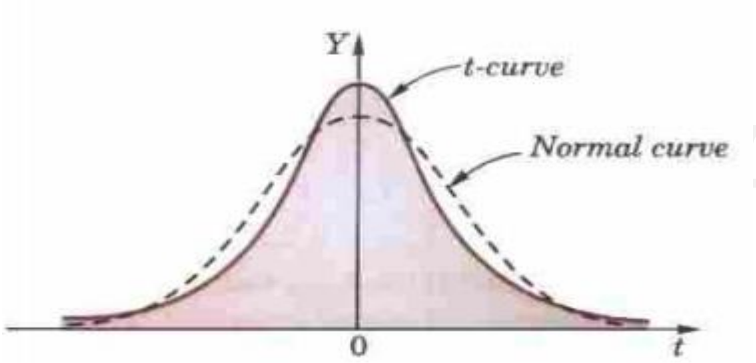
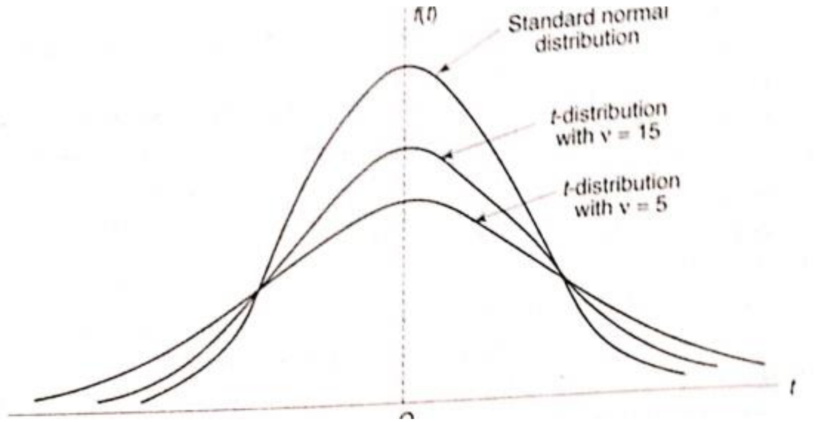
Small Samples

- n<30
- Assumptions made for large samples do not hold for small samples.
- Student's t-distribution is used for small samples.





- The concept of degrees of freedom is used.
- Definition-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 $\nu = n - k$ where n is the number of observations in the sample and k is the number of constraints imposed on them.

Small Sample Test :Steps

- H0: $\bar{x}$=μ or $\bar{\mathbf{x}}_1 = \bar{\mathbf{x}}_2$
- H1: $\bar{x} \neq \mu$ or $\bar{x} > \mu$ or $\bar{\mathbf{x}}_1 \neq \bar{\mathbf{x}}_2$ or $\bar{\mathbf{x}}_1 > \bar{\mathbf{x}}_2$
- Nature of the test is one/two tailed
- Degree of freedom=n-1 or n1+n2-2
- LOS is 5% or 1%
- Table value of t
- $\mathbf{t}_{\mathbf{cal}} = \frac{\bar{\mathbf{x}}-\mu}{\mathbf{s}/\ \sqrt{(\mathbf{n}-\mathbf{1})}}$
- (=or) $\frac{\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2}{\sqrt{\left\{\left(\frac{n_1 s_1{}^2+n_2 s_2{}^2}{n_1+n_2-2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right\}}}$ (samples are independent)
-
- $|\mathbf{t}_{\mathbf{cal}}| < (or >)$ table value of t so H0 (or H1) is accepted
- Conclusion

Problems on Small Sample: Testing Hypothesis
1.Test for Significance between sample mean and population mean

- Degree of freedom= n-1
- $t_{cal} = \frac{\bar{x}-\mu}{s/\ \sqrt{(n-1)}}$
- If sample observations are given

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

Ex1- A random sample of 16 observations has mean 27. The sum of the squares of the deviation from the mean is 140. Can it be a random sample from a population whose mean is 30 ?

Given: n=16, $\bar{x}$=27 , μ=30, $\sum(x-\bar{x})^2 = 140$

1. H0: $\bar{x}$=μ
2. H1: $\bar{x} \neq$ μ (as nature of the test is two tailed)
3. Degree of freedom =16-1=15
4. Let LOS α be 5% so critical value t_α=2.131
5. $t_{cal} = \frac{\bar{x}-\mu}{s/\ \sqrt{(n-1)}}$, $s = \sqrt{\left(\frac{\sum(x-\bar{x})^2}{n}\right)} = \sqrt{\frac{140}{16}} = 2.96$

   $= \frac{-3}{2.96/3.87}$ =3.93

6. $|t_{cal}| = 2.131 > t_\alpha$, null hypothesis is rejected

7. Conclusion: The random sample can not be regarded as the sample taken from the population whose mean is 30 .

Ex2- A company supplies toothpaste in packing of 100 grams. A random sample of 10 packing gave the following weights: 100.5, 100.3, 100.1, 99.8, 99.7, 99.7, 100.3, 100.4, 99.2, 99,3. Does the sample support the claim of the company that the packing weighs 100 grams?

Given: n=100,, μ=100, $\sum(x-\bar{x})^2 = 140$

1. To Calculate $\bar{x} = \frac{\sum x}{n}$, $s = \sqrt{\left(\frac{\sum(x-\bar{x})^2}{n}\right)}$

=99.93 =1.10

2. H0: $\bar{x}$=μ

3. H1: $\bar{x} \neq$ μ (as nature of the test is two tailed)

4. Degree of freedom =10-1=9

5. Let LOS α be 5% so critical value t_α=2.262

6. $t_{cal} = \frac{\bar{x}-\mu}{s/\ \sqrt{(n-1)}}$

=0.48

7. $|t_{cal}| = 0.48 < t_\alpha$, null hypothesis is accepted

8. Conclusion: the sample supports the claim of the company that the packing weighs 100 grams?

Exercise

..

Example *A random sample of size 16 values from a normal population showed a mean of 53 and a sum of squares of deviation from the mean equals to 150. Can this sample be regarded as taken from the population having 56 as mean? Obtain 95% and 99% confidence limits of the mean of the population.*

2.Test for Significance between two independent sample means

Degree of freedom= $n_1 + n_2$-2

(i) If samples standard deviations are given

$$t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left\{\left(\frac{n_1 s_1{}^2 + n_2 s_2{}^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right\}}}$$

(ii) If sample observations are given

$$t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left\{\left(\frac{\sum\{(x_1 - \bar{x}_1)^2 + (x_2 - \bar{x}_2)^2\}}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right\}}}$$

(iii) If unbiased standard deviations for the samples are given

$$t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left\{\left(\frac{(n_1 - 1) s_1{}^2 + (n_2 - 1) s_2{}^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right\}}}$$

(iv) If standard deviation of populations are given, we assume that $\bar{x}_1 - \bar{x}_2$ follow normal distribution with mean 0 and we use z-distribution.

Ex1-Sample of two type of electric bulbs were tested for length of life and following data were obtained: $n_1 = 8, n_2 = 7,\ \bar{x}_1 = 1134hrs,\ \bar{x}_2 =$1024hrs, s_1=35hrs, s_2=40hrs.Test at 5% LOS whether the difference in the sample means is significant.

Given: $n_1 = 8, n_2 = 7,\ \bar{x}_1 = 1134hrs,\ \bar{x}_2 =$1024hrs, s_1=35hrs, s_2=40hrs.

1. H0: $\bar{x}_1 = \bar{x}_2$
2. 2. H1: $\bar{x}_1 \neq \bar{x}_2$(as nature of the test is two tailed)

3. Degree of freedom =8+7-2=13

4. Let LOS α be 5% so critical value t_α=2.16

5.t_{cal} = $\dfrac{\bar{x}_1-\bar{x}_2}{\sqrt{\left\{\left(\frac{n_1 {s_1}^2+n_2 {s_2}^2}{n_1+n_2-2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right\}}} = \dfrac{110}{20.8} = 5.288$

6. $|t_{cal}| = 5.288 > t_\alpha$, null hypothesis is rejected

7. Conclusion: the difference in the sample means is significant.

Ex2- Two independent samples of 8 and 7 items respectively had the following values of the variable (weight in ounces):
sample1:19, 17, 15, 21, 16, 18,16, 14;sample2:15,14,15,19,15,18,16
Is the difference between the means of the sample significant?

$\overline{x1} = \frac{\sum x}{n} = 17, \overline{x2} = \frac{\sum y}{n}$=16 , $s_1{}^2$= $(2.12)^2$, $s_2{}^2$ = $(1.69)^2$

1. H0: $\bar{x}_1 = \bar{x}_2$
2. H1: $\bar{x}_1 \neq \bar{x}_2$ (as nature of the test is two tailed)

3. Degree of freedom =8+7-2=13

DEEPALI PHALAK

4. Let LOS α be 5% so critical value t_α=2.16

5.t_{cal} =t_{cal} = $\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left\{\left(\frac{n_1 s_1{}^2 + n_2 s_2{}^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right\}}} = \frac{17-16}{\sqrt{\left\{\left(\frac{8(2.12)^2 + 7(1.69)^2}{13}\right)\left(\frac{1}{8} + \frac{1}{7}\right)\right\}}} = 0.93$

6. $|t_{cal}| < t_\alpha$, null hypothesis is accepted

7. Conclusion: the difference in the sample means is not significant.

Ex3-The height of six randomly chosen sailors are in inches: 63, 65, 69, 68, 71, 72. The height of ten randomly chosen soldiers are in inches: 61, 62, 65, 66, 69, 69, 70, 71, 72, 73. Can we say the soldiers on an average taller than sailors?

Given: n_1=6, n_2=10 $\bar{x}_1 = \frac{\sum x_1}{n} = \frac{408}{6} = 68$ $\bar{x}_2 = \frac{\sum x_2}{n} = \frac{678}{10} = 67.8$

$$t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left\{\left(\frac{\sum\{(x_1-\bar{x}_1)^2+(x_2-\bar{x}_2)^2\}}{n_1+n_2-2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right\}}} = \frac{68-67.8}{2.014} = 0.099$$

1. H0: $\bar{x}_1 = \bar{x}_2$
2. H1: $\bar{x}_1 < \bar{x}_2$ (as nature of the test is one tailed)
3. Degree of freedom = 6+10-2 = 14
4. Let LOS α be 5% . Since it is one tailed test, we consider critical value t_α for 2 α %=10%.

   t_α =1.761
5. t_{cal} = .099
6. $|t_{cal}|$ = .099< t_α , null hypothesis is accepted
7. Conclusion: We can’t say the soldiers on an average are taller than sailors

Ex4-Two independent samples of size 8 and 10 have the means 950 and 1000 The standard deviation of the two populations are 80 and 100. Test the hypothesis that the populations have same mean.

Given: $n_1 = 8,\ n_2 = 10,\ \bar{x}_1 = 950,\ \bar{x}_2 = 1000\ \sigma_1$=80, σ_2=100

1. H0: $\bar{x}_1 = \bar{x}_2$
2. H1: $\bar{x}_1 \neq \bar{x}_2$(as nature of the test is two tailed)

3. Degree of freedom =8+10-2= 16

4. Let LOS α be 5% so critical value z_α=1.96

5. Formula used $z_{cal} = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}} = 1.178$

6. $|z_{cal}|$ =1.178 < 1.96= z_α for LOS α be 5% , null hypothesis is accepted

7. Conclusion: the populations have same mean. under 5% LOS.

Note : If standard deviation of populations are given, we assume that $\bar{x}_1 - \bar{x}_2$ follow normal distribution with mean 0 and we use z-distribution.

Exercise

1. **Example 85.** *Two independent samples of 8 and 7 items respectively had the following values of the variable (weight in ounces):*

Sample 1: 9 11 13 11 15 9 12 14

Sample 2: 10 12 10 14 9 8 10

Is the difference between the means of the sample significant?

3.Test for Significance between two not independent sample means(**paired t-test**)

- In such cases, sample is same for two tests
- We first find the differences of the corresponding values of the two sets of data then find the mean difference $\bar{x}$ and standard deviation s of the differences. We then define
- $t_{cal} = \frac{\bar{x}-\mu}{s/\sqrt{(n-1)}} = \frac{\bar{x}}{s/\sqrt{(n-1)}}, \; s = \sqrt{\left(\frac{\sum(x-\bar{x})^2}{n}\right)}$
- μ=0 is the null hypothesis and s is the standard deviation of the sample.

Q. Ten boys were given a test in statistics & their scores were recorded. They were given a month's special coaching & a second test was given to them in the same subject at the end of the coaching period. Test if the marks given below give evidence to the fact that coaching benefits the students

Marks in test1 : 70,68,56,75,80,90,68,75,56,58

Marks in test2 : 72,70,58,76,79,92,80,90,54,59

x= difference of the corresponding values of the two sets of data

=2, 2, ,2,1,-1,2,12,15,-2,1

$\bar{x} = \frac{\sum x}{n} = 3.4$; $s = \sqrt{\left(\frac{\sum(x-\bar{x})^2}{n}\right)}$ =5.54176

- H0: $\mu = 0$
- H1:$\mu > 0$
- nature of the test is one tailed
- Degree of freedom =10-1=9
- Let LOS α be 10%
- Table value of t is 1.383 $t_{\alpha(1\ tailed\ test)}$=$t_{2\alpha(2\ tailed\ test)}$
- Formula used t_{cal} =$\frac{\bar{x}}{s/\sqrt{(n-1)}} = 1.84057$
- $|t_{cal}|$ =1.84057< 1.383= t_α for LOS α be 5% , null hypothesis is not accepted
- Conclusion: the coaching benefits the students under 10% LOS.

Ex 2- A certain injection administered to 12 patients resulted in the following change of blood pressure 5,2,8, -1,3,0,6, -2,1,5,0,4. Can be concluded that the injection will be in general accompanied by an increase in blood pressure?

To calculate $\bar{x} = \frac{\sum x}{n}$ $\qquad s = \sqrt{\left(\frac{\sum(x-\bar{x})^2}{n}\right)}$

1. H0: μ=0
2. H1:μ > 0 (as nature of the test is one tailed)

3. Degree of freedom =12-1=11

4. Let LOS α be 5% so critical value for one tailed test= t_α at 10% LOS =1.796

5. Formula used $t_{cal} = \frac{\bar{x}}{s/\sqrt{(n-1)}} = 2.89$

6. $|t_{cal}| = 2.89 > 1.796 = t_\alpha$ for LOS α be 5% , null hypothesis is rejected

7. Conclusion: the injection will be in general accompanied by an increase in blood pressure under 5% LOS.

Exercise

- **1** An experiment was conducted on nine individuals. The experiment showed that due to smoking, the pulse rate increased in the following order:

  $$5,\ 3,\ 4,\ -1,\ 2,\ -3,\ 4,\ 3,\ 1.$$

  Can you maintain that smoking leads to an increase in the pulse rate?