TEST OF SIGNIFICANCE OF SMALL SAMPLES

Definition:

When the size of the sample (n) is less than 30, then that sample is called a small sample.

The following are some important tests for small samples:

(i) Student's 't' test

(ii) F-test

(iii) ψ^2- test

Student's t-Distribution:

A random variable T is said to follow Student's t-distribution or simply t-distribution, if its probability density function is given by,

$$\boxed{f(t) = \frac{1}{\sqrt{v}\ \beta\left(\frac{v}{2}, \frac{1}{2}\right)} \left(1 + \frac{t^2}{v}\right)^{-\frac{(v+1)}{2}}} \quad ; -\infty < t < \infty$$

where v is called the number of degrees of freedom of the t-distribution.

Properties of t-Distribution:

1. The probability curve of the t-distribution is similar to the standard normal curve and is

symmetric about $t = 0$, bell-shaped and asymptotic to the t-axis.

2. For sufficiently large value of n, the t-distribution tends to the standard normal distribution.
3. The mean of the t-distribution is zero.
4. The variance of t-distribution is $\frac{v}{v-2}$, if $n > 2$

   and is greater than 1 but it tends to 1 as $v \to \infty$
5. The variable ranges from $-\infty$ to ∞.

Uses of t-distribution:

The t-distribution is used to test the significance of the difference between

1. The mean of a sample and the mean of the population.
2. The means of two samples and
3. The coefficient of correlation in the small sample and that in the population, assumed zero.

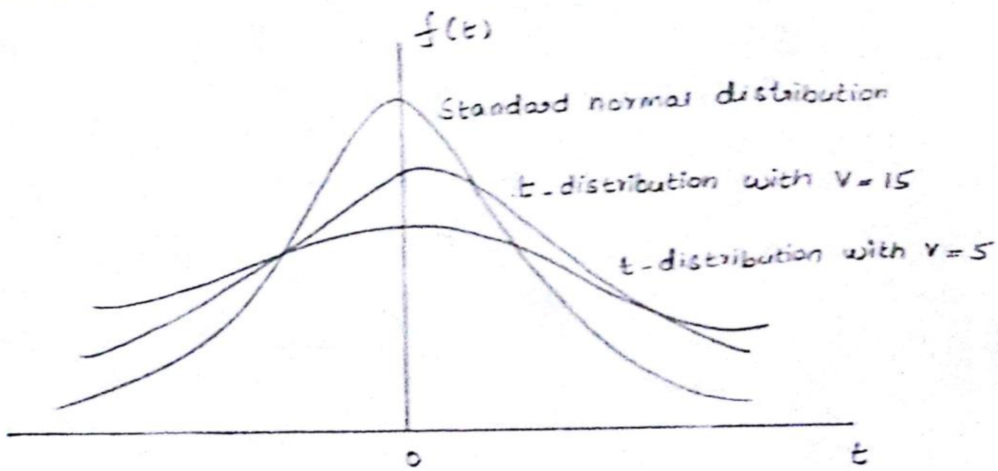
Assumptions for Student's t test:

The following assumptions are made in Student's t-test

1. The parent population from which sample is drawn is normal.

2. The sample observations are independent, that is, sample is random
3. The population Standard deviation σ is unknown
4. Sample size n is less than 30.

Graph of t-distribution:



Degrees of freedom:

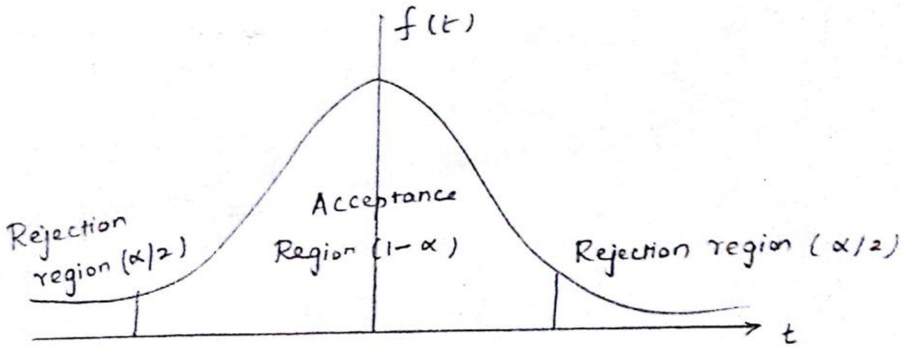
The number of independent variates used to compute the test statistic is known as the number of degrees of freedom of that statistic. In general, the number of degrees of freedom is given by $v = n - k$, where n is the number of observations in the sample and k is the number of constraints imposed on them or k is the number of values that have been found out and specified by prior calculations

Critical value of t :

The critical value or significant value of t at level of significance α and degrees of freedom v, for two tailed test, is given by

$$P\{|t| > t_v(\alpha)\} = \alpha$$

$$P\{|t| \le t_v(\alpha)\} = 1-\alpha .$$



The significant value of t at level of significance 'α' for a single-tailed test can be got from those of two-tailed test by referring to the values at '2α'. Critical value t_α are given in the tables, called t-table.

Test 1 :

Test of significance of the difference between sample mean and population mean :

Under the null hypothesis H_0 : The sample has been drawn from population with mean μ or there is no significant difference between the sample mean $\bar{x}$ and the population mean μ, we use the Statistics,

$$\boxed{t = \frac{\bar{x}-\mu}{S/\sqrt{n}} \quad (\text{or}) \quad \frac{\bar{x}-\mu}{s/\sqrt{n-1}}}$$

where $\bar{x}$ is the sample mean and

$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$ with degrees of freedom $(n-1)$.

At given level of significance α, and degrees of freedom $(n-1)$, we refer to t-table value t_α (two - tailed or one tailed).

If calculated t value is such that,

(i) $|t| < t_\alpha$, the null hypothesis is accepted

(ii) $|t| > t_\alpha$, the null hypothesis is rejected.

(7)

Problems

(1) Sandal powder is packed into packets by a machine. A random sample of 12 packets is drawn and their weights are found to be (in kg) 0.49, 0.48, 0.47, 0.48, 0.49, 0.50, 0.51, 0.49, 0.48, 0.50, 0.51, 0.48. Test if the average packing can be taken as 0.5 kg.

Solution:

Given: $n = 12$

Null hypothesis: H_0: The average packing can be taken as 0.5 kg. i.e., $H_0 : \mu = 0.5$

Alternative hypothesis: $H_1 : \mu \neq 0.5$ (Two-tailed test)

Calculation of Sample SD and Sample mean:

$$\bar{x} = \frac{\Sigma x}{n}$$

$$\bar{x} = \frac{0.49 + 0.48 + 0.47 + 0.48 + 0.49 + 0.50 + 0.51 + 0.49 + 0.48 + 0.50 + 0.51 + 0.48}{12}$$

$$\bar{x} = \frac{5.88}{12}$$

$$\boxed{\bar{x} = 0.49}$$

$$s^2 = \frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2$$

$$= [(0.2401)\times 3 + 4(0.2304) + (0.2209) + 2(0.2500) + 2(0.2601)] \times \frac{1}{12} - (0.49)^2$$

$$= 0.24025 - 0.2401$$

$$s^2 = 0.00015$$

$$\boxed{s = 0.012}$$

Test Statistic :

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}}$$

$$= \frac{0.49 - 0.5}{0.012/\sqrt{11}} = 2.7634$$

$$\boxed{t = 2.7634}$$

Table value :

At $\alpha = 5\%$ LOS, $v = n-1 = 11$ d.o.f,

the table value of t is given by,

$$\boxed{t_\alpha = 2.20}$$

Decision : Since $t > t_\alpha$, H_0 is rejected. $\therefore$ The average packing cannot be taken as 0.5 kg.

111

(16) A random sample of 10 boys has the following IQ's:
70, 120, 110, 101, 88, 83, 95, 98, 107, 100.
Do these data support the assumption of a population mean IQ of 100?

Solution:

Given: $n = 10$, $\mu = 100$

Null hypothesis: H_0: The mean IQ of the population can be assumed as 100. i.e., $H_0 : \mu = 100$.

Alternative hypothesis: $H_1 : \mu \neq 100$ (Two-tailed test)

Calculation of Sample mean and Sample SD:

$$\bar{x} = \frac{\Sigma x}{n} = \frac{972}{10}$$

$$\boxed{\bar{x} = 97.2}$$

$$s^2 = \frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2$$

$$= \frac{1834}{10} - (97.2)^2$$

$$\boxed{s = 13.72}$$

Test Statistic:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}}$$

$$= \frac{97.2 - 100}{13.72/\sqrt{9}}$$

$\boxed{|t| = 0.612}$

Table value :

At $\alpha = 5\%$ los , $\nu = n-1 = 9$ d.o.f , the table value of t is given by,

$\boxed{t_\alpha = 2.262}$

Decision :

Since $|t| < t_\alpha$, H_0 is accepted. Therefore the mean IQ of the population can be assumed as 100.