03/01/22 UNIT-2

DESCRIPTIVE STATISTICS: (IInd sub division)

construction of frequency distribution:-

32, 54, 38, 44, 68, 41, 30, 43, 46, 41, $\boxed{40}$, 31, $\boxed{40, 40}$, 36, 46, 48, 32, $\boxed{40}$, 17, 48, 47, 37, 52, 48, 47, 32, 26, 21, 41, 53, 33, 32, 50, 38, 33, 51, 43, 45, 32, $\boxed{40}$, 50, 31, 50, 31, 50, 48, 50, 55, 52, 45, 44.

Data	number (Tally marks)	Frequency.
15-20	I	1
20-25	I	1
25-30	II	2
30-35	~~IIII~~ ~~IIII~~	10
35-40	~~IIII~~ II	7
40-45	~~IIII~~ ~~IIII~~ II	12
45-50	~~IIII~~ ~~IIII~~ II	12
50-55	~~IIII~~ I	6
55-60	-	0
60-65	-	0
65-70.	I	1

52

ht \ W	15-20	20-25	25-30	30-35
1-3	I	I	~~IIII~~	
3-6				I
6-9				

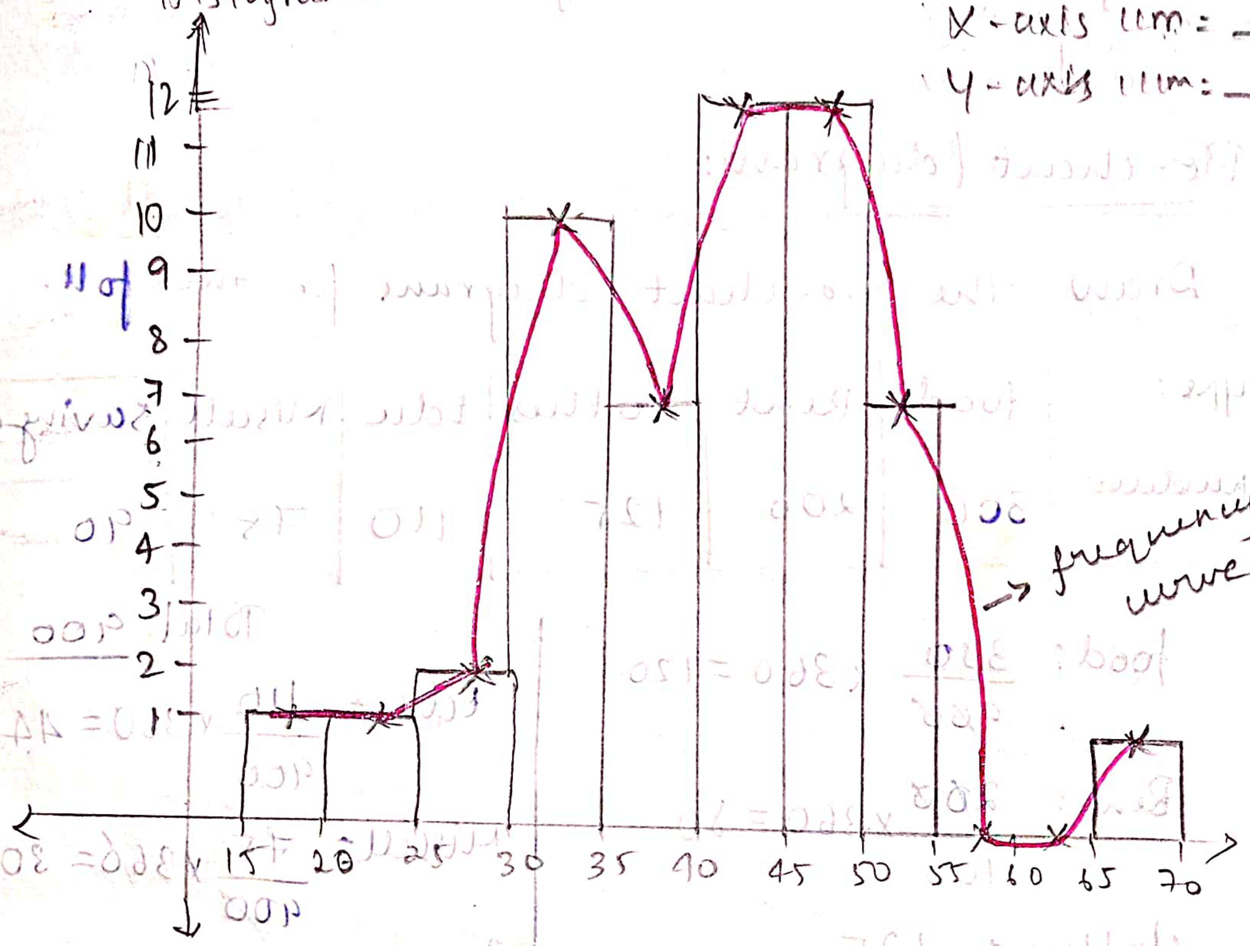
Graphical representation:

Scale

X-axis 1cm = ___

Y-axis 1cm = ___

Histogram

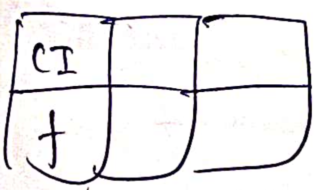


04/01/22 Recall

Tabulation of data.

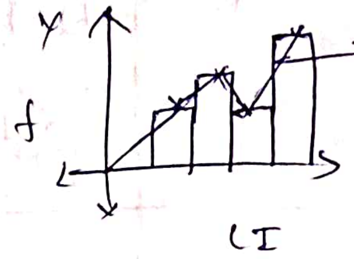
Univariate

Bivariate.



we have constructed histogram.

CI	...
f	...



→ frequency polygon / curve used to represent thickness of the ~~curve~~.

Primary data – Questionnary method.
secondary data

Pie-chart / diagram:

Draw the pie-chart / diagram for the foll.

Type:	food	Rent	clothes	Edu	Miscell	savings
Expenditure:	300	200	125	110	75	90

Total 900

food: $\frac{300}{900} \times 360 = 120$

Rent: $\frac{200}{900} \times 360 = 80$

clothes: $\frac{125}{900} \times 360 = 50$

Edu: $\frac{110}{900} \times 360 = 44$

Miscell: $\frac{75}{900} \times 360 = 30$

Savings: $\frac{90}{900} \times 360 = 36$

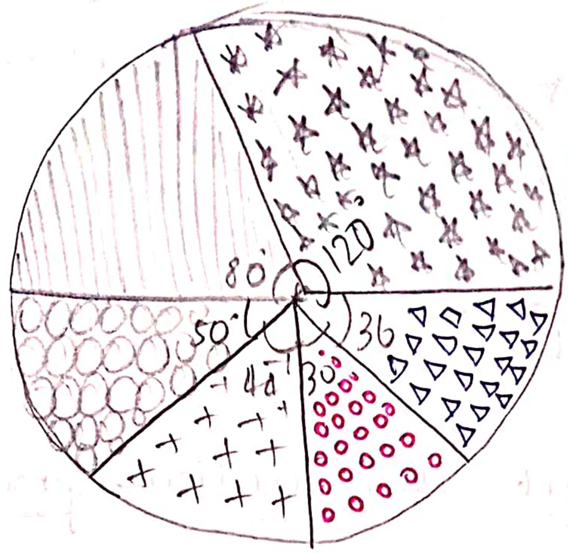
360

Pie-chart:

I^{st} construct a circle start at a pt. just take 120°

PTO

☆ - food
//// - Rent
0000 - clothes
+++ - Edu
o o o - Miscell
ΔΔΔ - savings.



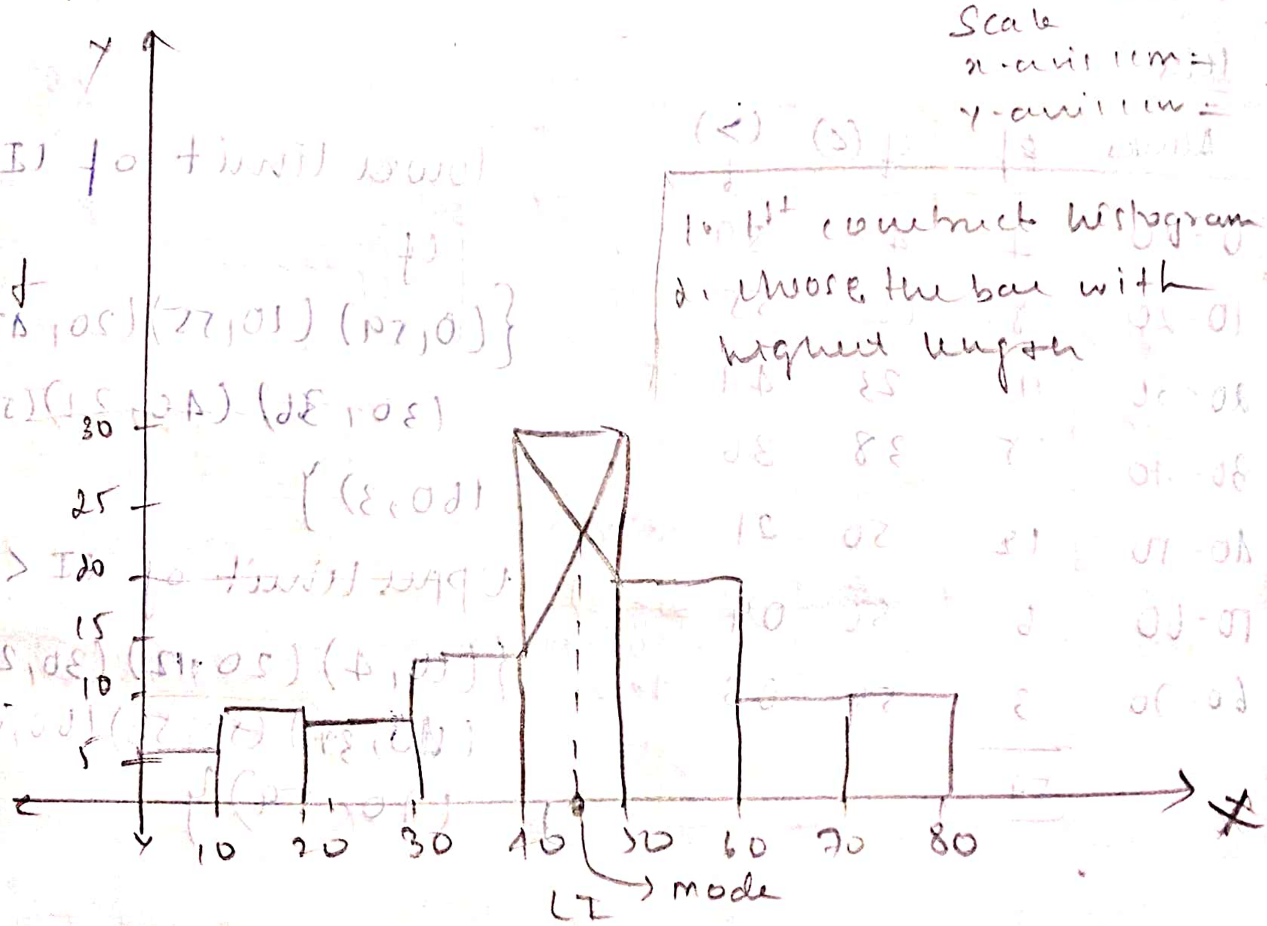
calculation of mode graphically:-

Eg:

CI:	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
f:	5	8	7	12	28	20	10	10

Scale
x-axis 1cm =
y-axis 1cm =

1. 1st construct histogram
2. choose the bar with highest length



$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

where $\Delta_1 = f_1 - f_0$

$\Delta_2 = f_1 - f_2$

x	f	f1	f2
10	7	15	22
20	8	15	24
30	7	16	24
40	9	17	
50	8		

method of grouping

Finding out the median graphically.

OGIVE curves.

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of Students :	4	8	11	15	12	6	3

Soln;

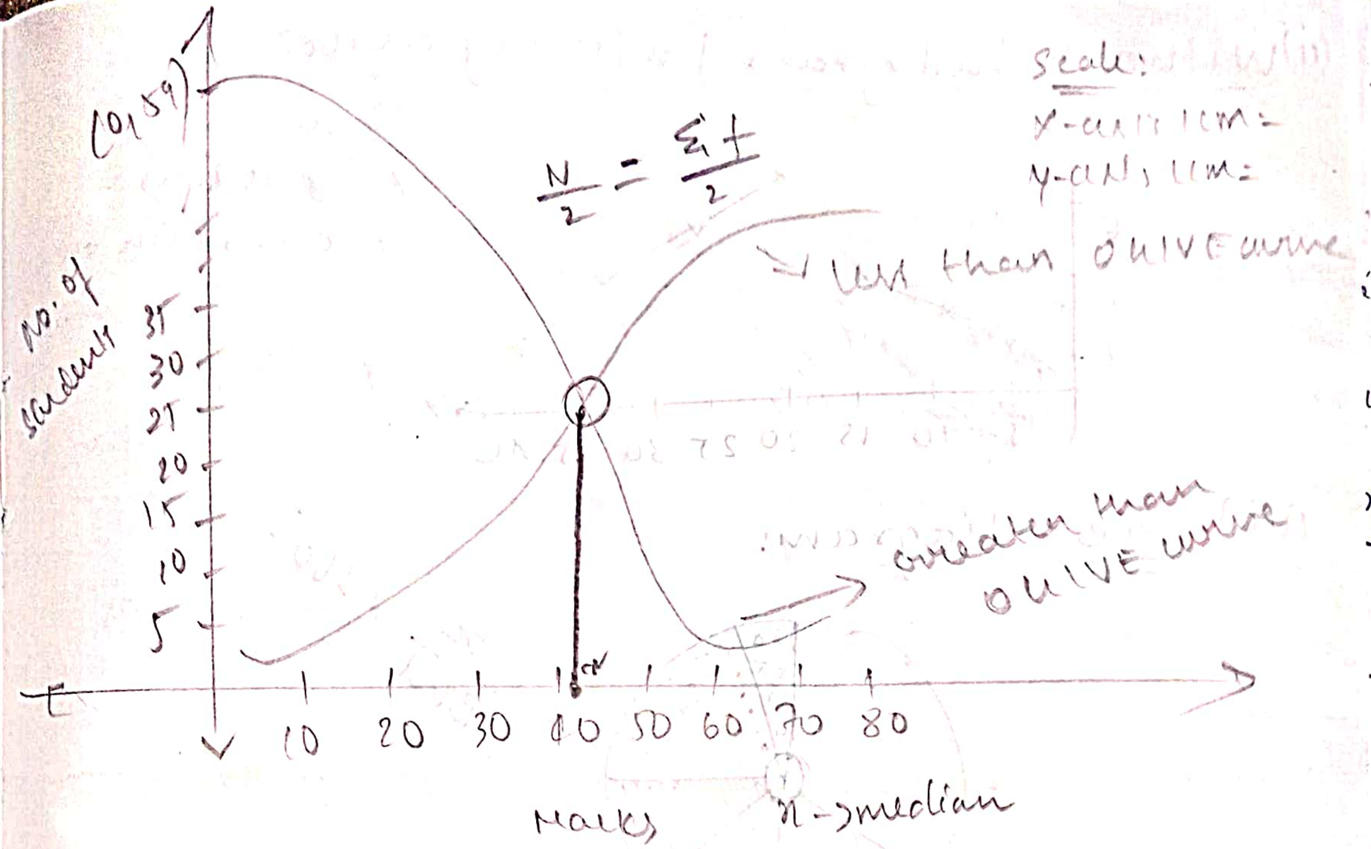
Marks	f	cf (<)	(>) cf
0-10	4	4	59
10-20	8	12	55 (59-4)
20-30	11	23	47 (55-8)
30-40	15	38	36 (47-11)
40-50	12	50	21 (36-15)
50-60	6	56	09 (21-12)
60-70	3	59	03 (9-6)
	59		

lower limit of CI > cf

{(0, 59) (10, 55) (20, 47) (30, 36) (40, 21) (50, 9) (60, 3)}

upper limit of CI < cf

{(10, 4) (20, 12) (30, 23) (40, 38) (50, 50) (60, 56) (70, 59)}



Recalling:-

* construction of pie-chart
* calculation of mode - graphically

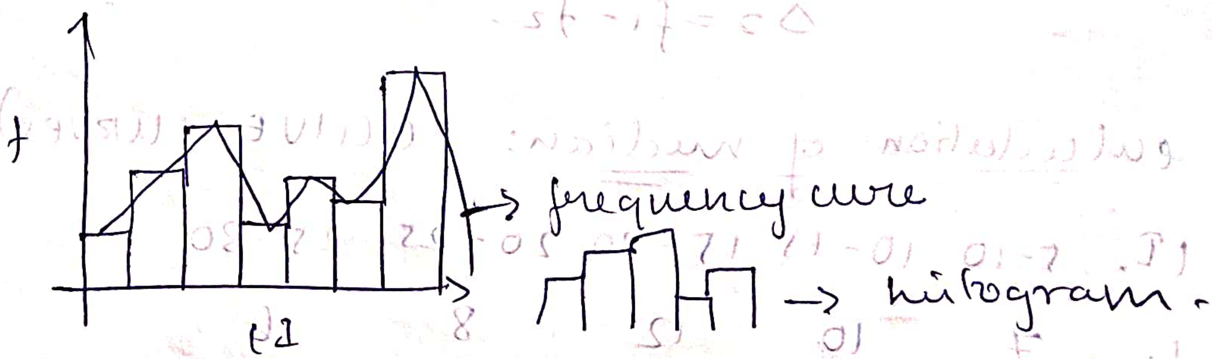
(Pre-requisites)

05/01/22

Recall:-

Graphical and diagrammatic representation of data:

(i) constructed the histogram & frequency curve (polygon).

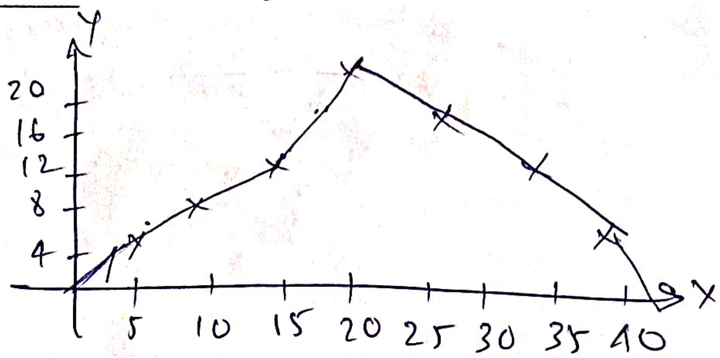


(ii) without histogram frequency curve:-

Scale

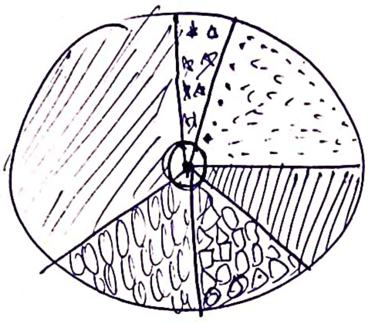
x - axis 1cm =

y - axis 1cm =



(iii) pie - diagram:

$\frac{x}{total} \times 360°$

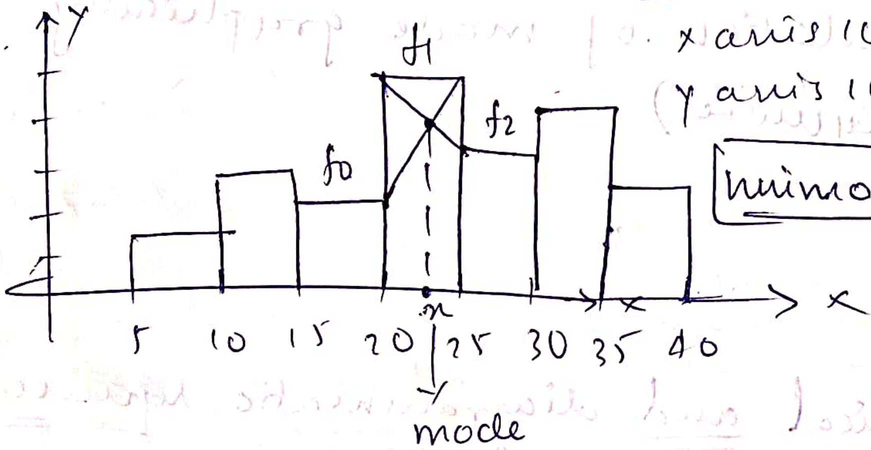


calculation of mode

scale

x axis 1cm = _

y axis 1cm = _



$\boxed{\text{unimodal}}$

$$\text{mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

$$\Delta_1 = f_1 - f_0$$

$$\Delta_2 = f_1 - f_2.$$

calculation of median: (OGIVE CURVES).

CI:	5-10	10-15	15-20	20-25	25-30
f:-	7	10	12	8	6

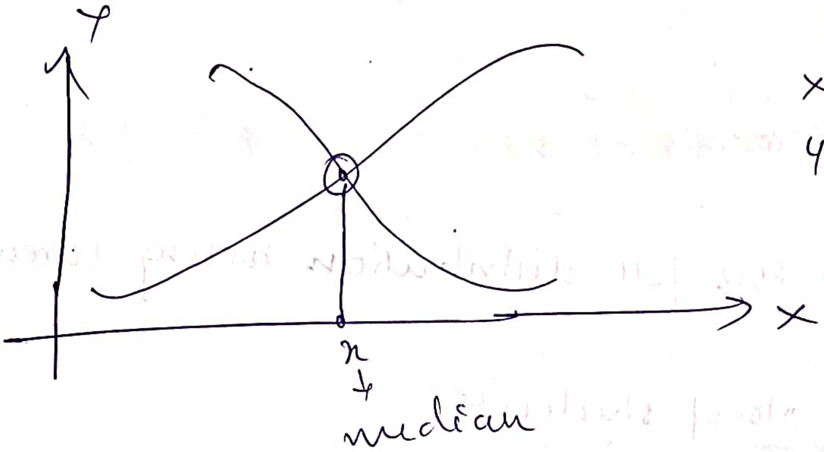
$<cf$:- 7 17 29 39 43

$>cf$: 43 36 26 14 6.

(5, 43) (10, 36) (15, 26) (20, 14) (25, 6) → one olive curve;

(10, 7) (15, 17) (20, 29) (25, 37) (30, 43) → another olive curve.

One is called as less than olive curve, and another is > than olive curve



Scale
x axis 1 cm =
y axis 1 cm =

lower limit → cf
upper limit <cf

Step 1: Alternatively we can construct either less than olive curve (or) > than olive curve

Step 2: Identify $\frac{N}{2} = \frac{\Sigma f}{2}$ on the curve

Step 3: Draw the ⊥ from the point on the curve i.e; $\frac{N}{2}$ to X-axis.

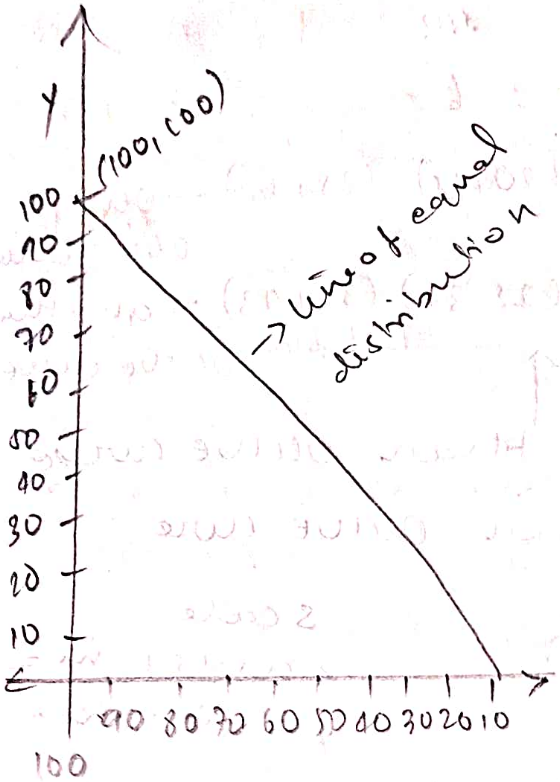
Lorenz curve:

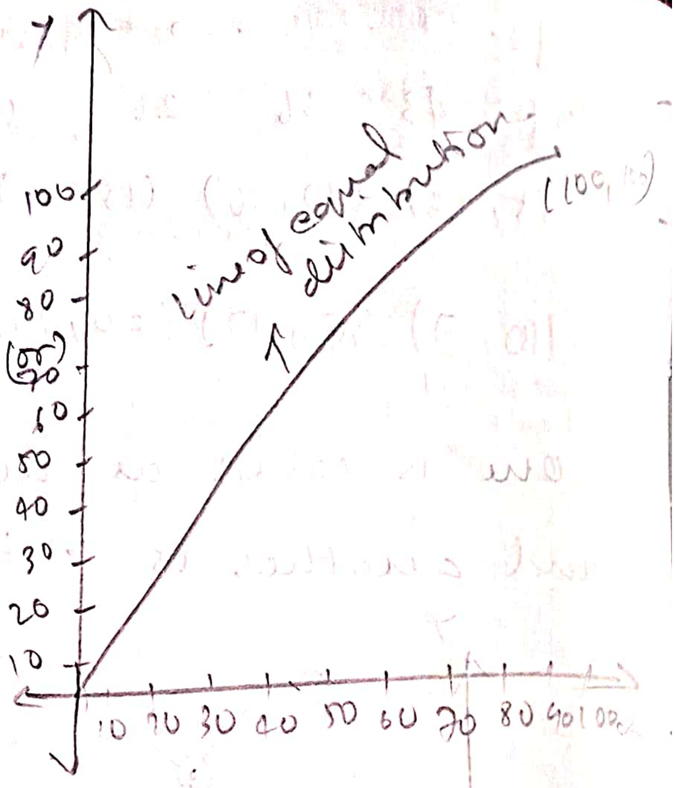
To measure the inequalities (variations) in the distribution.

Variability is data using Lorenz curve

$$\boxed{CV = \frac{\sigma}{\bar{x}} \times 100}$$

PTO





Eg: Represent the foll. distribution using lorenz curve.

Income	No. of students	
	Group A	Group B
10 ✓	15	8
30 ✓	5	6
35 ✓	1	6
45 ✓	2	2
80 ✓	2	3

Interpret the variability.

25
135
/200
1.20/200

cummulative percentage of income:

Income	cummulative income	cummulative %
10	10	$\frac{10}{200} \times 100 = 5$
30	40	$\frac{40}{200} \times 100 = 20$
35	75	$\frac{75}{200} \times 100 = 37.5$
45	120	$\frac{120}{200} \times 100 = 60$
80	200	$\frac{200}{200} \times 100 = 100$

cummulative % of group A.

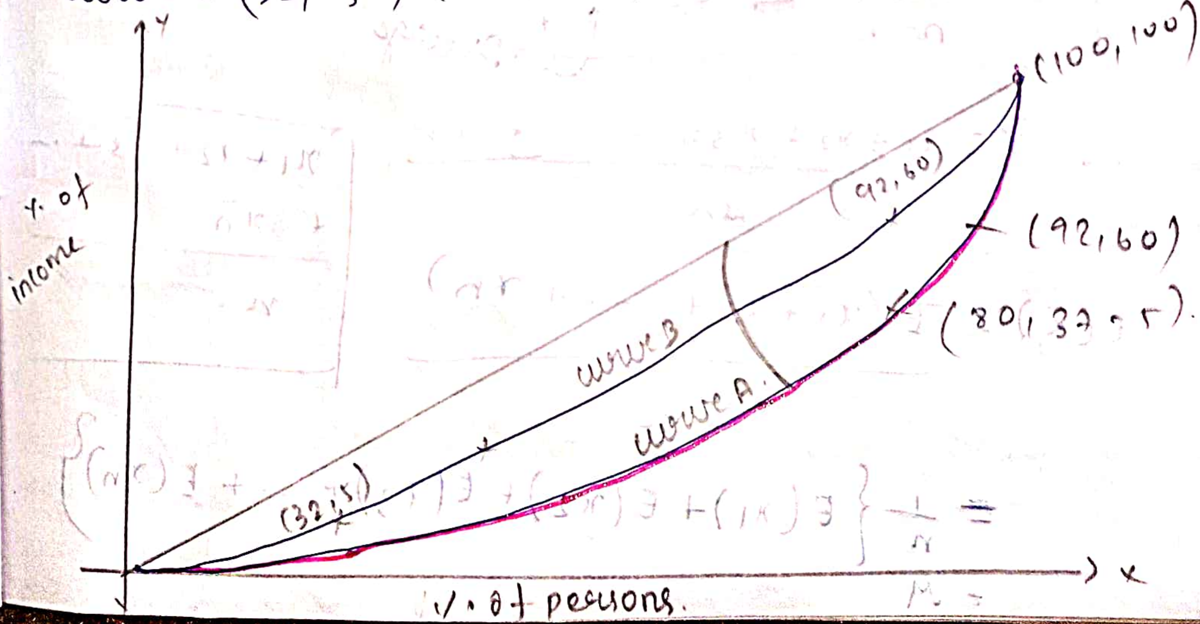
Group A	cummulative values	cummulative %.
15	15	$\frac{15}{25} \times 100 = 60$
5	20	$\frac{20}{25} \times 100 = 80$
1	21	$\frac{21}{25} \times 100 = 84$
2	23	$\frac{23}{25} \times 100 = 92$
2	25	$\frac{25}{25} \times 100. = 100$

curve A (60, 5) (80, 20) (84, 37.5) (92, 60) (100, 100).

cummulative % of Group B.

Group B	cummulative values	cummulative %.
8	8	$\frac{8}{25} \times 100 = 32$
6	14	$\frac{14}{25} \times 100 = 56$
6	20	$\frac{20}{25} \times 100 = 80$
2	22	$\frac{22}{25} \times 100 = 88.$
3	25	$\frac{25}{25} \times 100 = 100$

curve B (32, 5) (56, 20) (80, 37.5) (88, 60) (100, 100).



curve A is far from line of equal distribution and ∴ group A has greater variability.

06/01/22. CIA-2 Revision.

Small sample test:

sampling theory

large sample test — small sample test

χ^2 - test.

F - test.

1. the height of ~~the~~ 10 students are found to be 70, 67, 62, 68, 61, 68, 70, 64, 64, 66 inches. Is it reasonable to believe that the avg. height is > than 64 inches?

Soln:-

H_0: $\mu = 64$ inches

H_1: $\mu > 64$ inches.

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n}} \text{ with } n-1 \text{ d.f.}$$

$$E\left(\frac{ns}{n-1}\right)^2 = \sigma^2$$

$\bar{x} \rightarrow$ sample mean, μ = population mean.

$$\boxed{S^2 = \frac{1}{n-1} \sum (x - \bar{x})^2}$$

Average

$E(\bar{x}) = \mu$.

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

$$\frac{x_1 + x_2 + x_3 + \dots + x_n}{n}.$$

$$E(\bar{x}) = \frac{E(x_1 + x_2 + \dots + x_n)}{n}$$

$$= \frac{1}{n}\{E(x_1) + E(x_2) + E(x_3) + \dots + E(x_n)\}$$

$$= \mu$$

x: 70, 67, 62, 68, 66, 68, 70, 64, 64, 60.

$$\bar{x} = \frac{70+67+62+68+66+68+70+64+64+60}{10}$$

$$\bar{x} = \frac{66\cancel{0}}{1\cancel{0}} = 66.$$

$$\Sigma (x-\bar{x})^2 = (70-66)^2 + (67-66)^2 + (62-66)^2 + (68-66)^2 + (61-68)^2 + (70-66)^2 + (64-66)^2 + (61-66)^2 + (66-66)^2$$

$$= 90.$$

$$s^2 = \frac{90}{9} = 10$$

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{66-64}{\sqrt{10}/\sqrt{10}} = 2.$$

Table value $t = 1.83$ $\quad t = 2 > 1.83$.

Reject the null hypothesis at 5% level.

---

2. Increase of BP is recorded as 5, 2, 8, -1, 3, 0, 6, -2, 1, 5, 0, 4, & 6. (Paired t-test).

---

3. Increase in wt

food A	49	53	51	52	47	50	52	53
food B	52	55	52	53	50	54	54	53

can we conclude food A is better than food B?

Soln.

$H_0: \mu_x = \mu_y$

H_1: food A is better than the food B.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$\bar{x}_1 = \frac{49+53+51+52+47+50+52+53}{8}$$

$$= 50.875$$

$$\bar{x}_2 = \frac{52+55+52+53+50+54+54+53}{8}$$

$$= 52.875.$$

$$\Sigma (x_1 - \bar{x}_1)^2 = 30.875 \;;\; \Sigma (x_2 - \bar{x}_2)^2 = 16.875$$

$$s^2 = \frac{\Sigma (x_1 - \bar{x}_1)^2 + \Sigma (x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2} = 3.41.$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (n_1 + n_2 - 2)\,df.$$

$$= -2.17.$$

Calculated value -2.17 is less than table value -1.76 at 5% level.

Accept H_0.

06/01/22

1. When the 1st proof of 392 pages of a book of 1200 pages were read the distribution of printing mistakes were found as follows.

No. of mistakes (x): 0 1 2 3 4 5 6.

No. of pages : (f): 275 72 30 7 5 2 1 (392)

Fit a position distribution to the above data & test the gradness of fit.

soln:

$$f(r) = N \cdot \frac{e^{-\lambda} \lambda^{x}}{\lfloor x}$$ → pdf possion distribution

$N \to 392$

$$P(X = x) = \frac{e^{-\lambda} \lambda^{x}}{\lfloor x}$$ $\lambda \to$ mean

$$\lambda = \frac{0 \times 275 + 1 \times 72 + 2 \times 30 + 3 \times 7 + 4 \times 5 + 5 \times 2 + 6 \times 1}{392}$$

$\lambda = 0.482$

$$f(r) = 392 \times e^{-0.482} \frac{(0.482)^{r}}{\lfloor r} \quad r = 0, 1, 2, 3, 4, 5, 6$$

$f(0) = 242.1$

$f(1) = 116.7$

$f(2) = 28.1$

$f(3) = 4.5$

$f(4) = 0.5$

$f(5) = 0.1$

$f(6) = 0$

392

x	observed freq (O)	Expected freq (E)	$(O-E)^2$	$\frac{(O-E)^2}{E}$
0	275	242.1	1082.41	4.471
1	72	116.7	1998.09	17.121
2	30	28.1	3.61	0.128
3	7 } 15	4.5 } 5.1		
4	5	0.5	98.01	19.217
5	2	0.1		
6	1	0		40.937.

Ho: the value of O & E are equal.

H₁: The value of O & E are unequal.

$$\chi^2 = \sum_1 \frac{(O-E)^2}{E} \text{ with } (n-1) \text{ df}.$$

Here revised value of n is 4.

$\therefore \chi^2 = 40.937$ with 3 df.

Table value of χ^2 for 3 df at 5% level is 7.815. Cal value is > table value, we reject H_0

---

2. Two independent groups of 10 children were tested to find how many digits they could repeat from memory after hearing them. The results are

Group A: 8 6 5 7 6 8 7 4 5 6

Group B: 10 6 7 8 6 9 7 6 7 7.

Is the difference b/w the mean scores of the two groups significant?

Soln. H_0: mean scores of two groups are same.

H_1: " " " " " diff. significant.

$$t = \frac{\bar{x} - \bar{y}}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ with } (n_1+n_2-2) \text{ df}$$

$$S = \frac{\sum_1 (x-\bar{x})^2 + \sum_1 (y-\bar{y})^2}{n_1+n_2-2}$$

$\bar{x} = \frac{62}{10} = 6.2$, $\bar{y} = \frac{73}{10} = 7.3$.

P.T.O

x	$x-6.2$	$(x-6.2)^2$	y	$y-7.3$	$(y-7.3)^2$
8	1.8	3.24	10	2.7	7.29
6	-0.2	0.04	6	-1.3	1.69
5	-1.2	1.44	7	-0.3	0.09
7	0.8	0.64	8	0.7	0.49
6	-0.2	0.04	6	-1.3	1.69
8	1.8	3.24	9	1.7	2.89
7	0.8		7	-0.3	0.09
4	-2.2		6	-1.3	1.69
5	-1.2		7	-0.3	0.09
6	-0.2		7	-0.3	0.09
		15.6			16.1

$$s^2 = \frac{15.6 + 16.1}{10+10-2} = 1.76.$$

$$t = \left| \frac{6.2 - 7.3}{\sqrt{1.76\left(\frac{1}{10}+\frac{1}{10}\right)}} \right| = 1.8 \text{ with } 18 \text{ df}$$

Table value of t for 18 df at 5% level is 1.73,

$1.8 > 1.73$ Reject H_0 at 5% level.

3. Two independent samples of 8 & 7 items respectively had the foll. values of the variables.

I: 9 11 13 11 15 9 12 14

II: 10 12 10 14 9 8 10

Do estimate of population variance differ significantly?

Soln:

$H_0: \sigma_1 = \sigma_2$; $H_1: \sigma_1 \neq \sigma_2$

$$F = \frac{S_1^2}{S_2^2} \text{ with } (n_1 - 1, n_2 - 1) \text{ d.f.}$$

$$S_1^2 = \frac{1}{n-1} \sum (x - \bar{x})^2$$

$$\bar{x} = \frac{94}{8} = 11.75 \; ; \; \bar{y} = \frac{73}{7} = 10.42$$

x	$x - 11.75$	$(x - 11.75)^2$	y	$y - 10.42$	$(y - 10.4)^2$
9	-2.75		10	-0.42	0.1764
11	-0.75		12	1.58	2.4964
13	1.25		10	-0.42	0.1764
11	-0.75		14	3.58	12.8164
15	3.25		9	-1.42	2.0614
9	-2.75		8	-2.42	5.8564
12	0.25		10	-0.42	0.1764
14	2.25				23.7
		33.5			

$$S_1^2 = \frac{\sum (x - \bar{x})^2}{n_1 - 1} = \frac{33.5}{7} = 4.78$$

$$S_2^2 = \frac{\sum (y - \bar{y})^2}{n_2 - 1} = \frac{23.7}{6} = 3.95$$

$$F = \frac{S_1^2}{S_2^2} \; (n_1 - 1, n_2 - 1) \text{ d.f.}$$

$$= \frac{4.78}{3.95} = 1.21 \text{ with } (7, 6) \text{ d.f}$$

$F|_{(7,6)\,df} = 3.81$ (5% level) $1.21 < 3.81$

Accept H_0 at 5% level.

Q. Test the significance of the values of correlation co-eff 'r' obtained from samples of size n pairs from a bivariate normal table.

Soln: $r = 0.6$, $n = 38$.

$H_0: \rho = 0$, $H_1 = \rho \neq 0$

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \rightarrow \text{with } (n-2)\text{ df} = \frac{0.6\sqrt{36}}{\sqrt{1-0.6^2}} = \frac{0.6(6)}{\sqrt{1-0.36}}$$

Table value of t at 5% level for 36 df is

compare |t| with table value of t & decide the equal conclusion

$$= \frac{0.6(6)}{\sqrt{0.64}}$$

$$= \frac{\overset{0.3}{\cancel{0.6}}(\cancel{6})^3}{\cancel{0.8}}$$

$$= 4.5$$

07/01/22 Revision CIA - 2

Expection of a random variable:

[single dimensional rv]

$E(X) = \sum_i x_i p_i$ (Discrete case)

$\sigma^2 = var(x) = E(X - E(x))^2 = E(x^2) - (E(x))^2$

$E(x) = \int_{-\infty}^{\infty} x f(x)\, dx$ (continuous case).

1. If x and y are independent random variable, then $E(xy) = E(x) \cdot E(y)$

If x and y are independent r.v then $var(ax+by) = $ ___

Soln.

$$Var(x) = E(x^2) - (E(x))^2.$$

$$var(ax+by) = E((ax+by)^2) - [E(ax+by)]^2.$$

$$= E[a^2x^2 + 2ab\,xy + b^2y^2] - [aE(x) + bE(y)]^2$$

$$= a^2E(x^2) + 2abE(xy) + b^2E(y^2) - a^2(E(x))^2 - b^2(E(y))^2 - 2abE(x)E(y)$$

$$= a^2\{E(x^2) - (E(x))^2\} + b^2\{E(y^2) - (E(y))^2\}$$

$$= a^2 var(x) + b^2 var(y)$$

P.1) find E(x) & var(x) for the foll.

x:	1	2	3	4	5
P(x):	10/36	7/36	6/36	7/36	6/36.

1.

Soln:

$$E(x) = \sum_i x_i p_i = 1 \times \frac{10}{36} + 2 \times \frac{7}{36} + 3 \times \frac{6}{36} + \frac{4 \times 7}{36} + 5 \times \frac{6}{36}.$$

$$= \frac{1}{36}[10 + 14 + 18 + 28 + 30] = \frac{100}{36}$$

$$Var(x) = E(x^2) - (E(x))^2$$

$$E(x^2) = \sum_i x^2 p_i = \frac{1\times10}{36} + \frac{4\times7}{36} + \frac{9\times6}{36} +$$

$$\frac{16\times7}{36} + \frac{25\times6}{36}.$$

$$= \frac{1}{36}\left\{10+28+54+112+150\right\} = \frac{354}{36}.$$

$$var(x) = \frac{354}{36} - \left(\frac{100}{36}\right)^2 = \frac{36\times354 - 100\times100}{36\times36}.$$

$$= \_\_\_\_\_\_$$

continous distribution.

$$E(x) = \int_{-\infty}^{\infty} x f(x)\,dx$$

$$var(x) = E(x^2) - (E(x))^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \left[\int_{-\infty}^{\infty} x f(x)\,dx\right]^2.$$

Eg: The pdf of a r.v x is $f(x) = \begin{cases} \frac{3}{2} - x & 0<x<1 \\ 0 & \text{otherwise} \end{cases}$

find $E(x)$ & $var(x)$.

$$E(x) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 x\left(\frac{3}{2} - x\right)dx.$$

$$= \int_0^1 \left\{\frac{3x}{2} - x^2\right\}dx$$

$$= \left(\frac{3x^2}{4} - \frac{x^3}{3}\right)_0^1 = \frac{3}{4} - \frac{1}{3}$$

$$= \frac{9-4}{12} = \frac{5}{12}.$$

$var(X) = E(X^2) - (E(X))^2.$

$$E(X^2) = \int_0^1 x^2\left(\frac{3}{2} - x\right)dx = \int_0^1 \left(\frac{3x^2}{2} - x^3\right)dx$$

$$= \left\{\frac{x^3}{2} - \frac{x^4}{4}\right\}_0^1 = \frac{1}{2} - \frac{1}{2}$$

$$= \frac{1}{4}.$$

$$var(X) = E(X^2) - (E(X))^2$$

$$= \frac{1}{4} - \left(\frac{5}{12}\right)^2 = \frac{1}{4} - \frac{25}{144}$$

$$= \frac{36-25}{144} = \frac{11}{144}$$

---

3. Let $f(x_1, x_2) = \begin{cases} 21x_1^2x_2^3 & 0 < x_1 < x_2 < 1 \\ 0 & \text{elsewhere} \end{cases}$

be a joint PDF of r.v's X_1 & X_2. Find the conditional mean & variance of X_1 given $X_2 = x_2$, $0 < x_2 < 1$.

soln:

$E(X_1 / X_2 = x) = ?$

$var(X_1 / X_2 = x) = ?$

$$E(X_1X_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f(x_1, y_2)\,dx_1\,dx_2.$$

$$f(x_1, x_2) = \begin{cases} 21x_1^2x_2^3 & 0 < x_1 < x_2 < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

To find $f(x_1 | x_2) =$

$$f(x_1 | x_2) = \frac{f(x_1, x_2)}{f_2(x_2)}$$

$$f_2(x_2) = \int_0^{x_2} f(x_1, x_2)\,dx_1 = 21 x_2^3 \int_0^{x_2} x_1^2\,dx_1$$

marginal density for of X_2

$$= \cancel{21}^{7}\, x_2^3 \left(\frac{x^3}{\cancel{3}}\right)_0^{x_2}$$

$$= 7x_2^6 \qquad 0 < x_2 < 1$$

$$f(x_1|x_2) = \frac{21 x_1^2 x_2^3}{7 x_2^6} = \frac{3x_1^2}{x_2^3} \qquad 0 < x_1 < x_2, < 1$$

$0 < x_1 < x_2$

$0 < x_2 < 1$

conditional density for of X_1 given X_2

$$E(x_1|x_2) = \int_0^{x_2} x_1 f(x_1|x_2)\,dx_1$$

mean

$$= \int_0^{x_2} x_1 \left\{\frac{3x_1^2}{x_2^3}\right\} dx_1 = \frac{3}{x_2^3}\left\{\frac{x_1^4}{4}\right\}_0^{x_2}$$

$$= \frac{3x_2^4}{4x_2^3}$$

$$= \frac{3}{4} x_2 \qquad 0 < x_2 < 1$$

$$E(x_1^2|x_2) = \int_0^{x_2} x_1^2 f(x_1|x_2)\,dx_1$$

$$= \int_0^{x_2} x_1^2 \left\{\frac{3x_1^2}{x_2^3}\right\} dx_1 = \frac{3}{x_2^3}\int_0^{x_2} x_1^4\,dx_1$$

$$= \frac{3}{x_2^3}\left\{\frac{x_2^5}{5}\right\}$$

$$= \frac{3}{5} x_2^2 \qquad 0 < x_2 < 1.$$

$$\operatorname{var}(x_1 / x_2 = x_2) = \frac{3}{5} x_2^2 - \left(\frac{3}{4} x_2\right)^2 \quad 0 < x_2 < 1$$

$$= \frac{3}{5} x_2^2 - \frac{9}{16} x_2^2$$

$$= \frac{48 - 45}{80} x_2^2 .$$

Y \ X	x_1	x_2	x_3	...	x_n	Total
y_1	P_{11}	P_{12}	P_{13}	...	P_{1n}	a_1
y_2	P_{21}	P_{22}	P_{23}	---	P_{2n}	a_2
y_3	P_{31}	P_{32}	P_{33}	---	P_{3n}	a_3
⋮	⋮	⋮	---	---	---	⋮
y_m	P_{m1}	P_{m2}	P_{m3}	---	P_{mn}	a_m
Total f.	b_1	b_2	b_3	---	b_n	1

→ marginal density for of Y.

marginal density for of X.

$$\sum a_i = 1$$

$$\sum b_j = 1 .$$

$$f(Y / X = x_2)$$

conditional distribution $(Y/X) = ? \rightarrow n$

Conditional distribution $(X/Y) = ? \rightarrow m$