* sampling distribution :.

* population: The group of individuals under study is called population or universe. It may be finite or infinite

* Sampling : A part selected from the population is called a sample and the process of selection of sample is called sampling.

Note that :

* mean of population is denoted by 'μ'
* standard deviation of population is denoted by 'σ'
* Mean of sample is denoted by '$\bar{x}$'
* standard deviation of sample is denoted by 's'
* size of population is denoted by 'N'
* size of sample is denoted by 'n'

* Testing of hypothesis :

on the basis of sample information, we make certain decisions about population. In taking such decision we make assumptions these assumption are known as statical hypothesis.

* Null hypothesis: (Ho)

null hypothesis is no difference, thus we shall persume that there is no significant difference between the observed value and the expected value.

* Alternative hypothesis: (Ha)

It specifies a range of values rather than one value.

* levels of significance: (α)

It is expressed in the percentage as
5 % level of significance or
1 % level of significance.

* critical region:

The levels marked by probabilities 0.05 or 0.01 which decide the significance of an event are called level of significance and are expressed in percentage as 5 % level of significance or 1 % level of significance.
The corresponding region are called critical region.

* Two Tailed and one Tailed test :

The probability distribution of a sample statistic is normal distribution.

The z - curve is symmetrical as we know and the parts of the curve at the two ends are called the two tails of the curve.

If the rejection area lies on two sides i.e on the two tailes the test is called the two tailed test.

If on the other hand the rejection area lies on one side only the test is called one tailed test

Note that :

1) $\mu > \mu_0$ or $\mu < \mu_0$ is two tailed test

2) $\mu > \mu_0$ is Right Tailed Test (one tailed test)

3) $\mu < \mu_0$ is left tailed test (one tailed test

	level of significance	
	1 %	5 %
Two tailed test	$Z_\alpha = 2.576$	1.96
one tailed test	$Z_\alpha = 2.326$	1.64

Ex 1. A random sample of 50 items gives the mean 6.2 and variance 10.24. Can it be regarded as drawn from a normal population with mean 5.4 at 5% level of significance?

solution: Given: $n = 50$

$\bar{x} = 6.2$

$\mu = 5.4$

$s = \sqrt{10.24}$

i) Null Hypothesis (H_0): $\mu = 5.4$

Alternative Hypothesis H_a : $\mu \neq 5.4$

ii) Test statistic :

Since the population S.D. is unknown

But sample S.D. 's' is known

$$\therefore \quad Z = \left| \frac{\bar{x} - \mu}{s/\sqrt{n}} \right|$$

$$= \left| \frac{6.2 - 5.4}{\sqrt{10.24}/\sqrt{50}} \right| = \left| \frac{0.8}{3.2/7.07} \right|$$

$$= 1.77$$

iii) Level of significance : $\alpha = 0.05$ $\left(5\% = \frac{5}{100}\right)$

iv) Critical value: The value of Z_α at 5% level of significance from the table $= 1.96$

v> Decision: since; $|Z| = 1.77$ i.e calculated value which is less than the critical value $Z_\alpha = 1.96$

hence, the null hypothesis is accepted.

$\therefore$ The sample is drawn from the population with mean 5.4.

Ex. 2. A random sample of 400 members is found to have a mean of 4.45 cms can it be reasonably regarded as a sample from a large population whose mean is 5 cms and variance is 4 cms.

Solution: Given:

$n = 400$

$\bar{x} = 4.45$

$\mu = 5$

$\sigma = \sqrt{4} = 2$

i> Null Hypothesis $H_0 : \mu = 5$

Alternative hypothesis $H_a : \mu \neq 5$

ii> Test statistic :

$$Z = \left| \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \right|$$

Scanned by CamScanner

$$\therefore \quad Z = \left| \frac{4.45 - 5}{2/\sqrt{400}} \right| = \left| \frac{0.55}{2/20} \right|$$

$$\Rightarrow \quad Z = 5.5$$

iii) Level of significance : $\alpha = 0.05$ ($\because$ large sample)

iv) Critical value :

The value of Z_α at 5% level of significance from the table $= 1.96$

v) Decision : Since the computed value of $Z = 5.5$ is greater that the critical value $Z_\alpha = 1.96$

$\therefore$ The null hypothesis is rejected and the alternative hypothesis is accepted.

$\therefore$ The sample is not drawn from the above population.

* Distribution of the difference between means

– procedure to test the hypothesis:

Step 1. Given: sizes of two samples n_1, n_2 with mean $\bar{X}_1, \bar{X}_2$ respectively and means of populations μ_1, μ_2 and standard deviation of population σ_1, σ_2

Step 2. calculate $\bar{X}_1 - \bar{X}_2$

Step 3. find standard error (S. E)

$$S = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Step 4. find $Z = \left| \frac{\bar{X}_1 - \bar{X}_2}{S} \right|$

& take the decision.

Example: 1. The mean of two sample of sizes 1000 and 2000 respectively are 67.50 and 68.0 inches. Can the samples be regarded as drawn from the same population of standard deviation.

Solution: Given: $n_1 = 1000$, $n_2 = 2000$

$\bar{X}_1 = 67.50$, $\bar{X}_2 = 68.0$

$\sigma_1 = 2.5$ $\sigma_2 = 2.5$

i> Null Hypothesis H_0 : $\mu_1 = \mu_2$

Alternative hypothesis H_a : $\mu_1 \neq \mu_2$

ii> Calculation of statistic :

$$\overline{X}_1 - \overline{X}_2 = 67.5 - 68.0 = -0.5$$

Now, standard error (S.E.) is

$$S = \sqrt{\frac{\sigma_1^2}{n_1^2} + \frac{\sigma_2^2}{n_2^2}} = \sqrt{\frac{(2.5)^2}{1000} + \frac{(2.5)^2}{2000}}$$

$$= (2.5)\sqrt{\frac{1}{1000} + \frac{1}{2000}}$$

$$= 0.097$$

$$\therefore \quad Z = \left|\frac{\overline{X}_1 - \overline{X}_2}{S}\right| = \left|\frac{-0.5}{0.097}\right| = |-5.15| = 5.15$$

iii> Level of significance : $\alpha = 0.27\%$ (given)

iv> critical value :

The value of Z_α at 0.27% level of significance from the table is 3

v> Decision :

Note that the Computed value of $|Z| = 5.15$ is greater than the critical value $Z_\alpha = 3$

$\therefore$ The Hypothesis is Rejected.

$\therefore$ The sample cannot be regarded as drawn from the same population.

Ex. 2 The avarage of marks scored by 32 boys is 72 with standard deviation 8 while that of 36 girls is 70 with standard deviation 6. Test at 1 % level of significance whether the boys perform better than the girls.

Solution: Given:

$n_1 = 32$, $n_2 = 36$

$\overline{x}_1 = 72$, $\overline{x}_2 = 70$

$s_1 = 8$, $s_2 = 6$

i) Null Hypothesis H_0 : $\mu_1 = \mu_2$

Alternative Hypothesis H_a : $\mu_1 \neq \mu_2$

ii) Calculation of statistic : $\overline{x}_1 - \overline{x}_2 = 72 - 70 = 2$

the standard error (S.E)

$$s = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{(8)^2}{32} + \frac{(6)^2}{36}}$$

$$= \sqrt{3}$$

$$\therefore Z = \left| \frac{\overline{x}_1 - \overline{x}_2}{s} \right| = \left| \frac{2}{\sqrt{3}} \right| = 1.15$$

iii) Level of significance : $\alpha = 0.01$ $\left(1\% = \frac{1}{100}\right)$

iv) Critical value : The value of Z_α at 1 % level of significance from the table is 2.58

v) Decision!

since the computed value of $Z = 1.15$ is less than the critical value $Z_\alpha = 2.58$

Hence, the Null Hypothesis is accepted.

∴ Boys do not perform better than the Girls.

<u>Homework</u>:

<u>Ex 3</u>. Test the significance of the difference between the means of two normal population with the same standard deviation from the following data.

	Size	Mean	S.D
sample I	100	64	6
sample II	200	67	8