University Questions Solution set

Unit - III Statistics & Probability

Chapt 6: Statistics, Correlation & Regression

1. The first four moments of a distribution about the value 5 are 3, 30, 50 & 60. Obtain first four central moments & coefficient of Skewness & Kurtosis. (Nov 15)

→ $A = 5$, $\mu_1' = 3$, $\mu_2' = 30$, $\mu_3' = 50$, $\mu_4' = 60$.

Now, we calculate central moments.

$\mu_1 = 0.$

$\mu_2 = \mu_2' - (\mu_1')^2 = 30 - 3^2 = 21$

$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3$

$= 50 - 3(3)(30) + 2(3)^3$

$\mu_3 = -166$

$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6(\mu_1')^2\mu_2' - 3(\mu_1')^4$

$= 60 - 4(3)(50) + 6(3)^2(30) - 3(3)^4$

$\mu_4 = 1323$

Coefficient of Skewness is given by

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(-166)^2}{(21)^3}$$

$$\beta_1 = 2.9755$$

Coefficient of Kurtosis is given by

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{1323}{(21)^2}$$

$$\beta_2 = 3$$

Since β_2 is exactly three, the distribution is mesokurtic.

---

2. The first three moments of a distribution about the value 2 of a distribution are 1, 16 & −40. Find the mean, S.D. & skewness of the distribution. (May 16)

→ $A = 2, \mu_1' = 1, \mu_2' = 16, \mu_3' = -40.$

$$\text{Mean} = \bar{x} = A + \mu_1' = 2 + 1 = 3$$

$$\therefore \bar{x} = 3.$$

Now, we calculate central moments.

Scanned by CamScanner

$\mu_1 = 0$

$\mu_2 = \mu_2' - (\mu_1')^2 = 16 - 1 = 15$

$\mu_2 = 15$

$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3$

$= -40 - 3(1)(16) + 2(1)^3$

$\mu_3 = -86$

standard deviation $= \sqrt{\mu_2} = \sqrt{15} = 3.8730$

Coefficient of skewness is

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(-86)^2}{(15)^3} = 2.1914$$

---

3. If first four moments of a distribution about the value 5 are equal to $-4, 22, -117$ & 560, determine central moments & β_1, β_2. (Dec 16)

→ $A = 5,\ \mu_1' = -4,\ \mu_2' = 22,\ \mu_3' = -117,\ \mu_4' = 560.$

$\mu_1 = 0$

$\mu_2 = \mu_2' - (\mu_1')^2 = 22 - (-4)^2 = 6$

$\mu_2 = 6$

$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3$$

$$= (-117) - 3(-4)(22) + 2(-4)^3$$

$$\mu_3 = 19$$

$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6(\mu_1')^2\mu_2' - 3(\mu_1')^4$$

$$= 560 - 4(-4)(-117) + 6(-4)^2(22)$$

$$-3(-4)^4$$

$$\mu_4 = 32$$

Coeff. of skewness β_1 is

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{19^2}{6^3} = 1.6713$$

Coeff of kurtosis is

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{32}{6^2} = 0.8889$$

---

4. The 1st four moments about working mean 5 are 7, 70, 140, 175. Calculate central moments & hence β_1 & β_2. (May 17)

→ $\mu_1' = 7$, $\mu_2' = 70$, $\mu_3' = 140$, $\mu_4' = 175$.

$\mu_1 = 0$.

$$\mu_2 = \mu_2' - (\mu_1')^2 = 70 - 7^2 = 21$$

$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3$$

$$= 140 - 3(7)(70) + 2(7)^3$$

$$\mu_3 = -644$$

$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6(\mu_1')^2\mu_2' - 3(\mu_1')^4$$

$$= 175 - 4(7)(140) + 6(7)^2(70) - 3(7)^4$$

$$\mu_4 = 9632$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(-644)^2}{(21)^3} = 44.7830$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{9632}{(21)^2} = 21.8413$$

5. Calculate first three moments of following distribution about the mean:

(Dec 17)

x	f
0	1
1	8
2	28
3	56
4	70
5	56
6	28
7	8
8	1

Taking $A = 4$, $d = x - 24$

x	f	$d = x - \cancel{2}4$	fd	fd^2	fd^3
0	1	-4	-4	16	-64
1	8	-3	-24	72	-216
2	28	-2	-56	112	-224
3	56	-1	-56	56	-56
4	70	0	0	0	0
5	56	1	56	56	56
6	28	2	56	112	224
7	8	3	24	72	216
8	1	4	4	16	64
	$\Sigma f = 256$		$\Sigma fd = 0$	$\Sigma fd^2 = 512$	$\Sigma fd^3 = 0$

$$\mu_1' = \frac{\Sigma fd}{\Sigma f} = 0$$

$$\mu_2' = \frac{\Sigma fd^2}{\Sigma f} = \frac{512}{256} = 2$$

$$\mu_3' = \frac{\Sigma fd^3}{\Sigma f} = 0$$

$\boxed{\mu_1 = 0}$, $\mu_2 = \mu_2' - (\mu_1')^2 = 2$, $\boxed{\mu_2 = 2}$

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3$$

$$= 0 - 3(0) + 2(0)$$

$\boxed{\mu_3 = 0}$

Examples on Correlation

Properties :-

i. If $r(x,y) > 0$, then correlation is +ve.

ii. If $r(x,y) < 0$, then correlation is -ve,

if $r(x,y) = 0$, there is no correlation.

iii. It measures only linear relationship.

iv. It determines a single value which summarizes the extent of linear relationships.

Examples :-

1. Find coeff of correlation for the data:

$n=25$; $\Sigma x_i = 100$, $\Sigma y_i = 125$, $\Sigma x_i^2 = 250$, $\Sigma y_i^2 = 500$, $\Sigma x_i y_i = 522$.

→ The Karl-Pearson's coeff. of correlation

$$r(x,y) = \frac{\Sigma x_i y_i - n\bar{x}\bar{y}}{\sqrt{\left(\Sigma x_i^2 - n(\bar{x})^2\right)\left(\Sigma y_i^2 - n(\bar{y})^2\right)}}$$

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{100}{25} = 4$$

$$\bar{y} = \frac{\Sigma y_i}{n} = \frac{125}{25} = 5$$

$$\therefore r(x,y) = \frac{522 - 25 \times 4 \times 5}{\sqrt{(250 - 25 \times 16)(500 - 25 \times 25)}}$$

$$= \frac{522 - 500}{\sqrt{(-150)(-125)}}$$

$$\boxed{r = 0.160665}$$

2. Find coefficient of correlation for the following data:

x_i	10	14	18	22	26	30
y_i	18	12	24	6	30	36

→ Karl-Pearson's coeff. of correlation $r(x,y)$

$$r(x,y) = \frac{cov(x,y)}{\sigma_x \sigma_y}$$

$$= \frac{\frac{1}{n}\Sigma x_i y_i - \bar{x}\bar{y}}{\sqrt{\frac{1}{n}\Sigma x_i^2 - (\bar{x})^2}\sqrt{\frac{1}{n}\Sigma y_i^2 - (\bar{y})^2}}$$

$$\bar{x} = \frac{\Sigma x_i}{n}, \quad \bar{y} = \frac{\Sigma y_i}{n}, \quad n = 6.$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
10	18	180	100	324
14	12	168	196	144
18	24	432	324	576
22	6	132	484	36
26	30	780	676	900
30	36	1080	900	1296
$\Sigma x_i = 120$	$\Sigma y_i = 126$	$\Sigma x_i y_i = 2772$	$\Sigma x_i^2 = 2680$	$\Sigma y_i^2 = 3276$

$$\bar{x} = \frac{120}{6} = 20 \quad , \quad \bar{y} = \frac{126}{6} = 21$$

$$r(x,y) = \frac{\frac{1}{6}(2772) - (20)(21)}{\sqrt{\frac{1}{6}(2680) - (20)^2}\sqrt{\frac{1}{6}(3276) - (21)^2}}$$

$$\boxed{r(x,y) = 0.60}$$

3. Obtain correlation coeff. betn population density (per square miles) & death rate (per thousand persons) from following data:

Population density	200	500	400	700	300
Death rate	12	18	16	21	10

→ Let x_i = Population density

y_i = Death rate

$$r(x,y) = \frac{cov(x,y)}{\sigma_x \sigma_y}$$

$$= \frac{\Sigma u_i v_i - n(\bar{u}\bar{v})}{\sqrt{\Sigma u_i^2 - n(\bar{u})^2}\sqrt{\Sigma v_i^2 - n(\bar{v})^2}}$$

$\bar{u} = \frac{\Sigma u_i}{n}$, $\bar{v} = \frac{\Sigma v_i}{n}$, $n = 5$

$u_i = x_i - a$, $v_i = y_i - b$

$a = 400$, $b = 16$

x_i	y_i	$u_i = x_i - 400$	$v_i = y_i - 16$	u_i^2	v_i^2	$u_i v_i$
200	12	-200	-4	40,000	16	800
500	18	100	2	10000	4	200
400	16	0	0	0	0	0
700	21	300	5	90000	25	1500
300	10	-100	-6	10000	36	600
		$\Sigma u_i = 100$	$\Sigma v_i = -3$	$\Sigma u_i^2 = 150000$	$\Sigma v_i^2 = 81$	$\Sigma u_i v_i = 3100$

$\bar{u} = \frac{100}{5}$, $\bar{v} = \frac{-3}{5} = -0.6$.

$$\therefore r(x,y) = \frac{3100 - 5(2)(-0.6)}{\sqrt{150000 - 5(20)^2}\sqrt{81 - 5(-0.6)^2}} = 0.92298$$

Scanned by CamScanner

Examples On Regression

1) Obtain regression lines for the follo. data:

X	2	3	5	7	9	10	12	15
Y	2	5	8	10	12	14	15	16

Find estimate of i) Y when X = 6

& ii) X when Y = 20.

→ To find regression lines we require to calculate regression coeffs. b_{xy} & b_{yx}.

These coeffs. depends upon Σx, Σy, Σx^2, Σy^2, Σxy. So we prepare follo. table:

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
2	2	4	4	4
3	5	9	25	15
5	8	25	64	40
7	10	49	100	70
9	12	81	144	108
10	14	100	196	140
12	15	144	225	180
15	16	225	256	240
63	82	637	1014	797

$n =$ no. of pairs of observations $= 8$

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{63}{8} = 7.875$$

$$\sigma_x^2 = \frac{\Sigma x_i^2}{n} - (\bar{x})^2 = 17.6094$$

$$\bar{y} = \frac{\Sigma y_i}{n} = \frac{82}{8} = 10.25$$

$$\sigma_y^2 = \frac{\Sigma y_i^2}{n} - (\bar{y})^2 = 21.6875$$

$$\text{cov}(x,y) = \frac{\Sigma x_i y_i}{n} - \bar{x}\bar{y} = \frac{797}{8} - 7.875 \times 10.25$$

$$= 18.9063$$

$$b_{yx} = \frac{\text{cov}(x,y)}{\sigma_x^2} = \frac{18.9063}{17.6094} = 1.0736$$

$$b_{xy} = \frac{\text{cov}(x,y)}{\sigma_y^2} = \frac{18.9063}{21.6875} = 0.8718$$

Regression line of Y on X: $y - \bar{y} = b_{yx}(x - \bar{x})$

$$y - 10.25 = 1.0736(X - 7.875)$$

$$Y = 1.0736\,X + 1.7954$$

i) Estimate of y for $x=6$ can be obtained by substituting $x=6$ in above regression eqn.

$$Y = 1.0736 \times 6 + 1.7954 = 8.237$$

Regression line of X on Y:

$$X - \bar{X} = b_{xy}(Y - \bar{Y})$$

$$X - 7.875 = 0.8718(Y - 10.25)$$

$$X = 0.8718\,Y - 1.06095$$

ii) Estimate of x can be obtained by substituting $Y=20$ in above eqn

$$X = 16.37505.$$

2) Find the lines of regression for follo. data:

x	10	14	19	26	30	34	39
y	12	16	18	26	29	35	38.

& estimate y for $x = 14.5$ & x for $y = 29.5$

x	y	$u = x-26$	$v = y-26$	u^2	v^2	uv
10	12	-16	-14	256	196	224
14	16	-12	-10	144	100	120
19	18	-7	-8	49	64	56
26	26	0	0	0	0	0
30	29	4	3	16	9	12
34	35	8	9	64	81	72
39	38	13	12	169	144	156
		$\Sigma u = -10$	$\Sigma v = -8$	$\Sigma u^2 = 698$	$\Sigma v^2 = 594$	$\Sigma uv = 640$

Here $n = 7$, $\bar{u} = \frac{-10}{7} = -1.429$, $\bar{v} = \frac{-8}{7} = -1.143$

$\bar{u}^2 = 2.042$, $\bar{v}^2 = 1.306$

$$\mathrm{cov}(u,v) = \frac{1}{n}\Sigma uv - \bar{u}\bar{v}$$

$$= \frac{1}{7}(640) - (1.429)(1.143) = 89.795$$

$$\sigma_u^2 = \frac{1}{n}\Sigma u_i^2 - \bar{u}^2 = \frac{1}{7}(698) - 2.042 = 97.672$$

$$\sigma_u = 9.883$$

$$\sigma_v^2 = \frac{1}{n}\Sigma v_i^2 - \bar{v}^2 = \frac{1}{7}(594) - 1.306 = 83.551$$

$$\sigma_v = 9.14$$

$$r = r(x,y) = r(u,v) = \frac{\mathrm{cov}(u,v)}{\sigma_u \sigma_v} = 0.9941$$

$$r\frac{\sigma_y}{\sigma_x} = r\frac{\sigma_v}{\sigma_u} = 0.9194$$

$$r\frac{\sigma_x}{\sigma_y} = r\frac{\sigma_u}{\sigma_v} = 1.0749$$

$$\bar{x} = a + \bar{u} = 26 - 1.429 = 24.571$$

$$\bar{y} = b + \bar{v} = 26 - 1.143 = 24.857$$

Regression line of y on x is

$$y - \bar{y} = b_{yx}(x - \bar{x})$$

$$y - 24.857 = 0.9194(x - 24.571) \quad \text{— (i)}$$

Regression line x on y is

$$x - 24.571 = 1.0749(y - 24.857) \quad \text{— (ii)}$$

To estimate y for $x = 14.5$

put $x = 14.5$ in (i) gives $y = 15.5977$

Estimate of x for $y = 29.5$ is obtained from (ii)

$$x = 24.571 + 1.0749(29.5 - 24.857)$$

$$x = 29.56176$$

University Question paper sol^n on Binomial Distribution

(Probability)

1. An average box containing 10 articles is likely to have 2 defectives. If we consider a consignment of 100 boxes, how many of them are expected to have three or less defective. (May 2016)

→ $n = 10, \quad p = \frac{2}{10} = 0.2, \quad q = 0.8$

p(three or less defective)

$= p(r \leq 3)$

$= p(r=0) + p(r=1) + p(r=2) + p(r=3)$

$= 10C_0 (0.2)^0 (0.8)^{10} + 10C_1 (0.2)^1 (0.8)^9$

$+ 10C_2 (0.2)^2 (0.8)^8 + 10C_3 (0.2)^3 (0.8)^7$

$= 0.1074 + 0.2684 + 0.302 + 0.2013$

$= 0.8791$

for consignment of 100 boxes, the probability to have 3 or less defective is

$100 \times p(r \leq 3) = 87.9126$

2. If mean & variance of a binomial distribution are 4 & 2 respectively. find probability of :

i) exactly 2 successes &

ii) less than 2 successes. (Dec 16)

→ mean = $np = 4$

var. = $npq = 2$

$\therefore q = \frac{2}{4} = \frac{1}{2}$

$\Rightarrow p = \frac{1}{2}$

$\therefore n = 8$

p(exactly 2 successes)

$= p(r=2)$

$= {}^8C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^6$

$= 0.1094$

ii) p(less than 2 successes)

$= p(r < 2)$

$= p(r=0) + p(r=1)$

$= {}^8C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^8 + {}^8C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^7$

$= 0.0039 + 0.03125$

$= 0.0351$

3) If mean & variance of a binomial distn are 12 & 3 resp. find $P(r \geq 1)$

(Dec 17)

$np = 12$, $npq = 3$

$q = \frac{3}{12} = \frac{1}{4}$

$\therefore p = 1 - \frac{1}{4} = \frac{3}{4}$

$n = 12 \times \frac{4}{3} = 16$

$n = 16$

$$p(r \geq 1) = 1 - p(r<1)$$

$$= 1 - p(r=0)$$

$$= 1 - 16C_0 \left(\frac{3}{4}\right)^0 \left(\frac{1}{4}\right)^{16}$$

$$= 0.9999$$

$$= 1$$

Poisson Distribution:-

Poisson distribution is a particular limiting form of Binomial distribution.

$$P(r) = \frac{z^r}{r!} e^{-z} \quad , r = 1, 2, 3, 4, \ldots$$

is called Poisson Distn with discrete random variate.

When no. of trials goes on increasing & p (prob. of success) tends to zero such that mean is constant (finite).

$\therefore np = z$

Mean = $z = np$

Variance $(\sigma^2) = z = np$

S.D. $(\sigma) = \sqrt{np}$

1. <u>Examples:-</u>

1. If the prob. that concrete cube fails is 0.001. Determine prob. that out of 1000 cubes ① exactly two ⓘⓘ more than one cubes will fail.

→ $p = 0.001$, $n = 1000$

mean $(z) = np = 0.001 \times 1000 = 1$

By Poisson's distⁿ,

$$P(r) = \frac{e^{-z} z^{r}}{r!}$$

ⓘ Exactly two cubes will fail:

$$P(r=2) = \frac{(1)^2 e^{-1}}{2!} = 1.3591$$

ⓘⓘ More than one cube will fail:

$$P(r>1) = 1 - p(r=0,1) = 1 - [p(r=0) + p(r=1)]$$

$$= 1 - \left[\frac{e^{-1}(1)^0}{0!} + \frac{e^{-1}(1)^1}{1!}\right]$$

$$= 0.2642$$

2) An insurance company found that only 0.01 % of the population is involved in certain type of accident each year. If its 1000 policy holders were randomly selected from population. What is prob. that not more than two of its clients are involved in such an accident next year? (Given: $e^{-0.1} = 0.9048$)

→ Here $p = \frac{0.01}{100} = 10^{-4}$, $n = 1000$

$$z = np = 0.1$$

$$p(r) = \frac{e^{-z} z^r}{r!} = \frac{e^{-0.1} (0.1)^r}{r!}$$

$$p(\text{not more than } 2) = p(r=0) + p(r=1) + p(r=2)$$

$$= e^{-0.1}\left[\frac{(0.1)^0}{0!} + \frac{(0.1)^1}{1!} + \frac{(0.1)^2}{2!}\right]$$

$$= 0.9998$$

3. Assume that prob. of an individual coal miner being died in a mine accident during a year is $\frac{1}{2400}$. Calculate prob. that in mine employing 200 miners, there will be at least one will died by accident in a year.

→ $p = \frac{1}{2400}$, $n = 200$

$z = np = 0.0833$

By Poisson's dist^n,

$P(r \geq 1) = 1 - P(r = 0)$

$= 1 - \left[e^{-0.0833} \cdot \frac{(0.0833)^0}{0!} \right]$

$= 0.07995$

4. In a random variable has a Poisson dist^n such that $P(1) = P(2)$. Find

(i) Mean of dist^n

(ii) $P(4)$

$\rightarrow$ $$P(r) = \frac{e^{-z} z^{r}}{r!}$$

$$P(r=1) = P(r=2)$$

$$\frac{e^{-z} z}{1!} = \frac{e^{-z} z^{2}}{2!}$$

$$\Rightarrow \frac{z}{2} = 1$$

$$\Rightarrow z = 2$$

$\therefore$ mean $= 2$

$$P(r=4) = \frac{e^{-z} z^{4}}{4!} = 0.09$$

4) If the prob. that an individual suffers a bad reaction from a certain injection is 0.001, determine the prob. that out of 2000 individual, more than 2 individuals, will suffer a bad reaction.

$\rightarrow$ $p = 0.001$, $n = 2000$

$$z = np = 2$$

$$p(r) = \frac{e^{-z} z^r}{r!}$$

The prob. ~~of~~ more than 2 individuals

$= p(r > 2)$

$= 1 - p(r \leq 2)$

$= 1 - [p(r=0) + p(r=1) + p(r=2)]$

$= 0.3233$

5. In a Poisson distⁿ, if $p(r=1) = 2\,p(r=2)$ find $p(r=3)$

→ $$p(r) = \frac{e^{-z} z^r}{r!}$$

$p(r=1) = 2\,p(r=2)$

$$\therefore \frac{e^{-z} z^1}{1!} = 2\,\frac{e^{-z} z^2}{2!}$$

$\Rightarrow z = 1$

$$\therefore p(r=3) = \frac{z^3 e^{-z}}{3!} = \frac{e^{-1}\, 1^3}{3!} = 0.0613$$

University question paper sol^n on

Poisson Distribution

1. In a certain factory turning out razor blades, there is a small chance of 0.002 for any blade to be defective. The blades are supplied in packets of 10. Use Poisson's dist^n to calculate approximate no. of packets containing no defective & none defective resp. in a consignment of 1000 packets.

(May 17)

→ $p = 0.02$, $n = 10$

$\Rightarrow z = np = 0.2$

P(no defective & one defective)

$= P(r \le 1)$

$= P(r=0) + P(r=1)$

$= \dfrac{e^{-0.2}(0.2)^0}{0!} + \dfrac{e^{-0.2}(0.2)^1}{1!} = 0.9825$

In consignment of 1000 packets, the approx no. of packets containing no defective & one defective is

$1000 \times p(r \leq 1)$

$= 1000 \times 0.9825$

$= 982.5$

Normal Distribution :-

Let p be the prob. of success

q be the prob. of failure

If prob. of success is very small as compare to n the no. of trials which is very large then use normal dist^n.

And is given by

$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

where, μ: mean

x: variable assumes all values from $-\infty$ to ∞.

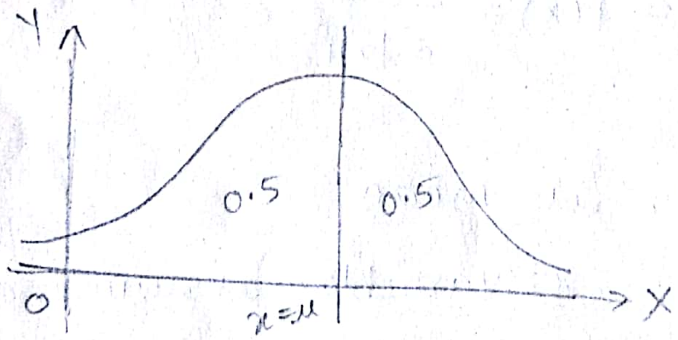
σ: S.D. > 0

$f(x)$: normal density function.

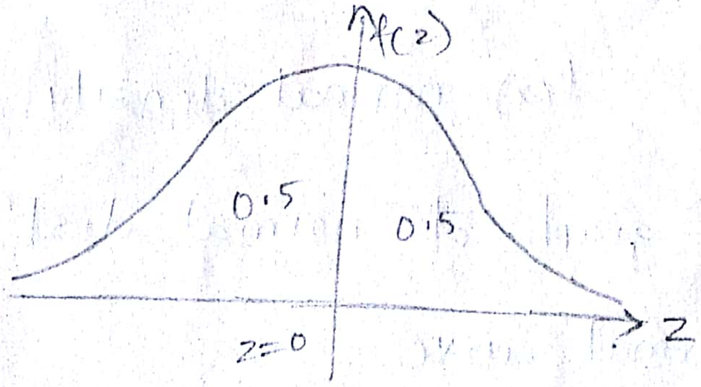
The graph of normal dist^n is called normal curve.

Properties of normal distribution:-

1. The mean, mode & median all are equal
2. The curve is symmetric about the mean μ.
3. Exactly half of the values are to the left of the centre & exactly half values are to the right.
4. The total area under the curve is 1.



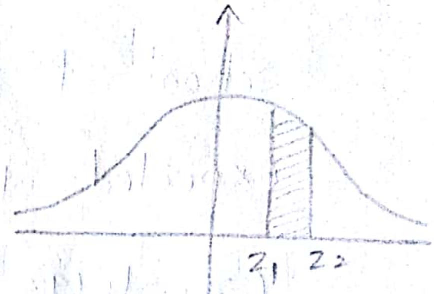
$\Downarrow$



let $z = \dfrac{x-\mu}{\sigma}$ be the normal variable

1. $A(z = z_1) = A(z=0 \text{ to } z = z_1)$

2. $P(0<z<z_1) = A(z=0 \text{ to } z=z_1)$

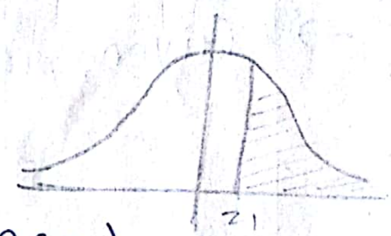


$A(z_2) - A(z_1)$

3. $P(z_1 < z < z_2) = A(z_2) - A(z_1)$

$= A(0<z<z_2)$

$- A(0<z<z_1)$



$A(z>z_1)$

4. $P(z>z_1) = 0.5 - A(z_1)$

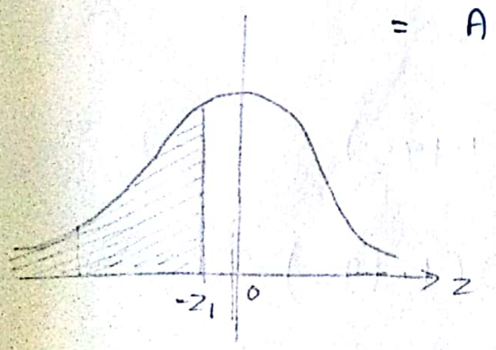
5. $P(z<-z_1) = 0.5 - A(-z_1) = 0.5 - A(z_1)$

6. $P(-z_2 < z < -z_1) = A(-z_2) - A(-z_1)$

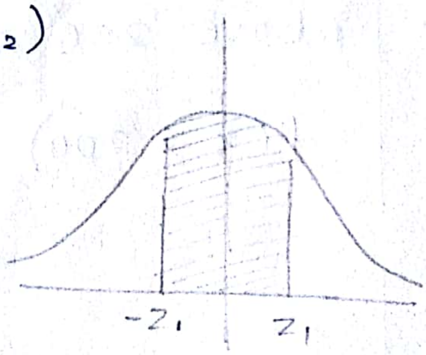
$= A(z_2) - A(z_1)$

7. $P(-z_1 < z < z_2) = A(-z_1) + A(z_2)$

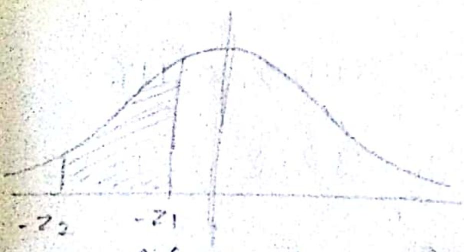
$= A(z_1) + A(z_2)$



$P(z<-z_1)$



$A(-z_1 < z < z_1)$



$A(-z_2 < z < -z_1)$

Examples :-

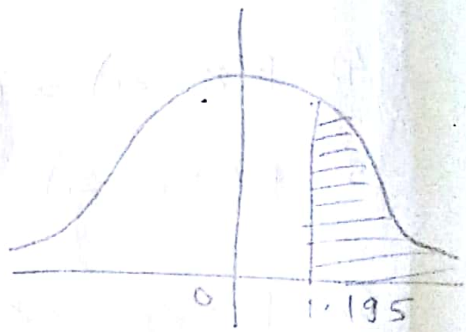
1. Heights of students follow normal distⁿ with mean 190 cm & var. 70 cm^2. In a school of 800 students how many are expected to have heights above 200 cm

($z = 1.195$, $A = 0.3850$)

→ Given, mean $= \mu = 190$

$\sigma^2 = 70 \Rightarrow \sigma = \sqrt{70}$

$x = 200$



Let $z = \dfrac{x-\mu}{\sigma} = \dfrac{200-190}{\sqrt{70}} = 1.195$

The prob. of expected to have height above 200 cm is

$P(x > 200) = P(z > 1.195)$

$= 0.5 - A(1.195)$

$= 0.5 - 0.3840 = 0.116$

Expected no. of students $= 800 \times 0.116$

$= 92.8 = 93$ students.

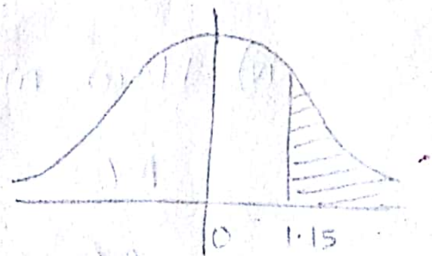
2. Assume the mean height of soldiers to be 68.22 inches with the variance of 10.81 inches (square). How many soldiers in a regiment of 10,000 would you expect to be over 6 feet tall. $A(z=1.15)=0.3749$

→ $\mu = 68.22$, $\sigma^2 = 10.8 \Rightarrow \sigma = \sqrt{10.8} = 3.28$

$x = 6$ feet $= 6 \times 12 = 72$ inches

$$z = \frac{x-\mu}{\sigma} = \frac{72-68.22}{3.28} = 1.15$$

$P(x > 72) = P\left(\frac{x-\mu}{\sigma} > 1.15\right)$



$= P(z > 1.15)$

$= 0.5 - P(0 < z < 1.15)$

$= 0.5 - A(z = 1.15)$

$= 0.5 - 0.3749$

$= 0.1251$

∴ Expected no. of soldiers above 6 feet are

$10,000 \times 0.1251 = 1251$

3) In a sample of 1000 cases, the mean of certain test is 14 & S.D. is 25. Assuming the distribution is normal. Find:

i) How many students score between 12 & 15.

Given: $A(z = 0.08) = 0.0319$

$A(z = 0.04) = 0.0160$

ii) How many scores above 18

Given $A(z = 0.16) = 0.0636$

iii) How many scores below 8

Given $A(z = 0.24) = 0.0948$

iv) How many scores exactly 16.

$A(z = 0.06) = 0.0239$

$A(z = 0.1) = 0.0398$

→ $\mu = 14, \quad \sigma = 25$

$$z = \frac{x - \mu}{\sigma}$$

i. Score betⁿ 12 & 15

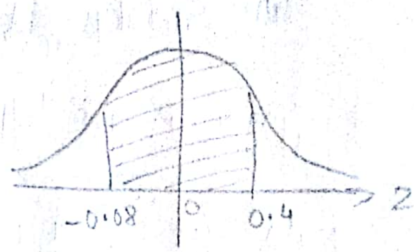
$x_1 = 12, \quad x_2 = 15$

$$z_1 = \frac{x_1 - \mu}{\sigma} = -0.08, \quad z_2 = \frac{x_2 - \mu}{\sigma} = 0.04$$

prob. of students scores betⁿ 12 & 15

$= P(12 < x < 15) = P(-0.08 < z < 0.4)$

$= P(-0.08 < z < 0) + P(0 < z < 0.4)$



$= A(z = -0.08) + A(z = 0.4)$

$= A(z = 0.8) + A(z = 0.4)$ ... (Due to symmetry)

$= 0.0319 + 0.0160$

$= 0.0479$

$\therefore$ No. of students having score betⁿ 12 & 15

$= 1000 \times 0.0479 \approx 48$

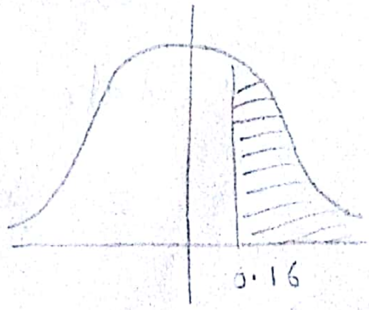
ii) Score above 18 :

$x = 18,$

$z = \dfrac{x - \mu}{\sigma} = 0.16$

Prob. who score above 18 is

$P(x > 18) = P(z > 0.16)$



$= 0.5 - A(z = 0.16)$

$= 0.5 - 0.0636 = 0.4364$

$\therefore$ no. of students score above 18 is

$1000 \times 0.4364 \approx 437$

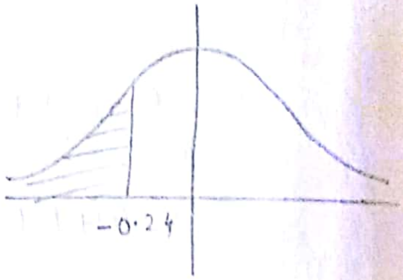
iii) Scores below 8 :

$x = 8$

$$z = \frac{x-\mu}{\sigma} = -0.24$$

Prob. who scores below 8 is

$P(x<8) = P(z<-0.24)$



$= 0.5 - A(z=-0.24)$

$= 0.5 - A(z=0.24)$ ... Due to symmetry

$= 0.5 - 0.0948 = 0.4052$

$\therefore$ no. of students scoring below 8 is

$1000 \times 0.4052 \approx 406$

iv) Scores exactly 16 marks :

($\because$ for exact point, take average of two adjusant points)

take $x_1 = 15.5$, $x_2 = 16.5$

$$z = \frac{x-\mu}{\sigma}, \quad z_1 = \frac{x_1-\mu}{\sigma} = 0.06$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = 0.1$$

Prob. who scores exactly 16 marks

$= P(15.5 < x < 16.5)$

$= P(0.06 < z < 0.1)$

$= A(z = 0.1) - A(z = 0.06)$

$= 0.0398 - 0.0239$

$= 0.0159$

$\therefore$ No. of students who scores exactly 16 marks

$= 1000 \times 0.0159 \approx 16$