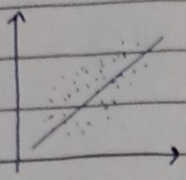
Date——
Page——

A

Regression lines:

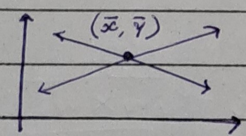


Y on X -

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

X on Y -

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



Regression coefficient:

$$b_{yx} = r \cdot \frac{\sigma_y}{\sigma_x} = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}$$

$\sigma \rightarrow$ standard deviation

$r \rightarrow$ correlation coefficient

$$b_{xy} = r \cdot \frac{\sigma_x}{\sigma_y} = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum y^2 - (\sum y)^2}$$

$$\boxed{-1 \le r \le 1}$$

$$\boxed{r = \pm\sqrt{b_{xy} \cdot b_{yx}}}$$

$$\begin{cases} b_{xy} > 0, \; b_{yx} > 0 \Rightarrow r > 0 \\ b_{xy} < 0, \; b_{yx} < 0 \Rightarrow r < 0 \end{cases}$$

- find Σ of data
- find b_{xy}, b_{yx}
- find corr. coeff
- find regression lines

Q. $n = 9$

X	Y	XY	X^2	Y^2
1	9	9	1	81
2	8	16	4	64
3	10	30	9	100
4	12	48	16	144
5	11	55	25	121
6	13	78	36	169
7	14	98	49	196
8	16	128	64	256
9	15	135	81	225
45	108	597	285	1356.

$$b_{yx} = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n\Sigma x^2 - (\Sigma x)^2}$$

$$= \frac{9(597) - 45(108)}{9(285) - (45)^2}$$

$$\boxed{b_{yx} = 0.915}$$

$$b_{xy} = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n(\Sigma y^2) - (\Sigma y)^2}$$

$$= \frac{9(597) - 45(108)}{9(1356) - (108)^2}$$

$$\boxed{b_{xy} = 0.95}$$

corr. coeff $\Rightarrow r = \pm\sqrt{b_{xy} * b_{yx}}$

$= \pm\sqrt{0.95 \times 0.95}$

$$\boxed{r = +0.95}$$

Regression line $\Rightarrow$

Y on X: $(Y - \bar{Y}) = b_{yx}(X - \bar{X})$

$\bar{Y} = \frac{108}{9} = 12$, $\bar{x} = 5$

$(y - 12) = 0.95(x - 5)$

$y = 0.95x - 4.75 + 12$

$$\boxed{Y = 0.95x + 7.25}$$

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

$\bar{x} = 5$, $\bar{Y} = 12$

$(x - 5) = 0.95(y - 12)$

$$\boxed{x = 0.95y - 6.4}$$

$\boxed{-1 \le r \le 1}$

Date
Page

Q. Find mean values of X & Y.
Find correlation coeff, r.
Find s.d. of Y.

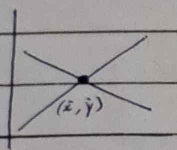
$var(x) = 9 = \sigma_x^2$

regression eqⁿ : $8x - 10y + 66 = 0$ —— ①

$40x - 18y - 214 = 0$ —— ②

⇒ s.d. of X $= \sigma_x = 3$

(i)



$8\bar{x} - 10\bar{y} + 66 = 0$ $\quad$ $40\bar{x} - 18\bar{y} - 214 = 0$

$4\bar{x} - 5\bar{y} + 33 = 0$ $\quad$ $20\bar{x} - 9\bar{y} - 107 = 0$

$$20\bar{x} - 25\bar{y} + 165 = 0$$
$$-\ 20\bar{x} - 9\bar{y} - 107 = 0$$
$$-16\bar{y} = -272$$
$$\bar{y} = 136/8 = \boxed{17 = \bar{y}}$$

$$4\bar{x} - 5\times17 + 33 = 0$$
$$4\bar{x} = 52$$
$$\boxed{\bar{x} = 13}$$

(ii) let eqⁿ ① be Y on X, $y = \frac{-66 - 8x}{-10} = \frac{33 + 4x}{5} = \frac{33}{5} + \overset{b_{yx}}{\frac{4x}{5}}$

let eqⁿ ② be X on Y, $x = \frac{214 + 18y}{40} = \frac{107 + 9y}{20} = \frac{107}{20} + \overset{b_{xy}}{\frac{9y}{20}}$

correlation coeff, $r = \pm\sqrt{b_{xy} \cdot b_{yx}}$

$= +\sqrt{\frac{4}{5} \cdot \frac{9}{20}} = \sqrt{\frac{36}{100}} = \frac{6}{10} = 0.6$

(iii) s.d. of Y, σ_y $\quad b_{yx} = r \frac{\sigma_y}{\sigma_x}$

$$\frac{4}{5} = 0.6 \times \frac{\sigma_y}{3}$$

$$\sigma_y = \frac{4}{\cancel{5}} \times \frac{3}{\cancel{6}} \times 10 = 4$$

Date
Page

Q: For given lines of regression,

$3x - 2y = 5$ — ①

$x - 4y = 7$ — ②

Find regression coeff & coeff of correlation.

Let ① be X on Y

$x = \boxed{\frac{2}{3}} y + \frac{5}{3}$

Let ② be Y on X

$y = \boxed{\frac{1}{4}} x - \frac{7}{4}$

Regression coeff.

$\Rightarrow b_{yx} \Rightarrow \frac{1}{4}$ — regression coeff of x

$b_{xy} \Rightarrow \frac{2}{3}$ — regression coeff of y

$r = \pm \sqrt{2/3 \cdot 1/4} = \sqrt{1/6}$

$r = \cancel{0.6854}\ 0.4082$

Q: Following data about sales expenditure of a firm is given (in crores)

	sales (x)	ad expenditure (y)
mean	40	6
σ	10	~~1.5~~ 15

Coeff of correlation is 0.9

(i) find regression lines

(ii) What is add expenditure if firm proposes sales target of ₹60 crore

$x = 60$ crore (independent)

$y \Rightarrow$ (dependent)

$r = 0.9$

$\bar{x} = 40$, $\sigma_x = 10$, $\bar{y} = 6$, $\sigma_y = \cancel{1.5}\ 1.5$

~~$b_{xy} = r\left(\frac{\sigma_x}{\sigma_y}\right) = 0.9 \times \frac{10}{1.5} = \frac{9 \times 10^2}{15} = 6$~~

$b_{yx} = r\frac{\sigma_y}{\sigma_x} = 0.9 \times \frac{15}{10} = \cancel{\frac{9 \times 15}{1000} = \frac{144}{1000} = 0.144}$

$= \frac{9 \times 15}{100} = \frac{135}{1000} = 0.135$

$\begin{array}{r} \overset{2\,3}{135} \\ \times\ 6 \\ \hline 810 \end{array}$ $\begin{array}{r} \overset{1\,2}{135} \\ \times\ 4 \\ \hline 540 \end{array}$ $\begin{array}{r} \overset{2\,2}{144} \\ \times\ 6 \\ \hline 8.64 \\ 0.24 \\ \hline 8.40 \end{array}$

Date
Page

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

$\cancel{(y-6) = 0.144(x-40)}$

$\cancel{y-6 = 0.144x - 5.76}$

$\cancel{y = 0.144x - 0.24}$

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

$\cancel{(x-40) =}$

$\cancel{(y-\bar{y}) = 1.35(x-\bar{x})}$

$(y-\bar{y}) = 0.135\,(x-\bar{x})$

$(y-6) = 0.135\,(x-40)$

$y-6 = 0.135x - 5.4$

$y = 0.135x + 0.6$

$y = 8.10 + 0.6 \qquad (x = 60)$

$y = 8.7$

Probability distribution

discrete | continuous

Date
Page

Bernoulli distribution :

- discrete probability distribution
- models single binary expt with 2 possible outcomes (either 0 or 1, success or failure, true or false)

PROBABILITY MASS FUNCTION OF BERNOULLI :

$$\begin{cases} p, & x=1 \quad \text{(success)} \\ q=1-p, & x=0 \quad \text{(failure)} \end{cases}$$

$p \rightarrow$ probability of success

$q \rightarrow$ probability of failure

mean & variance -

μ = mean $= p$

$\sigma^2 = pq$

Binomial distribution :

- discrete probability distribution
- gives only 2 possible results in an expt — success or failure.
- ~~too~~ there are two parameters — n, p

$n \rightarrow$ number of trial

$p \rightarrow$ probability of success

$$\boxed{q = 1-p}$$

Properties -

- there are two possible outcomes — success or failure, true or false
- there is "n" number of independent trials
- probability of success or failure remains same for each trial
- only number of success is calculated out of n number of trial
- every trial is an ~~St~~ independent trial which means

No of trial $\begin{cases} 1, & \text{Bernoulli} \\ >1, & \text{Binomial} \end{cases}$

Date
Page

outcomes of one trial doesn't affect outcome of other trial.

mean & variance -

mean = μ = np

variance = σ^2 = npq

PROBABILITY MASS FUNCTION OF BINOMIAL:

$$P(x, n, p) = {}^{n}C_{x}\, p^{x}\, q^{n-x}$$

$$P(x,n,p) = \frac{n!}{x!(n-x)!} \cdot p^{x} \cdot (1-p)^{n-x}$$

Q. If a coin is tossed 5 times, find probab of -

Binomial

(i) getting exactly two heads

(ii) atleast four heads

(iii) atmost two heads

(i) $n = 5, \ x = 2, \ p = q = 1/2$

$$P(x,n,p) = \frac{5!}{2!\ 3!} \cdot \left(\tfrac{1}{2}\right)^2 \left(\tfrac{1}{2}\right)^3$$

$$= \frac{5 \times 4}{2} \cdot \left(\tfrac{1}{2}\right)^5$$

$$= 10 \cdot \frac{1}{32} = 5/16$$

R

Date
Page

(ii) 4 heads $n = 5, \quad x = 4, \quad p = q = 1/2$

$$P(n, x, p) = \frac{5!}{4!\,1!} \cdot (1/2)^4 (1/2)^1$$

$$= 5 \cdot 1/32 = 5/32$$

5 heads $n = 5, \quad x = 5, \quad p = q = 1/2$

$$P(n, x, p) = 1 \cdot (1/2)^5 = 1/32$$

≥ 4 heads $\therefore P(X \geq 4) = P(X = 4) + P(X = 5)$

req. $P(n, x, p) = 5/32 + 1/32 = 6/32 = 3/16$

(iii) 1 head $n = 5, \quad x = 1, \quad p = q = 1/2$

$$P(n, x, p) = \frac{5!}{1!\,4!} \cdot (1/2)^1 (1/2)^4$$

$$= 5 \cdot 1/32 = 5/32$$

2 heads $n = 5, \quad x = 2, \quad p = q = 1/2$

$$P(n, x, p) = \frac{5!}{3!\,2!} \cdot (1/2)^2 (1/2)^3$$

$$= \frac{5 \times 4}{2} \cdot 1/32 = 10/32$$

0 head $n = 5, \quad x = 0, \quad p = q = 1/2$

$$P(n, x, p) = \frac{5!}{0!\,5!} (1/2)^5 (1/2)^0$$

$$= 1/32$$

≤ 2 head $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$

$$= 1/32 + 5/32 + 10/32 = 16/32 = 1/2$$

Scanned with AnyScanner

Q: Coin tossed 10 times. Probab of getting exact 6 heads or atleast 6 heads

#

6 heads exact

$n=10, \quad x=6, \quad n=p=1/2$

$P(n,x,p) = \frac{10!}{6!\,4!} \times \left(\frac{1}{2}\right)^6 \left(\frac{1}{2}\right)^4$

$= \frac{10 \times 9 \times 8 \times 7}{4 \times 3 \times 2} \times \left(\frac{1}{2}\right)^{10}$

$= 210 \cdot \left(\frac{1}{2}\right)^{10}$

$= \frac{210}{1024} = \frac{105}{512}$

atleast 6 heads

$n=10, \quad x=7, \quad n=p=1/2$

$P(n,x,p) = \frac{10!}{7!\,3!} \left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^3$

$= \frac{10 \times 8 \times 9}{6} \cdot \frac{1}{2^{10}}$

$= \frac{120}{2^{10}}$

$n=10, \quad x=8, \quad n=p=1/2$

$P(n,x,p) = \frac{10!}{8!\,2!} \times \frac{1}{2^{10}}$

$= \frac{10 \times 9}{2} \times \frac{1}{2^{10}}$

$= 45 \times \frac{1}{2^{10}}$

$n=10, \quad x=9, \quad n=p=1/2$

$P(n,x,p) = \frac{10!}{9!\,1!} \left(\frac{1}{2^{10}}\right)$

$= 10 \cdot \frac{1}{2^{10}}$

$x=10, \quad x=10, \quad n=p=1/2$

$P(n,x,p) = \frac{10!}{10!\,0!} \left(\frac{1}{2}\right)^{10}$

$= \frac{1}{2^{10}}$

$P(x \geq 6) = P(x=6) + P(x=7) + P(x=8) + P(x=9) + P(x=10)$

$\frac{210 + 120 + 45 + 10 + 1}{2^{10}} = \frac{386}{1024} = \frac{193}{512}$

Poisson Distribution:

* used for calc possibilities for an event with the avg rate of values.
* discrete probability distribution — means random variable take only specific values in a given list of number, probably ∞.
* measures how many times ~~how many times~~ an event is likely to occur within 'x' (period of time).
* it's basically time limiting process of binomial distribution.
* used under certain condition —
  1. number of trials, n tends to ∞.
  2. probability of success, tends to 0.
  3. $\boxed{np = \lambda}$ (λ is finite & constant)

In poisson distribution,

$\mu \Rightarrow E(x) = \lambda$

$\sigma^2 = E(x) = \lambda$

Poisson distribution formula:

$$\boxed{P(X = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}} \quad , \quad \lambda \rightarrow \text{avg rate of values}$$

Q: Random variable X has a poisson distribution with parameter λ, s.t. $P(X=1) = 0.2\,P(X=2)$

Find $P(X=0)$

Using Poisson Distribution,

$$P(X=x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$P(X=1) = 0.2\,P(X=2)$$

$$\frac{e^{-\lambda}\cdot\lambda^{1}}{1!} = 0.2 \times \frac{e^{-\lambda}\lambda^{2}}{2!}$$

$$\frac{\lambda}{1} = 0.2 \times \frac{\lambda^{2}}{2}$$

$$\lambda = 10$$

$$P(X=0) = \frac{e^{-10}(10)^{0}}{0!} = e^{-10} = (2.718)^{-10}$$

$$= 0.0000453999$$

Q: Determine probability of 4 staff members being ~~adjacent~~ absent on any day with 800 staff members, given that probability of any staff member being absent on any day is 0.001.

$$n = 800 \quad , \quad p = 0.001$$

$$\lambda = np = 0.8$$

$$P(X=4) = \frac{e^{-0.8}(-0.8)^{4}}{4!}$$

$$= \cancel{\frac{2.2255409285 \times 0.4096}{24}} = \frac{0.4493289641 \times 0.4096}{24}$$

$$= \cancel{0.0379825652} = 0.0076685477$$

Date
Page

Q. An avg of 0.61 soldiers died by horse kicks per year in each Prussian army corps. Calc the probability that exactly two soldiers died in VII Army Corps in 1898. Assume that horse kick deaths follows a poisson distributio.

$\lambda = 0.61$

$X = 2$

$$P(X=2) = \frac{e^{-0.61}(0.61)^2}{2}$$

$$= \frac{0.5433508691 \times 0.3721}{2}$$

$$= \frac{0.2021808584}{2}$$

$$= 0.1010904292$$

Q. If the probab of bad rxn is 0.002, determine chance that out of 1000 persons more than 3 will suffer bad rxn.

$p = 0.002$ $\quad np = 1000(0.002) = 2$

$X = 0, 1, 2, 3$

$$P(X=0) = \frac{e^{-0.002} \times (0.002)^0}{0!} = 0.998001997$$

$$P(X=1) = \frac{e^{-0.002} \times (0.002)^1}{1!} = 0.001996004$$

$$P(X=2) = \frac{e^{-0.002} \times (0.002)^2}{2} = 0.000001996$$

$$P(X=3) = \frac{e^{-0.002} \times (0.002)^3}{6} =$$

Scanned with AnyScanner

$$P(X=x) = \frac{e^{-\lambda} \cdot \lambda^{x}}{x!}$$

$$P(X=0) = \frac{e^{-2} \cdot (+2)^{0}}{0!} = 0.1353352832$$

$$P(X=1) = \frac{e^{-2} \cdot (+2)^{1}}{1!} = 0.2706705665$$

$$P(X=2) = \frac{e^{-2} \cdot (2)^{2}}{2!} = 0.2706705665$$

$$P(X=3) = \frac{e^{-2} \cdot (2)^{3}}{3!} = 0.1804470443$$

$$P(X>4) = 1 - \left(P(X=0) + P(X=1) + P(X=2) + P(X=3)\right)$$

$$= 0.1428765395$$

Q3: If 1% of total machines made are defective, find probab that <3 machines are defective in a sample of 100 machines.

$n = 100$ $\qquad$ $\lambda = np = 1$

$p = 0.01$

$$P(X=0) = \frac{e^{-1} \cdot 1^{0}}{0!} = e^{-1}$$

$$P(X=1) = \frac{e^{-1} \cdot 1^{1}}{1} = e^{-1}$$

$$P(X=2) = \frac{e^{-1} \cdot 1^{2}}{2} = \frac{e^{-1}}{2}$$

$$P(X<3) = e^{-1} + e^{-1} + \frac{e^{-1}}{2}$$

$$= \frac{5}{2} e^{-1} = \frac{5}{2}(0.3678794412)$$

$$= 0.919698603$$

R

CONTINUOUS PROBABILITY DISTRIBUTION

- fundamental concepts in probab theory that deals with random variables taking values within a continuous range.

Uniform distribution -

- a continuous probab distribution that assigns equal probab to all outcomes within a specified interval
- a random variable 'X' is said to have continuous probability distribution over an interval (a, b) if its probability density function is constant, say 'k', over entire range, i.e.,

$$\begin{cases} k, & a < x < b \\ 0, & \text{otherwise} \end{cases}$$

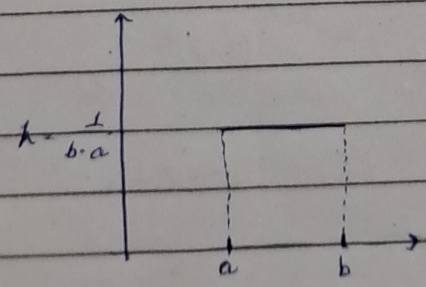
since total probability is always $= 1$,

we have,

$$\int_a^b f(x)\,dx = k\int_a^b dx = 1$$

$$\boxed{k = \frac{1}{b-a} = f(x)}$$

$$\therefore f(x) = \begin{cases} \dfrac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases}$$



mean = median $= \frac{a+b}{2}$, $var(x) = \frac{(b-a)^2}{12}$, s.d. $= \sqrt{var(x)}$

Date
Page

$$P(a<x<b) = \frac{d-c}{b-a} = \int_a^b f(x)\,dx$$

"conditional probability"

Q. If X is uniformly distributed in $(-1,4)$, find mean, var(x), s.d. & median.

$$\text{mean} = \text{median} = \frac{-1+4}{2} = 1.5$$

$$\cancel{s.d.}\ \text{var}(x) = \frac{(4+1)^2}{12} = \frac{25}{12} = 2.08$$

$$\text{s.d.} = \sqrt{\text{var}(x)} = \sqrt{2.08}$$

Q. Using uniform distribution probab density func for random variable X in interval $(0,20)$, find probab for range $(3<x<16)$.

$$P(3<x<16) = \frac{d-c}{b-a} = \frac{16-3}{20-0} = 13/20$$

Q. If x is uniformly distributed, mean $=1$, var $= 4/3$. find $P(X<0)$.

$$\frac{a+b}{2} = 1 \qquad \frac{(b-a)^2}{12} = 4/3$$

$a+b=2 \qquad b-a = \pm 4$

$a+b=2$
$b-a=4$
$2b = 6$
$b = 3$
$\cancel{a=-1}$ $b=3$

$a+b=2$
$b-a=-4$
$2a=6$
$a=3$
$b=-1$

R

Date
Page

~~P(-1<x<0)~~ $P(-1 < x < 3)$

$$k = \frac{1}{b-a} = \frac{1}{3-(-1)} = \frac{1}{4}$$

$$P(x < 0) = \int_{-1}^{0} f(x)\,dx = \frac{1}{4}[x]_{-1}^{0} = \frac{1}{4}[0+1] = \frac{1}{4}$$

R

Exponential distribution -

- continuous probab distribution
- models time b/w events in a Poisson process where events occur independently @ constant rate

Probability density func^n -

$$\boxed{f(x) = \lambda e^{-\lambda x}}, \quad \lambda \geq 0,\ x \geq 0 \qquad (\lambda \rightarrow \text{mean})$$

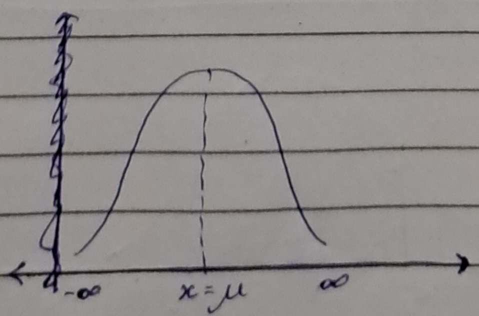
~~mean~~, $\boxed{\mu = \frac{1}{\lambda}}$ ~~variance~~, $\boxed{\sigma^2 = \frac{1}{\lambda^2}}$

Normal distribution - properties, area under curve, formula

- continuous probab distribution
- aka Gaussian distribution
- widely used in statistics for solving complex mathematical problem. It's represented graphically by a bell-shaped curve.
- a random variable 'x' is said to have a normal distribution with parameters $\underset{\text{(mean)}}{\mu}$ and $\underset{\text{(s.d.)}}{\sigma}$, if its

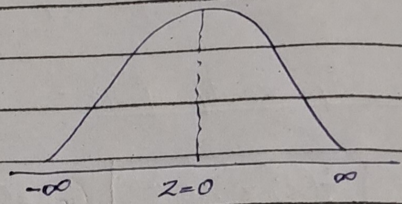
probability density func^n is given by -

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}} \qquad -\infty \leq x, \mu \leq \infty \qquad \sigma \geq 0$$



STANDARD NORMAL VARIATE :

$$\boxed{z = \frac{x-\mu}{\sigma}}$$ for $\mu = 0$, $\sigma = 1$



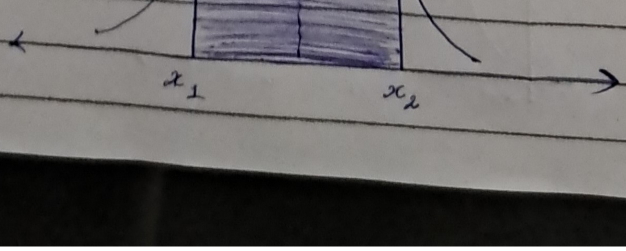
Probability density func^n -

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2}}$$ $-\infty \le x \le \infty$

AREA UNDER NORMAL CURVE :

- Curve of any continuous probab distributed is constructed so that area under curve bounded by two ordinates x_1 & x_2 equals probability that random variable X assumes b/w x_1 & x_2, which is given by -

$$\boxed{P(x_1 \le x \le x_2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx}$$



Scanned with AnyScanner

R

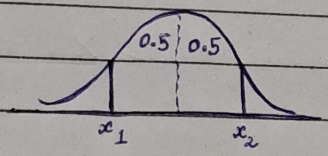
- converting $P(x) \to P(z)$, $\sigma = 1$

$$P(z_1 < z < z_2) = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{1}{2}z^2} \, dz$$

PROPERTIES OF NORMAL DISTRIBUTION :

any 5 or 6. (from LMS)

- area under curve = 1.
- graph is symmetric. Probability of each part is 0.5



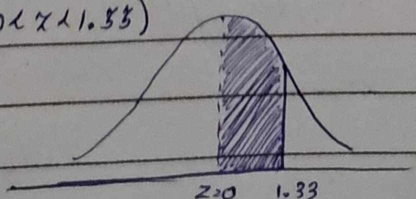
- bell- shaped curve
- mean = median = mode

Q. If X is a random variable with normal distribution having $\mu = 50$, $\sigma = 15$. find the probability b/w $50 < x < 70$.

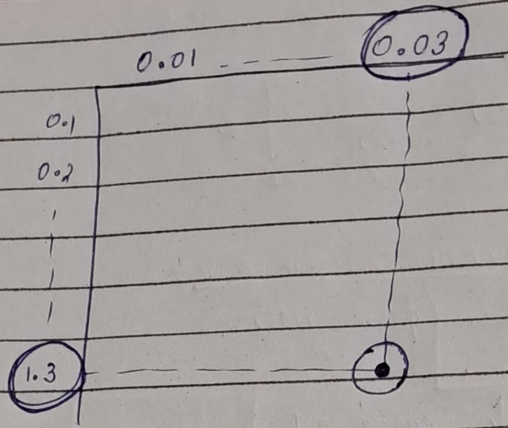
$$z = \frac{x - \mu}{\sigma} = \frac{x - 50}{15}$$

$$\underset{(x=50)}{z_1} = \frac{50-50}{15} = 0 \qquad \underset{(x=70)}{z_2} = \frac{70-50}{15} = \frac{4}{3} = 1.33$$

$P(50 < x < 70) = P(0 < z < 1.33)$



$\boxed{1.33 \rightarrow 1.3 + 0.03}$



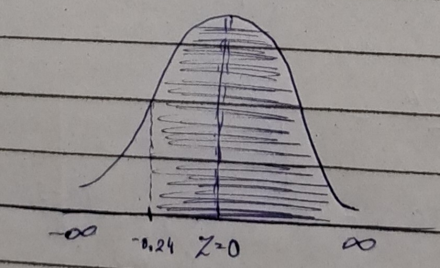
$P(0 < z < 1.33) = .9082$

Q3; $\mu = 527$, $\sigma = 112$ for a random variable X. $P(X > 500)$ w... normally distributed?

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

$$Z_1 = \frac{500 - 527}{112} = \frac{-27}{112} = -0.24$$

$$P(X > 500) = P(Z > -0.24)$$



$= 1 - P(Z < -0.24) = 1 - .4052 = 0.5948$

(or) $= P(Z > 0.24) = 0.5948$