Normal distribution: A random variable X is said to have normal distribution if its density function is given by

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \times e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad ; -\infty < x < \infty,$$

$$-\infty < \mu < \infty,$$

$$\sigma > 0.$$

Note: We take the normal variate

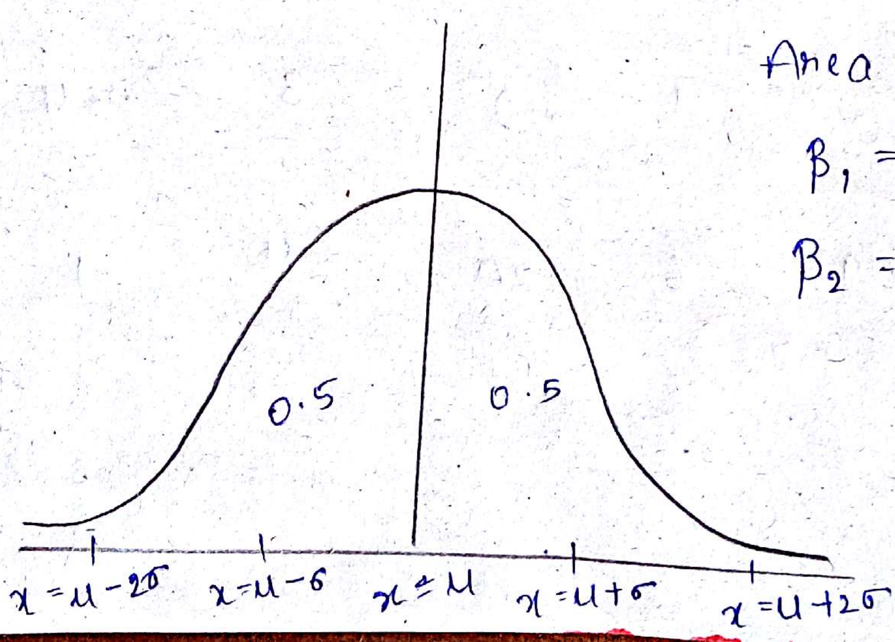
$$z = \frac{x-\mu}{\sigma} \quad ; \quad \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-z^2/2}$$

Normal curve: It is a bell shaped curve symmetrical about $x = \mu$ line. The total (prop.) probability or total area under this curve is 1. Left of $x = \mu$ is 0.5 and right of $x = \mu$ is 0.5

Area = 1

$\beta_1 = 0$

$\beta_2 = 3$



To find probability of normal curve:

step ①: perform the change of scale

$z = \frac{x - \mu}{\sigma}$ and find z_1 & z_2 values

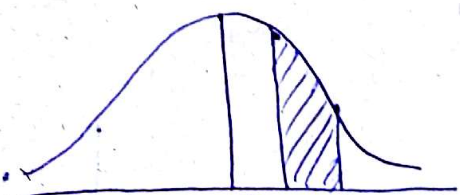
corresponding to x_1 & x_2

step ②:

(a) ~~To find probability of P~~

(a) $P(x_1 \le X \le x_2) = P(z_1 \le z \le z_2)$

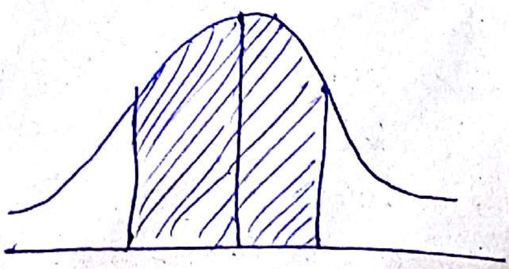
case (i): If both z_1 and z_2 are of same side, i.e. positive side or negative side



$$\boxed{P(z_1 \le z \le z_2) = |A(z_2) - A(z_1)|}$$

case (ii): If both z_1 and z_2 are of opposite side

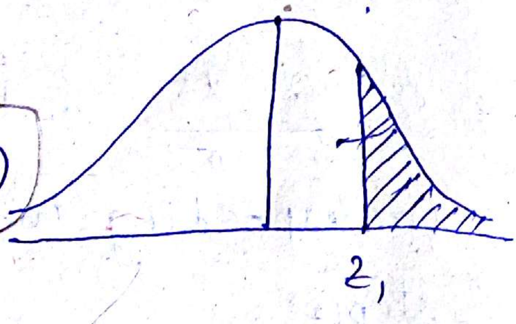
$$\boxed{P(z_1 \le z \le z_2) = A(z_1) + A(z_2)}$$



(b) $p(z > z_1)$

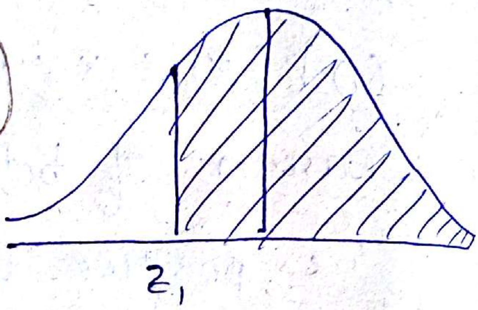
case (i): if $z_1 > 0$

$$\boxed{p(z > z_1) = 0.5 - A(z_1)}$$



case (ii): If $z_1 < 0$

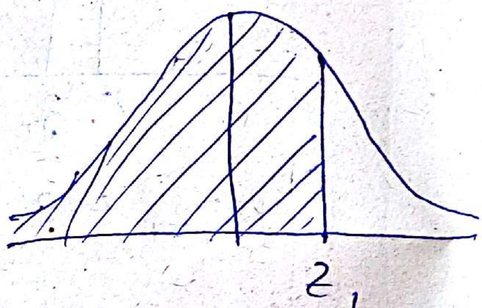
$$\boxed{P(z > z_1) = 0.5 + A(z_1)}$$



(c) $p(z < z_1)$

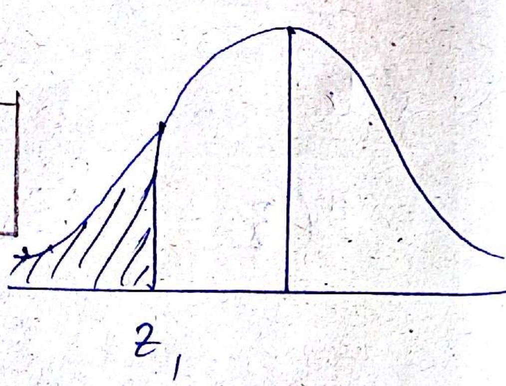
case (i): If $z_1 > 0$

$$\boxed{p(z < z_1) = 0.5 + A(z_1)}$$



case (ii): if $z_1 < 0$

$$\boxed{P(z < z_1) = 0.5 - A(z_1)}$$



31) For a normally distributed variate with mean 1 & standard deviation 3. Find Prob. that i) $3.43 \le x \le 6.19$

ii) $-1.43 \le x \le 6.19$

Sol: $\boxed{\mu = 1 \;;\; \sigma = 3 \text{ . for normal}}$

i) $x_1 = 3.43$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{3.43 - 1}{3} = 0.81$$

$x_2 = 6.19$

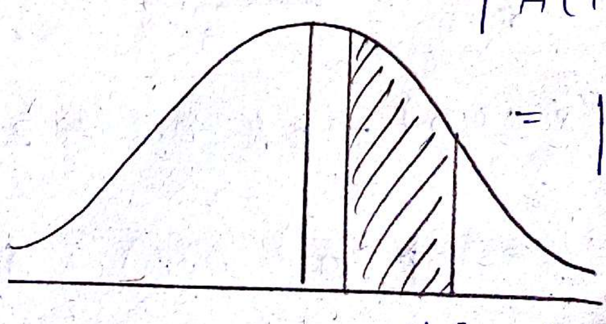
$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{6.19 - 1}{3} = 1.73$$

$$P(3.43 \le x \le 6.19) = P(0.81 \le z \le 1.73)$$

$$= |A(1.73) - A(0.81)|$$

$$= |0.4582 - 0.2910|$$

$$= 0.1672$$



1.70 | 0.03 →

ii) $x_1 = -1.43$

0.8 | 0.01

$$z_1 = \frac{-1.43 - 1}{3} = -0.81$$

$x_2 = 6.19$

$$z_2 = \frac{6.19 - 1}{3} = 1.73$$

$$P(-0.81 \le z \le 1.73) = A(-0.81) + A(1.73)$$

$$= A(0.81) + A(1.73)$$

$$= 0.2910 + 0.4582$$

$$= 0.7472.$$

(32) If X is a normal variate with mean 30, standard deviation 5, find the prob. that i) $26 \le X \le 40$ ii) $X \ge 45$

Sol: Given $\mu = 30$, and $\sigma = 5$

i) $x_1 = 26$

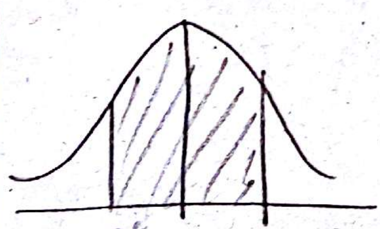
$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{26-30}{5} = -0.8$$

$x_2 = 40$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{40-30}{5} = 2$$

$$P(26 \le x \le 40) = p(-0.8 \le z \le 2)$$

$$= \cancel{P}\ A(-0.8) + A(2)$$



$z_1 = -0.8$ $\quad z_2 = 2$

$$= A(0.8) + A(2)$$

$$= 0.2881 + 0.4772$$

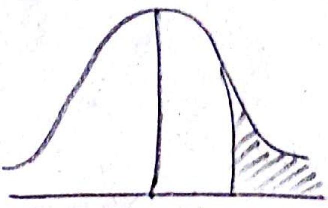
$$= 0.7653$$

ii) $\cancel{x_1} =$ $P(X \ge 45)$

$x_1 = 45$

$$z_1 = \frac{45-30}{5} = 3$$

$$P(X \geq 45) = P(Z \geq 3)$$
$$= 0.5 - A(3)$$
$$= 0.5 - 0.4987$$
$$= 1.3 \times 10^{-3}$$
$$= 0.0013$$



$z = 3$

Prob

04-04-17

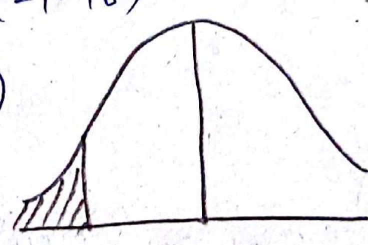
(33) If X is the normal variate. Find the area for which i) To the left of $z = -1.78$

ii) To the right of $z = -1.45$

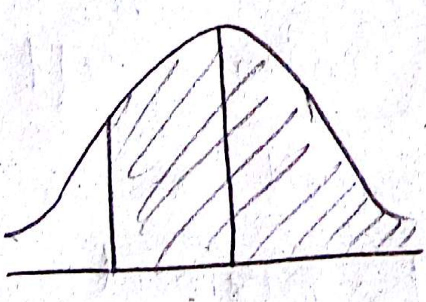
iii) corresponding to $-0.8 \leq z \leq 1.53$

iv) To the left of $z = -2.52$, and to the right of $z = 1.83$.

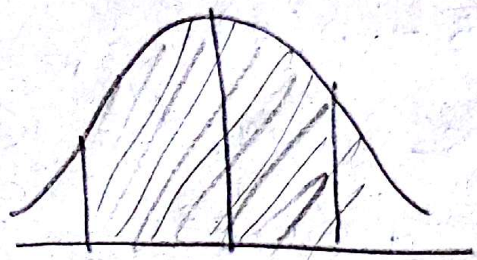
Sol:) i) $P(z < -1.78) = 0.5 - A(-1.78)$
$$= 0.5 - A(1.78)$$
$$= 0.5 - 0.4625$$
$$= 0.0375$$



ii) $P(z > -1.45) = 0.5 + A(-1.45)$
$$= 0.5 + A(1.45)$$
$$= 0.5 + 0.4265$$
$$= 0.9265$$



iii) $P(-0.8 \le z \le 1.53) =$



$= A(-0.8) + A(1.53)$

$= A(0.8) + A(1.53)$

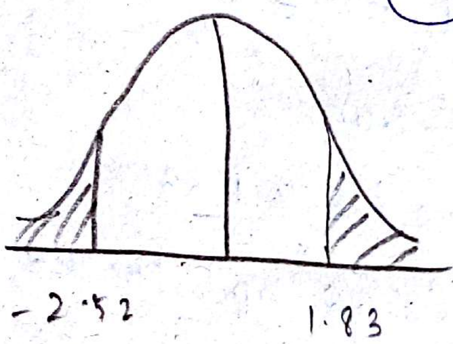
$= 0.2881 + 0.4370$

$= 0.7251$

iv)

$P(z < -2.52) + p(z > 1.83)$

%



$= [0.5 - A(-2.52)] +$
$[0.5 - A(1.83)]$

$= 1 - [A(-2.52) + A(1.83)]$

$= 1 - [A(2.52) + A(1.83)]$

$= 1 - [0.4941 + 0.4664]$

$= 0.0395$.

(34) In a normal distribution, 7% of items are under 35 and 89% are under 63. Determine the mean & variance of the distribution.

over - right
under - left

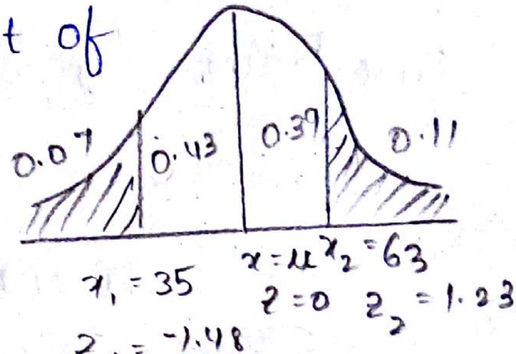
sol: let μ be the mean and σ be the standard deviation of normal distribution

2.5
0.02

Given that 7% of items are under 35

$\Rightarrow$ the area to the left of

$x_1 = 35$ is 7%



$\Rightarrow P(X < 35) = 7/100$

$= 0.07$

Also, given that 89% of items are under 63

$\Rightarrow$ area to the left of $x_2 = 35$ is 89%.

(or)

$\Rightarrow$ ~~or~~ 11% of items are over 63

which means the area 11% to the right of $x_2 = 63$

$89\% = \frac{89}{100} = 0.89$

$\Rightarrow P(X > 63) = 1 - P(X < 63)$

$= 1 - 0.89$

$= 0.11$

$\Rightarrow$ we have x values, we have to find z values from table.

By taking reverse look up from normal tables, we have $z_1 = -1.48$, $z_2 = 1.23$

~~$z_1 = -a_T$ Edel~~

~~$z_1 = 0.4306$ & $z_2 =$~~

Look 0.43 area in the table & know z values

$\Rightarrow \quad z_1 = -1.48 \quad \& \quad z_2 = 1.23.$

we have $\quad z_1 = \dfrac{x_1 - \mu}{\sigma}$

$$\Rightarrow \quad -1.48 = \frac{35 - \mu}{\sigma}$$

$$\Rightarrow \quad -1.48\sigma = 35 - \mu$$

$$\mu - 1.48\sigma = 35 \quad \text{— (1)}$$

$$z_2 = \frac{x_2 - \mu}{\sigma} \quad \Rightarrow \quad 1.23 = \frac{63 - \mu}{\sigma}.$$

$$\Rightarrow \quad 1.23\sigma = 63 - \mu.$$

$$\Rightarrow \quad \mu + 1.23\sigma = 63. \quad \text{— (2)}$$

$$\begin{array}{r} \mu + 1.48\sigma = 35 \\ - \qquad\qquad\quad - \\ \hline 2.\overset{71}{\cancel{83}}\sigma = 28 \end{array}$$

$$\cancel{\sigma = 10.64}6$$

$$\boxed{\sigma = 10.332}$$

eq (1) $\Rightarrow \quad \mu - 1.48(10.332) = 35$

$$\mu = 35 + 15.29$$

$$\boxed{\mu = 50.29}$$

Mean $= 50.29$

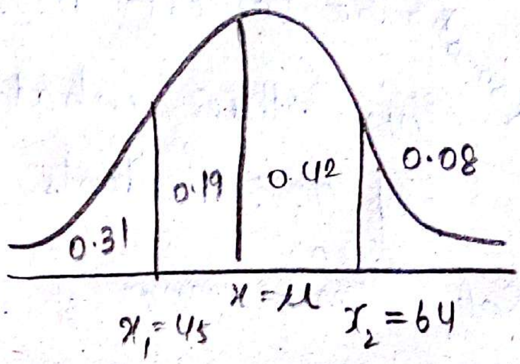
Variance $(\sigma^2) = (10.332)^2 = 106.75$

(35) In a normal distribution, 31% of items are under 45 and 8% are over 64. find the mean and variance of the distribution.

Sol:

$z_1 = -0.50$

$z_2 = 1.41$



$$z_1 = \frac{x_1 - \mu}{\sigma}$$

$$-0.5 = \frac{45 - \mu}{\sigma}$$

$$\mu - 0.5\sigma = 45 \quad — (1)$$

$$z_2 = \frac{x_2 - \mu}{\sigma} \Rightarrow 1.41 = \frac{64 - \mu}{\sigma}$$

$$\mu + 1.41\sigma = 64 \quad — (2)$$

$$\mu + 0.5\sigma = 45$$

$$- \qquad\qquad -$$

$$1\,\cancel{.91\sigma} = 0$$

$$\sigma = 9.947$$

eq (1) $\Rightarrow \mu = 45 + 0.5(9.947)$

$$\boxed{\begin{array}{l} \text{Mean} \quad \mu = 49.973. \\ \text{Variance } (\sigma^2) = 98.942 \end{array}}$$

6) The marks obtained in mathematics by 1000 students is normally distributed with mean 78% and standard deviation 11%. Determine i) How many students got marks above 90%.

ii) what was the highest mark obtained by the lowest 10% students.

$x \to$ marks

iii) within what limits, ~~wil~~ did the middle of 90% students lie?

Sol: Given:

$n = 1000$; $\mu = 78\% = 0.78$

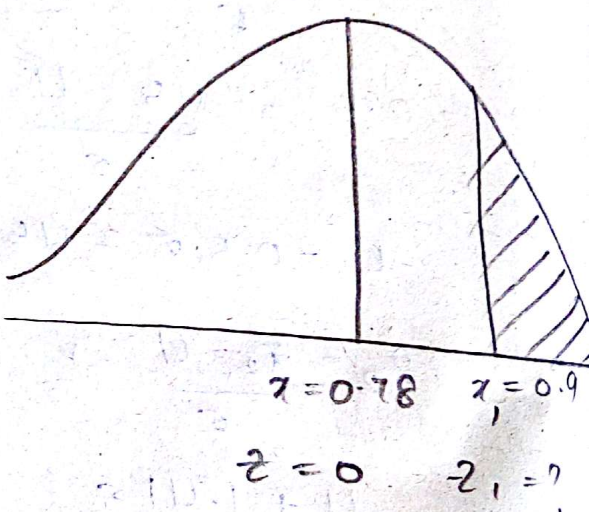
$\sigma = 11\% = 0.11$

$\sigma^2 = 0.0121$

i) $P(X > 0.9) =$

$x_1 = 90\% = 0.9$

$x_1 = 0.9$



$$z_1 = \frac{x_1 - \mu}{\sigma}$$

$$= \frac{0.9 - 0.78}{0.11}$$

$z_1 = 1.09$

$P(X > 0.9) = P(z > 1.09)$

$= 0.5 - A(1.09)$

$= 0.5 - 0.3621$

$= 0.1379$

The no. of students above 90% marks is

$= 1000 \times 0.1379$

$= 137.9$

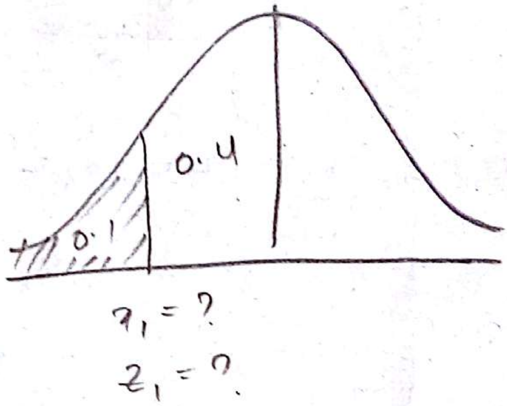
$= 138$ students

ii) The 10% area to the left of z corresponds to the lowest 10% students

we need the marks i.e. value of x exactly on 10% line

~~...~~,

By reverse look up, we get z value from table



i.e. $z_1 = -1.28$

we have $z_1 = \dfrac{x_1 - \mu}{\sigma}$ $\Rightarrow$ ~~$x =$~~

$= -1.28 = \dfrac{x_1 - 0.78}{0.11}$

$x_1 = 0.6392$ (marks in decimal)

$\Rightarrow$ $x_1 = 0.639.2 \times 100$

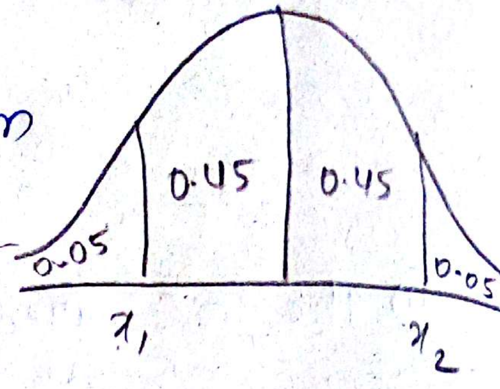
$= 63.92$ ~~or~~

$x_1 = 64$ % (marks in ~~...~~ %)

iii) middle 90%

$\Rightarrow$ 45% on either sides

By reverse look up from table,



$z_1 = -1.64$ &

$z_2 = 1.64$

$z_1 = \frac{x_1 - \mu}{\sigma} \Rightarrow -1.64 = \frac{x_1 - 0.78}{0.11}$

$\Rightarrow x_1 = 0.5996$ (in decimal)

$x_1 = 0.5996 \times 100$

$x_1 = 59.96$

$x_1 = 60\%$ (in %)

$z_2 = \frac{x_2 - \mu}{\sigma} \Rightarrow 1.64 = \frac{x_2 - 0.78}{0.11}$

$x_2 = 0.9604$

$x_2 = 0.9604 \times 100$

$x_2 = 96.04$

$x_2 = 96\%$

B/w 60% & 96%, middle of 90% students lie.

Normal Approximation to Binomial Distribution:

Nrml distribution can be used to Approximate Binomial Distribution suppose no. of sucess ranges from x_1 to x_2 then, we find $z_1 = \frac{\left(x_1 - \frac{1}{2}\right) - \mu}{\sigma}$

$$z_2 = \frac{\left(x_2 + \frac{1}{2}\right) - \mu}{\sigma}$$

then we can find the probability using normal distribution.

① Find the probability that out of 100 patients b/w 84 and 95 (inclusive) will survive a heart operation. Given that chance of survival is 0.9.

Ans:

$n = 100$

$p = 0.9$

$x_1 = 84 \qquad x_2 = 95$

$$z_1 = \frac{\left(x_1 - \frac{1}{2}\right) - \mu}{\sigma} = \frac{\left(84 - \frac{1}{2}\right) - n \cdot p}{n \cdot p \cdot q}$$

$$= \frac{\left(84 - \frac{1}{2}\right) - 90}{\sqrt{9}}$$

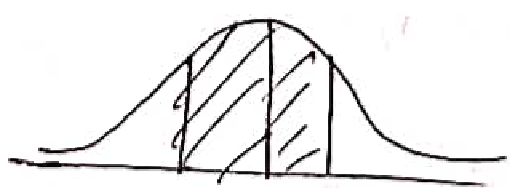
$$= \frac{\left(84 - \frac{1}{2}\right) - 90}{3}$$

$$= -2.16$$

$$z_2 = \frac{\left(95 + \frac{1}{2}\right) - 90}{3} = 1.83$$

$P(84 \leq x \leq 95)$

$P(-2.16 \leq z \leq 1.83)$



$A(z_1) + A(z_2) \Rightarrow 0.4846 + 0.4664$

$= 0.951$

(2) 8 coins are tossed together find the probability of getting 1 to 4 heads.

Sol: $n = 8 \quad p = \frac{1}{2} \quad q = \frac{1}{2}$

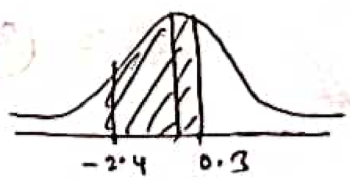
$x_1 = 1 \quad x_2 = 4$

$z_1 = \frac{\left(1 - \frac{1}{2}\right) - 4}{\sqrt{2}} = -2.474$

$\mu = np = 4$

$\sigma = npq = 2$

$z_2 = \frac{\left(4 + \frac{1}{2}\right) - 4}{\sqrt{2}} = 0.353$



$A(z_1) + A(z_2)$

$= 0.4932 + 0.1368$

$= 0.63$

* Find the probability by guess work a student can correctly answer 25 to 30 questions in a multiple choice quiz consisting of 80 questions. Assume that in each question with '4' choices only '1' choice is correct and student has no knowledge of the subject.

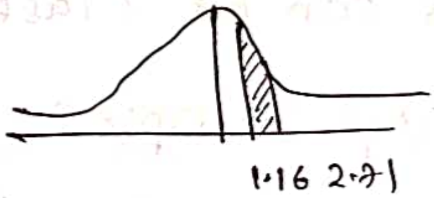
Sol: $n = 80 \quad p = \frac{1}{4} \quad q = 3/4$

$x_1 = 25 \quad x_2 = 30$

$$z_1 = \frac{\left(25 - \frac{1}{2}\right) - 20}{\sqrt{15}} = 1.161$$

$$Z_2 = \frac{\left(30 + \frac{1}{2}\right) - 20}{\sqrt{15}} = 2.711$$

$\Rightarrow A(Z_2) - A(Z_1)$

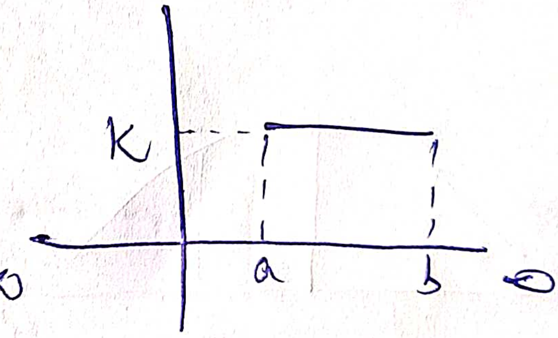


$\Rightarrow 0.4966 - 0.37720$

$= 0.1196$

Uniform distribution:-

Uniform distribution is a continuous distribution.
A random variable X is said to be uniformly distributed over the interval $-\infty < a < b < \infty$
If its probability density function is constant over the entire range. As the distribution is as rectangle it is also called as rectangular distribution



K = constant.

It is also called as rectangular distribution.

Note: To find k

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \Rightarrow \int_a^b f(x)\,dx = 1$$

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

$$k(b-a) = 1$$

$$\boxed{k = 1/(b-a)}$$

→ Probability distribution function ⇒ $f(x) = \begin{cases} \frac{1}{b-a}, & a<x<b \\ 0, & \text{otherwise} \end{cases}$

Measures of uniform distribution:

1) Mean, μ:

$$\mu = \int_{-\infty}^{\infty} x f(x)\,dx = \int_a^b xk\,dx \Rightarrow k\left(\frac{x^2}{2}\right)_a^b = \frac{1}{b-a}\left(\frac{b^2-a^2}{2}\right)$$

$$= \frac{(b+a)\cancel{(b-a)}}{2\cancel{(b-a)}} \Rightarrow \boxed{\frac{b+a}{2} = \mu}$$

2) Variance:

$$\int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2 = \int_a^b x^2 k\,dx - \mu^2$$

$$= k\left.\frac{x^3}{3}\right|_a^b + \mu^2 \Rightarrow \frac{b^3-a^3}{3(b-a)} - \frac{(a+b)^2}{4}$$

$$= \frac{\cancel{(b-a)}(b^2+ab+a^2)}{3\cancel{(b-a)}} - \frac{(a+b)^2}{4}$$

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

$$= \frac{4b^2+4ab+4a^2-3a^2-3b^2-6ab}{12}$$

$$= \frac{a^2-2ab+b^2}{12} = \boxed{\frac{(b-a)^2}{12} = \sigma^2}$$

To find the probability between a1 to a2:-

$$P(a_1 < x < a_2) = \int_{a_1}^{a_2} f(x)\,dx = \int_{a_1}^{a_2} \frac{1}{b-a}\,dx = \frac{1}{b-a}(x)_{a_1}^{a_2}$$

$$= \frac{a_2 - a_1}{b-a}.$$

1) A bus is uniformly late between 2 to 10 mins, How long can you expect wait. What is SD (σ). If it's greater than 7 min late, you'll be late for work what is the probability of you being late.

Sol $a=2,\ b=10$

$\mu = ?$ (Expectation) $\quad f(x) = \frac{1}{b-a} = \frac{1}{10-2} = \frac{1}{8} = 0.125$

$$\mu = \frac{a+b}{2} = \frac{2+10}{2} = 6$$

$$S.D = \sqrt{\sigma^2} \quad \sigma^2 = \frac{(b-a)^2}{12} = \frac{(10-2)^2}{12} = \frac{\cancel{64}^{16}}{\cancel{12}_{3}} = 5.33$$

$= 2.309$

$$P(x>7) = \int_7^{10} f(x)\,dx = \int_7^{10} 0.125\,dx = 0.125 \times 3$$
$$= 0.375$$

2) The amount of time a person must wait for a train to arrive in a certain town is uniformly distributed btw 0 to 40. i) probability density funct?

ii) Draw a graph of $f(x)$

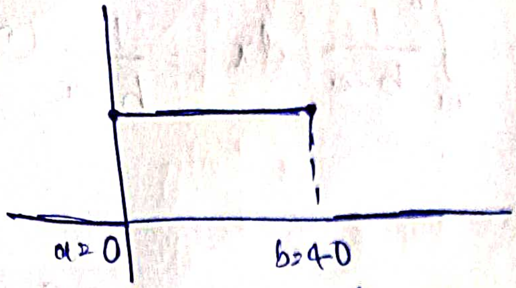
iii) What is the probability that a person must wait less than 8 mins

iv) What is the probability that a person must waite more than 30 mins.

v) find $P(10 < x < 20)$, $P(x > 45)$ vi) μ, σ^2

$a=0,\ b=40$ i) $f(x)=\frac{1}{b-a}=\frac{1}{40}=0.025$

ii)



iii) $P(X<8)=\int_{0}^{8} f(x)\,dx=\int_{0}^{8} 0.025\,dx=0.025(8)=0.2$

iv) $P(X>30)=\int_{30}^{40} f(x)\,dx=\int_{30}^{40} 0.025\,dx=0.025\times 10=0.25$

v) $P(10<X<26)=\int_{10}^{26} 0.025\,dx=0.025\times 16=0.4$

vi) $P(X>45)=0$

vi) $\mu=\frac{a+b}{2}=\frac{40}{2}=20$ $\sigma^2=\frac{(b-a)^2}{12}=\frac{(40)^2}{12}=\frac{40\times 40}{12}=133.$

Standard Deviation $\Rightarrow \sigma=11.54$

3) The amount of time that it takes a student to complete a chemistry test is uniformly distributed btw 20 & 45 min

i) probability density function ii) graph of $f(x)$

iii) what is the prob that a student will take > 36

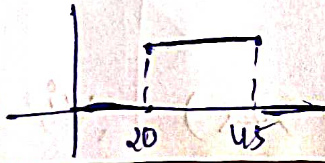
iv) " " " " " " " btw 26 & 35 mins to complete the test

v) Determine μ, σ^2 vi) probability that student will take more than 50 mins

Sol $a=20 \quad b=45$

i) $f(x)=\frac{1}{45-20}=\frac{1}{25}=0.04$

ii)



iii) $P(X > 36) = \int_{36}^{45} f(x)dx = \frac{45-36}{45-20} = 0.04 \times 9 = \underline{\underline{0.36}}$

iv) $P(26 < X < 35) = \int_{26}^{35} f(x)dx = 9 \times 0.04 = \underline{\underline{0.36}}$

v) $\mu = \frac{a+b}{2} = \frac{20+45}{2} = \frac{65}{2} = 32.5$

$\sigma^2 = \frac{(b-a)^2}{12} = \frac{(45-20)^2}{12} = \frac{25^2}{12} = \underline{\underline{52.08}}$

vi) $P(X > 50) = \underline{\underline{0}}$

Sampling

Population: The totality of observations with which we are concerned, whether this number be finite or infinite, constitutes what we call population. The size of the population is denoted by N

Sample: A portion of the population which is examined with a view to determining the population characteristics is called a sample. The size of the sample is denoted by n

Different methods of sampling
Some important methods of sampling are discussed below.

I. Probability Sampling Methods

1. Random sampling or Probability sampling

It is the process of drawing a sample from a population in such a way that each member of the population has an equal chance of being included in the sample. The sample obtained by the process of random sampling is called a random sample.

If N is the size of a population and n is the sample size, then

(i) The number of sample with replacement = N^n

(i) The number of sample without replacement = ${}^NC_n$

2. Stratified sampling or Stratified Random sampling

This method is useful when the population is heterogeneous. In this type of sampling, the population is first sub divided into several parts or small groups called *strata*according to some relevant characteristics so that each stratum is more or less homogeneous. Each stratum is called a sub-population. Then a small sample called sub-sample is selected from each stratum at random. All the sub samples are combined together to form the stratified sample which represents the population properly. The process of obtaining and examining a stratified sample with a view to estimating the characteristic of the population is known as *Stratified Sampling.*

3. Systematic Sampling or Quasi – Random Sampling

As the name suggests this means forming the sample in some systematic manner by taking items at regular intervals. In this method, all the units of the population are arranged in some order. If the population size is finite, all the units of the population are arranged in some order. Then from the first k items, one unit is selected at random. This unit and every k^{th} unit of the serially listed population combined together constitute a systematic sample. This type of sampling is known as Systematic Sampling.

II. Non – ProbabilitySampling Methods

1. Purposive Sampling or Judgment Sampling

When the choice of the individual items of a sample entirely depends on the individual judgment of the investigator (or sampler), it is called a purposive or *Judgment Sampling.* For example, if a sample of 20 students is to be selected from a class of 100 to analyze the extra-curricular activities of the students, the investigator would select 20 students who, in his judgment, would represent the class.

2. Sequential Sampling

It consists of a sequence of sample drawn one after another from the population depending on the results of previous sample. If the result of the first sample leads to a decision which is not acceptable, the lot from which the sample was drawn is rejected. But if the result of the first sample is acceptable, no new sample is drawn. But if the first sample leads to no clear decision, a second sample is drawn and as before if required a third sample is drawn to arrive at a final decision to accept or reject the lot. This process is called *Sequential Sampling*

Classification of Samples

Samples are classified in two ways.

1. Large Sample: if the size of the sample (n) ≥ 30, the sample is said to be large sample

2. Small Sample: if the size of the sample (n) < 30, the sample is said to be small sample

Parameter and Statistics

Parameter is statistical measures of the population. Ex: population mean (μ), population variance (σ^2)

Statistic is statistical measures of the Sample. Ex: Sample mean ($\bar{x}$), population variance (s^2)

The Sample Mean :

If X_1, X_2 ……, X_n represents a random sample of size n, then the sample mean is defined by the statistic $\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n}$

The sample Variance:

If X_1, X_2 ……, X_n represents a random sample of size n, then the sample variance is defined by the statistic $s^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}$

Central Limit Theorem

If $\bar{x}$be the mean of a sample of size n drawn from a population with mean μ and S.D. σthen the standardized sample mean $z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$ is a random variable whose distribution function approaches that to the standard normal distribution N(z;0,1) as $n \to \infty$

STANDARD ERROR (S.E.) OF A STATISTIC

1) S.E. ($\bar{x}$) = $\sigma / \sqrt{n}$
2) S.E. of sample proportion $p = \sqrt{\frac{PQ}{n}}$ where Q=1-P
3) S.E. $\left(\overline{x_1} - \overline{x_2}\right) = \sqrt{\frac{\sigma_1}{n_1} + \frac{\sigma_2}{n_2}}$ where $\overline{x_1}$ and $\overline{x_2}$ are the means of two random samples of sizes n_1 and n_2 drawn from two populations with S.D. σ_1 and σ_2 respectively.
4) S.E. $\left(s_1 - s_2\right) = \sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}$

Finite Population: Consider a finite population of size N with mean μ and S.D. σ. Draw all possible samples of size n without replacement, from this population. Then

(i) The mean of the sampling distribution of means (for N>n) is $\mu_{\overline{X}} = \mu$

(ii) The variance is $\sigma_{\overline{X}}^{\ 2} = \dfrac{\sigma^2}{n}\left(\dfrac{N-n}{N-1}\right)$

Note : The factor $\left(\dfrac{N-n}{N-1}\right)$ is called the finite population **correction factor.**

1. What is the value of correcting factor if n = 5 and N = 200

Sol. Given N = the size of the finite population = 200

n = the size of the sample = 5

$\therefore$ Correction factor = $\dfrac{N-n}{N-1}$ = $\dfrac{200-5}{200-1} = \dfrac{195}{199} = 0.98$

2. What is the value of correcting factor if n = 10 and N = 1000

Sol. Given N = the size of the finite population = 1000

n = the size of the sample = 10

$\therefore$ Correction factor = $\dfrac{N-n}{N-1}$ = $\dfrac{1000-10}{1000-1} = \dfrac{990}{999} = 0.991$

3. A population consists of five numbers 2, 3, 6, 8 and 11. Consider all possible samples of size two which can be drawn with replacement from this population. Find

(a) The mean of the population
(b) The S.D. of the population
(c) The mean of the sampling distribution of means and
(d) The S.D. of the sampling distribution of means (i.e., the standard error of means)

Sol.

(a) Mean of the population is given by $\mu = \dfrac{2+3+6+8+11}{5} = \dfrac{30}{5} = 6$

(b) Variance of the population σ^2 is given by

$$\sigma^2 = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{n}$$
$$= \frac{(2-6)^2 + (3-6)^2 + (6-6)^2 + (8-6)^2 + (11-6)^2}{5}$$
$$= \frac{16+9+0+4+25}{5} = 10.8$$

(c) Sampling with replacement (infinite population) :

The total no. of samples with replacement is

$N^n = 5^2 = 25$ samples of size 2

Here N = population size and n = sample size listing all possible samples of size 2 from population 2, 3, 6, 8, 11 with replacement we get 25 samples

$$\begin{Bmatrix} (2,2) & (2,3) & (2,6) & (2,8) & (2,11) \\ (3,2) & (3,3) & (3,6) & (3,8) & (3,11) \\ (6,2) & (6,3) & (6,6) & (6,8) & (6,11) \\ (8,2) & (8,3) & (8,6) & (8,8) & (8,11) \\ (11,2) & (11,3) & (11,6) & (11,8) & (11,11) \end{Bmatrix}$$

Now compute the arithmetic mean for each of these 25 samples. The set of 25 means $\bar{x}$ of these 25 samples, gives rise to the distribution of means of the samples known as sampling distribution of means.

The samples means are

$$\begin{Bmatrix} 2 & 2.5 & 4 & 5 & 6.5 \\ 2.5 & 3 & 4.5 & 5.5 & 7 \\ 4 & 4.5 & 6 & 7 & 8.5 \\ 5 & 5.5 & 7 & 8 & 9.5 \\ 6.5 & 7 & 8.5 & 9.5 & 11 \end{Bmatrix}$$

And the mean of sampling distribution of means is the mean of these 25 means.

$$\mu_{\overline{X}} = \frac{Sum\, of\ all\ sample\, means}{25} = \frac{150}{25} = 6$$

(d) The variance of the sampling distribution of means is obtained by subtracting the mean 6 from each number and squaring the result, adding all 25 members thus obtained, and dividing by 25

$$\sigma_{\overline{X}}^{\ 2} = \frac{(2-6)^2 + \ldots\ldots\ldots\ldots + (11-6)^2}{25} = \frac{135}{25} = 5.4$$

$$\therefore\ \sigma_{\overline{X}} = \sqrt{5.40} = 2.32$$

Clearly $\sigma_{\overline{X}}^{\ 2} = \dfrac{\sigma^2}{n} = \dfrac{10.8}{2} = 5.4 \Rightarrow \sigma_{\overline{X}} = \sqrt{5.4} = 2.32$

4. Solve the above examples without replacement to find $\mu_{\overline{X}}$ and $\sigma_{\overline{X}}$

Sol.

$\mu = 6$ and $\sigma = 3.29$

Sampling without replacement (finite population) :

The total no. of samples without replacement is ${}^{N}C_n = {}^{5}C_2 = 10$ samples of size 2

$$\begin{Bmatrix} (2,3) & (2,6) & (2,8) & (2,11) \\ (3,6) & (3,8) & (3,11) & \\ (6,8) & (6,11) & & \\ (8,11) & & & \end{Bmatrix}$$

The corresponding sample means are

$$\begin{Bmatrix} 2.5 & 4 & 5 & 6.5 \\ 4.5 & 5.5 & 7 & \\ 7 & 8.5 & & \\ 9.5 & & & \end{Bmatrix}$$

The mean of the sampling distribution of means is

$$\mu_{\overline{X}} = \frac{Sum\ of\ all\ sample\ means}{25} = \frac{(2.5+4+......8.5+9.5)}{10} = \frac{60}{10} = 6$$

The variance of sampling distribution of means

$$\sigma_{\overline{X}}^{\ 2} = \frac{(2.5-6)^2 + ... + (9.5-6)^2}{10} = \frac{40.5}{10} = 4.05$$

Showing that $\sigma_{\overline{X}}^{\ 2} = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{10.8}{2}\left(\frac{5-2}{5-1}\right) = 4.05$

5(H.W). A population consists of 6 numbers 5, 10, 14, 18, 13, and 24. Consider all possible samples of size two which can be drawn without replacement from this population. Find

(a) The mean of the population
(b) The S.D. of the population
(c) The mean of the sampling distribution of means and
(d) The S.D. of the sampling distribution of means (i.e., the standard error of means)

6. The mean height of students in a college is 155cms and S.D. is 15. What is the probability that the mean height of 36 students is less than 157 cms.

Sol.

μ = mean of the population
= mean height of students of a college = 155 cm
σ = S.D. of population = 15 cms
n = sample size = 36
$\overline{x}$ = Mean of samples = 157 cms

Now $z = \frac{\overline{x}-\mu}{\sigma/\sqrt{n}} = \frac{157-155}{15/\sqrt{36}} = 0.8$

$\therefore P(\overline{x} \le 157) = P(z < 0.8) = 0.5 + A(0.8) = 0.5 + 0.2881 = 0.7881$

Thus the probability that the mean height of 36 students less than 157 = 0.7881

7. A random sample of size 100 is taken from an infinite population having the mean μ = 76 and the variance σ^2 = 256. What is the probability that $\overline{x}$ will be between 75 and 78.

Sol.

μ = mean of the population = 76
σ^2 = variance of population = 256 i.e. σ = 16
n = sample size = 100
$\overline{x}_1$ = 75

Now $z_1 = \frac{\overline{x}_1-\mu}{\sigma/\sqrt{n}} = \frac{75-76}{16/\sqrt{100}} = -0.625$

And when $\bar{x}_2$ = 78

Now $z_2 = \dfrac{\bar{x}_1 - \mu}{\sigma/\sqrt{n}} = \dfrac{78-76}{16/\sqrt{100}} = 1.25$

$\therefore P\left(75 \le \bar{x} \le 78\right) = P\left(-0.625 \le z \le 1.25\right) = A\left(-0.625\right) + A\left(1.25\right) = 0.2334 + 0.3944 = 0.628$

8. A random sample of size 64 is taken from an infinite population having the mean 45 and the S.D. 8. What is the probability that x will be between 46 and 47.5

Sol.

μ = mean of the population = 45

σ = S.D. of the population = 8

n = sample size = 64

$\bar{x}_1$ = 46

Now $z_1 = \dfrac{\bar{x}_1 - \mu}{\sigma/\sqrt{n}} = \dfrac{46-45}{8/\sqrt{64}} = 1$

And when $\bar{x}_2$ = 47.5

Now $z_2 = \dfrac{\bar{x}_2 - \mu}{\sigma/\sqrt{n}} = \dfrac{47.5-45}{8/\sqrt{64}} = 2.5$

$\therefore P\left(46 \le \bar{x} \le 47.5\right) = P\left(1 \le z \le 2.5\right) = A\left(2.5\right) - A\left(1\right) = 0.4938 - 0.3413 = 0.1525$

9(H.W). A random sample of size 64 is taken from a normal population with μ = 51.4 and σ = 68. What is the probability that the mean of the sample will (a) exceed 52.9, (b) fall between 50.5 and 52.3, (c) 50.6