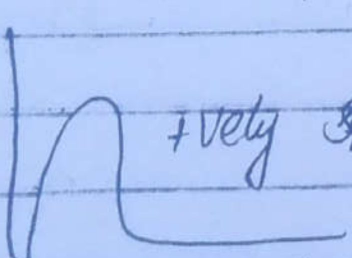
$$\beta_2 = \frac{m_4}{(m_2)^2}$$

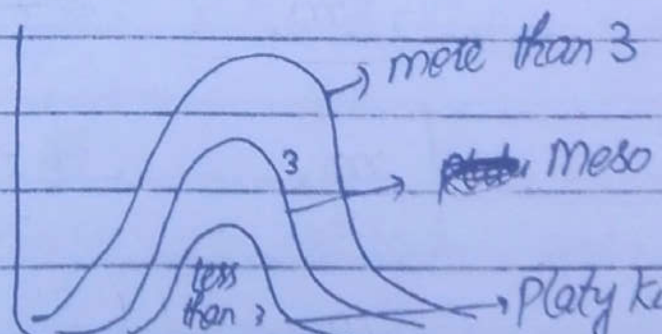
$$\beta_2 = \frac{1058.84}{(22.56)^2}$$

$$\boxed{\beta_2 = 2.08}$$

$\beta_1 = 0.53 \rightarrow \beta_1 \neq 0$

$\beta_2 = 2.$[redacted]$8 \rightarrow \beta_2 ? = 3$





$\beta_1 \nearrow$ =

Lecture # 5

6-2-17

Formulas by which moment about any value and zero are converted into moments about mean.

$$m_1 = m_1' - m_1'$$
$$m_2 = m_2' - (m_1')^2$$
$$m_3 = m_3' - 3m_2'm_1' + 2(m_1')^3$$
$$m_4 = m_4' - 4m_3'm_1' + 6m_2'(m_1')^2 - 3(m_1')^4$$

Question.

$$m_1' = 1.3$$
$$m_2' = 2.6$$
$$m_3' = 3.1$$
$$m_4' = 4.2$$

Calculate the β_1 and β_2. Also explain results

$$m_1 = m_1' - m_1'$$
$$m_1 = 1.3 - 1.3$$
$$\boxed{m_1 = 0}$$

$$m_2 = m_2' - (m_1')^2$$
$$m_2 = 2.6 - (1.3)^2$$
$$m_2 = 2.6 - 1.69$$
$$\boxed{m_2 = 0.91}$$

$$m_3 = m_3' - 3m_2'm_1' + 2(m_1')^3$$
$$= 3.1 - 3(2.6)(1.3) + 2(1.3)^3$$
$$= 3.1 - 10.14 + 4.394$$
$$\boxed{m_3 = -2.646}$$

$$m_4 = m_4' - 4m_3'm_1' + 6m_2'(m_1')^2 - 3(m_1')^4$$

$$= 4.2 - 4(3.1)(1.3) + 6(2.6)(1.3)^2 - 3(1.3)^4$$

$$= 4.2 - 16.12 + 26.364 - 8.568$$

$$= 30.364 - 24.688$$

$$\boxed{m_4 = 5.876}$$

$$\beta_1 = \frac{(m_3)^2}{(m_2)^3}$$

$$\beta_1 = \frac{(-2.646)^2}{(0.91)^3}$$

$$\beta_1 = \frac{7.00}{0.75}$$

$$\boxed{\beta_1 = 9.33}$$

$$\beta_2 = \frac{m_4}{(m_2)^2}$$

$$\beta_2 = \frac{5.876}{(0.91)^2}$$

$$\beta_2 = \frac{5.876}{0.8281}$$

$$\boxed{\beta_2 = 7.09}$$

Q: Calculate β_1, β_2 and C.V.

Classes	f	x	$(x-\bar{x})$	$(x-\bar{x})^2$	$(x-\bar{x})^3$	$(x-\bar{x})^4$	fx
0-2	2	1	-4.75	22.56	-107.17	509.06	2
2-4	5	3	-2.75	7.56	-20.796	57.19	15
4-6	9	5	0.75	0.56	0.42	0.316	45
6-8	12	7	1.25	1.56	1.953	2.441	84
8-10	3	9	3.25	10.56	34.32	111.56	27
10-12	1	11	5.25	27.56	144.70	759.69	11
	$\Sigma f = 32$	$\Sigma x = 36$					$\Sigma fx = 184$

$$C.V = \frac{S.d}{\bar{X}} \times 100 \qquad \text{Ans} = \%$$

$$\Sigma f = \cancel{x}\ n$$

$$\bar{X} = \frac{\Sigma fx}{\Sigma f} = \frac{184}{32}$$

$$\boxed{\bar{X} = 5.75}$$

$fx^2 =$	$fx \times x$
	$2 \times 1 = 2$
	$15 \times 3 = 45$
	$45 \times 5 = 225$
	$84 \times 7 = 588$
	$27 \times 9 = 243$
	$11 \times 11 = 121$
	$\Sigma fx^2 = 1224$

$$S.D = \sqrt{\frac{\Sigma fx^2}{\Sigma f} - \left(\frac{\Sigma fx}{\Sigma f}\right)^2}$$

$$S.D = \sqrt{\frac{1224}{32} - \left(\frac{184}{32}\right)^2}$$

$$S.D = \sqrt{38.25 - (5.75)^2}$$

$$= \sqrt{38.25 - 33.06}$$

$$S.D = \sqrt{5.1875}$$

$$\boxed{S.D = 2.277}$$

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

$$C.V = \frac{2.277}{5.75} \times 100$$

$$\boxed{C.V = 39.6\ \%}$$

Lecture #6
20-2-17

Regression and Correlation / Association

Regression ⇒ Relation b/w independent and dependent variable.

Define as dependence of one variable on the independence of one or more variable.

$$\text{dep} \leftarrow \boxed{\begin{array}{l} Y = \alpha + \beta X_i + \varepsilon_i \\ i = 1, 2, 3, \ldots n \end{array}}$$

Error term $\leftarrow E(\varepsilon_i) = 0$

$$Y = a + bx \quad \text{Indep.}$$

$$\text{dep} \leftarrow \boxed{\begin{array}{l} Y = a_{yx} + b_{yx}\, x \\ X = a_{xy} + b_{xy}\, Y \end{array}}$$

a = Intersept , b = Regression cofficient

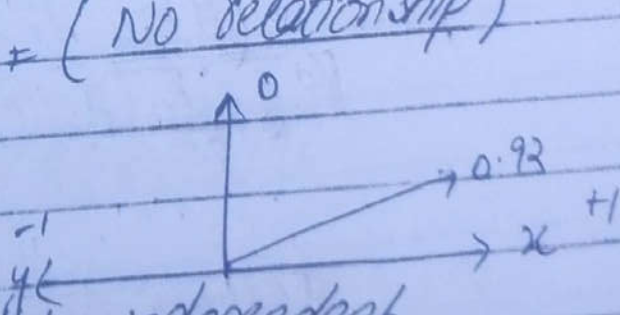
Correlation ⇒ Strong relation b/w dependent and independent variable.

Qualitative :- attributes.

Correlation = -1 to $+1$

+ve (direction same) ; 0 = (No relationship)

-ve (direction opposite)



1 independent

linear ⇒ 1 dependent and 1 independent

Multiple ⇒ Independent 1 and dependent may be many more.

Properties of Correlation:-

⇒ Correlation always lies b/w -1 to $+1$.

⇒ Co-efficient correlation is symmetrical with respect to x and y. $\gamma_{xy} = \gamma_{yx}$

⇒ Correlation coefficient is independent from change of origin and scale. $\gamma_{xy} = \gamma_{uv}$

$$u = \frac{x-A}{h}, \quad v = \frac{y-B}{K}$$

⇒ Correlation coeffient is the geometric mean of two regression coefficients

Formula Of Correlation

$\gamma_{xy} = \gamma_{yx}$

$$\gamma_{xy} = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}}$$

$$\gamma = \frac{\Sigma (x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma (x-\bar{x})^2 \cdot \Sigma (y-\bar{y})^2}}$$

$\gamma_{xy} = \gamma_{uv}$

$u = \frac{x-A}{h}, \quad v = \frac{y-B}{k}$

$$\gamma = \frac{\Sigma (\cancel{A}+hu-\cancel{A}-h\bar{u})(\cancel{B}+kv-\cancel{B}-k\bar{v})}{\sqrt{\Sigma (\cancel{A}+hu-\cancel{A}-h\bar{u})^2 \cdot \Sigma (\cancel{B}+kv-\cancel{B}-k\bar{v})^2}}$$

$x = A+hu, \; y = B+kv$

$\bar{x} = A+h\bar{u}, \; \bar{y} = B+k\bar{v}$

$$= \frac{hk\,\Sigma (u-\bar{u})(v-\bar{v})}{\sqrt{h^2k^2\,\Sigma (u-\bar{u})^2 \cdot \Sigma (v-\bar{v})^2}}$$

$$\gamma_{xy} = \frac{\Sigma (u-\bar{u})\cdot(v-\bar{v})}{\sqrt{\Sigma (u-\bar{u})^2 \cdot \Sigma (v-\bar{v})^2}} = \gamma_{uv}$$

$x = 6, 11, 9, 13, 5, 16, 8$ $\qquad n = 7$

$y = 9, 11, 13, 12, 17, 6, 18$ $\qquad \Sigma y = 86$

$$r_{xy} = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}}$$

~~r_{xy}~~ = ?

x	y	xy	x^2	y^2
6	9	54	36	81
11	11	121	121	121
9	13	117	81	169
13	12	156	169	144
5	17	85	25	289
16	6	96	256	36
8	18	144	64	324
$\Sigma x = 68$	$\Sigma y = 86$	$\Sigma xy = 773$	$\Sigma x^2 = 752$	$\Sigma y^2 = 1164$

$$r_{xy} = \frac{7(773) - (68)(86)}{\sqrt{[7(752) - (68)^2][7(1164) - (86)^2]}}$$

$$r_{xy} = \frac{5411 - 5848}{\sqrt{[5264 - 4624][8148 - 7396]}}$$

$$r_{xy} = \frac{-437}{\sqrt{(640)(752)}} = \frac{-437}{\sqrt{481280}}$$

$$r_{xy} = \frac{-437}{693.74} = \boxed{-0.62}$$