In case of large samples, sampling distribution approaches a normal distribution & values of simple statistic are considered best estimates of parameters in a popⁿ. It will no longer be ~~possible~~ ~~to~~ possible to assume the statistics computed from small samples are normally distributed. As such, new technique has been devised for small samples which involves the degree of freedom ~~tohe~~.

* No. of degrees of freedom:

It is number of values in set which may be assigned arbitrarily. eg: if $x_1 + x_2 + x_3 = 15$ & we assign any value to two of variables (say x_1, x_2) then values of x_3 will be known. Two variables are therefore free & independent choices for finding third. Hence these are degrees of freedom (d.f) If there are n observations, the d.f are (n-1). In other words, while finding the mean of small sample one degree of freedom is used up & (n-1) d.f are left to estimate the popⁿ variance

Date: / /

Students t distribution :-

The t - distribution is also known as Student's t distribution is the prob distribution that estimates the pop^n parameter when sample size is small & pop^n S.D is unknown. It resembles the normal distribution & as sample size increases t distribution looks more normally distributed with values of mean & S.D 0 & 1 resp.

Consider a small sample of size n drawn from a normal population with mean μ & S.D σ. If $\bar{x}$ & s be sample mean & S.D then

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$ is random variable

Date: / /

having t distribution with $\nu = n-1$ df

with pdf

$$y(t) = \frac{1}{\sqrt{\nu}\, B\left(\frac{1}{2}, \frac{\nu}{2}\right)} \cdot \frac{1}{\left(1 + \frac{t^2}{\nu}\right)^{\frac{\nu+1}{2}}}$$

$-\infty < t < \infty$

Here sample variance

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

Thus for small sample $(n < 30)$ &

with σ unknown, a natural statistic

for inference on pop^n mean μ is

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

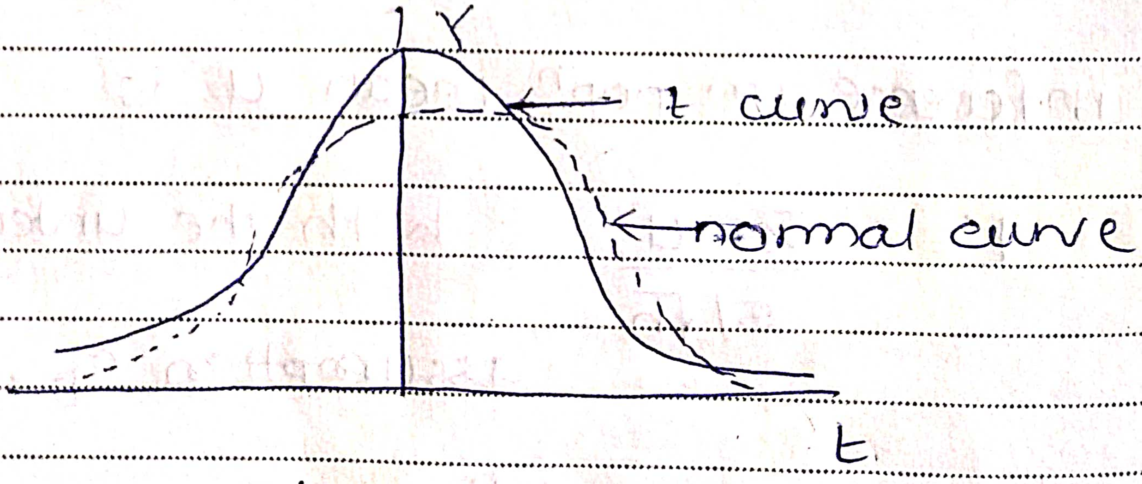
with the underlying assumption of sampling

from normal population-

Date: / /

Properties: -

① This curve is symmetrical about line $t=0$ like normal curve. since only even powers of t appears in ⊛ But it is more peaked than normal curve with same S.D. The t curve approaches horizontal axis less rapidly than normal curve. Also t curve attains the max^m value at $t=0$ so that its mode coincides with mean.



Date: / /

The t-distribution curve is similar to normal curve. While variance for N.D is 1, the variance for t-distribution is more than one since it depends on parameter ν so t distribution is more variable

As $n \to \infty$ the variance of t distribution approaches to 1.

As $\nu = (n-1) \to \infty$ t distribution approaches std. normal distribution. [The shape of t distribution depends on d.f. The curves with more df are taller & have thinner tails] Infact for $n \geq 30$, std. normal distribution provides good approximation to t-distribution.

~~[illegible]~~

iii) The Probability p that value of t will exceed t_0 is $\int_{t_0}^{\infty} y \, dx$. The values of t_0 have been tabulated for various values of p for various values of ν from 1 to 30.

iv] Moments about mean :-

All moments of odd order about mean are zero due to symmetry about line $t=0$. Even ordered moments about mean are

$$\mu_2 = \frac{\nu}{\nu-2}, \quad \mu_4 = \frac{3\nu^2}{(\nu-2)(\nu-4)} \quad \ldots$$

Date: / /

The t distribution is extensively used in test of hypothesis about one mean or about equality of two means when σ is unknown.

* significance Test of sample mean:

Given a random sample $x_1, x_2 \ldots x_n$ from normal population we have to test hypothesis that mean of Pop^n is μ

For this first calculate

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

where $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$ & $s^2 = \dfrac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$

Date: / /

Then find value of P for the df from table.

If calculated value of $t > t_{0.05}$ then difference bet^n $\bar{x}$ & μ is said to be significant at 5% LOS.

If $t > t_{0.01}$ the diff. is said to significant at 1% LOS.

If $t < t_{0.05}$ the data is significant with hypothesis that μ is mean of pop^n.

Date: / /

A machinist is making engine parts with axle diameter of 0.7 inch. A random sample of 10 parts shows mean diameter 0.742 inch with S.D of 0.04 inch. On the basis of this sample would you say that the work is inferior?

→ $n = 10$, $\bar{x} = 0.742$, $S = 0.04$

$\mu = 0.7$

Taking hypothesis that the product is not inferior ie there is no significant difference betn $\bar{x}$ & μ.

$t = \dfrac{\bar{x} - \mu}{S/\sqrt{n}} = 3.16.$

& $\nu = n - 1 = 9$

Date: / /

From table $t_{0.05} = 2.26$.

$t_{cal} = 3.16 > t_{0.05}$

$\therefore$ diff. is significant at 5 % LOS.

so hypothesis is rejected.

ie work is inferior.