Example 1

In a study of time and motion of factory, the supervisor estimates that the standard deviation to be 0.95 seconds. If you want to be 95% confident that the error will not exceed 0.10 second, what should be the size of the sample to estimate population mean?

Solution:

Standard deviation (SD) = 0.95

Sampling error(E) = 0.10

Level of Confidence (1 - α) = 95%

Sample size (n) = ?

For 1- α = 95%, Z_α = 1.96 from standard normal table.

Sample size(n) = $\left(\frac{Z_\alpha \times SD}{E}\right)^2 = \left(\frac{1.96\times0.95}{0.10}\right)^2 = 346.70 \cong 347$

Classwork

How large sample should be taken to keep the risk of error being ± 5 is 0.0456? It is provided that standard deviation is 20.

Solution

Standard deviation (SD) = 20

Risk (α) = 0.0456

Error (E) = 5

Level of confidence (1 - α) = 1 – 0.0456 = 0.9544 = 95.44%

For (1 - α) = 95.44%, Z_α = 2.00 from standard normal table

Sample size(n) = $\left(\frac{Z_\alpha \times SD}{E}\right)^2 = \left(\frac{2.00\times20}{5}\right)^2 = 64$

Example 2

It is desired to estimate the proportion of the junior executives who change their first job within the first five years. This proportion is to be estimate within 3% of error and 0.95 degree of confidence is to be used. A study revealed that 30% of such junior executives changed their first job within 5 years.

(a) How large a sample is required to update the study?

(b) How large should be the sample if the no such previous estimates are available

Solution:

Level of confidence (1 - α) = 0.95 = 95%

Sampling Error (E) = 3% = 0.03

Previous estimate (P) = 30% = 0.30 and Q = 1 – P = 1 – 0.30 = 0.70

For 1- α = 95%, Z_α = 1.96 from standard normal table.

(a) Sample size(n) = $\left(\frac{Z_\alpha}{E}\right)^2$ x PQ = $\left(\frac{1.96}{0.03}\right)^2$ x 0.30 x 0.70 = 896.37 ≅ 896

(b) When previous estimate is not available, we assume P = Q = 0.5

Then

Sample size(n) = $\left(\frac{Z_\alpha}{E}\right)^2$ x PQ = $\left(\frac{1.96}{0.03}\right)^2$ x 0.50 x 0.50 = 1067.11 ≅ 1067

NUMERICAL EXAMPLES OF CONFIDENCE INTERVAL AND HYPOTHESIS TESTING

Q1)Random sample drawn from two countries given the following data relating to the heights of a adult males:

Males	Country A	Country B
Mean (height in inches)	67.42	67.25
S.d(in inches)	2.58	2.50

No of sample	1000	1200

Obtain 98% confidence interval for $\mu_1 - \mu_2$.

Solution:

Given:

n_1=1000 and n_2=1200 , $\overline{X}_1 = 67.42$ and $\overline{X}_2 = 67.25$

S_1=2.58 and S_2 = 2.50

Now,

$$S.E(\overline{X}_1 - \overline{X}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$S.E(\overline{X}_1 - \overline{X}_2) = \sqrt{\frac{2.58^2}{1000} + \frac{2.50^2}{1200}} = 0.1089$$

$Z_{0.02}$ = 2.33(from standard normal table)

So, confidence interval for difference between two mean is

C. I. for($\mu_1 - \mu_2$) = $(\overline{X}_1 - \overline{X}_2) \pm Z_\alpha \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

C. I. for($\mu_1 - \mu_2$) = (67.42-67.25)± 2.33*0.1089

C. I. for($\mu_1 - \mu_2$) = -0.0837, 0.4237)

Q2)A store keeper wanted to buy a large quantity of light bulbs from two brands labeled A and B. He bought 100 bulbs from each brand and found by testing that brand A had mean life time 1120hrs and S.D 75hrs and brand B had mean life time 1062hrs and SD 82 hrs. Find the 99% confidence limits for the difference in the average life if bulbs from the two brands.

Answer(29.33, 86.67)

Q3)A random sample of 12 records gives the average of 163.99 minutes with standard deviation of 3.043 minutes. Find the 95% confidence limits for population mean if population consists of 100units.

Solution:

Sample size (n)=12 Population size(N)=100 Sample mean($\bar{X}$) = 163.99 minutes

Sample standard deviation(s)=3.043hours

Confidence interval (1-α)=95%

Level of significance (α)=5%=0.05

[Note: in this question sample size is less than 30 so problem is related to t test. S.E($\bar{X}$) = $\frac{s}{\sqrt{n-1}}\sqrt{\frac{N-n}{N}}$

S.E($\bar{X}$) = $\frac{3.043}{\sqrt{12-1}}\sqrt{\frac{100-12}{100}}$ = 0.8606

C.I for $\mu = \bar{X} \pm t_{\alpha, v}$ S.E.($\overline{X}$)

C.I for μ = 163.99 $\pm$ $t_{0.05, 12-1}$ 0.8606

From t-table $t_{0.05, 11}$ = 2.20

C.I for μ =(162.097, 165.883)

Q4)A sample of 20 bulbs, drawn at random from a batch, and discovers that the mean life of the sample bulb is 990hours with a standard deviation of 22hours. Find confidence interval for mean.

Solution Given,

Sample size (n)=20 Sample mean($\bar{X}$) = 990 hours Sample standard deviation(s)=22hours

Confidence interval (1-α)=95% (If not given 95% is taken)

Level of significance (α)=5%=0.05

S.E($\bar{X}$) = $\sqrt{\frac{s^2}{n-1}} = \frac{s}{\sqrt{n-1}} = \frac{22}{\sqrt{20-1}}$ = 5.0471

We find Confidence interval for mean as,

C.I for $\mu = \bar{X} \pm t_{\alpha, v}$ S.E.($\overline{X}$)

C.I for μ = 990 $\pm$ $t_{0.05, 20-1}$ 5.0471

C.I for μ = 990 $\pm$ $t_{0.05, 19}$ 5.0471

From t-table $t_{0.05, 19}$ = 2.093

C.I for μ = 990 $\pm$ 2.093* 5.0471

Solving this you will get answer

Q5)A random sample of 10 bulbs have the following life in months: 24,26, 32, 28,20,20,23,34,30 and 43. Obtain 95% fiducially limit for the population mean.

Solution:

Here sample size (n) = 10

X	X^2
24	576
26	676
32	1024
28	784
20	400
20	400
23	529
34	1156
30	900
43	1849
280	8294

Here, $\sum X=280$ and $\sum X^2 = 8294$ and $\bar{X} = 28$

We have,

Or, $S^2 = \frac{1}{n-1}[\sum X^2 - \frac{(\sum X)^2}{n}]$

Or, $S^2 = \frac{1}{10-1}[8294 - \frac{(280)^2}{10}$

Or, $S^2 = 50.444$

$S=7.102$

Now,

$S.E(\bar{X}) = \sqrt{\frac{S^2}{n}} = \frac{S}{\sqrt{n}} = 7.102/\sqrt{10} = 2.246$

C.I for $\mu = \bar{X} \pm t_{\alpha, v}$ S.E.($\bar{X}$)

C.I for $\mu = 28 \pm t_{0.05, 10-1}$ 2.246

C.I for $\mu = 28 \pm t_{0.05, 9}$ 2.246

From t-table $t_{0.05, 9} = 2.262$

C.I for $\mu = 28 \pm 2.262 * 2.246$

Solving this you will get answer

impQ6)400 laptops were taken as sample from a large consignment of Sony Electronics Company and 20 laptops were found to be damaged

(a)Find the standard error of proportion of damaged laptops.

(b) Estimate 95% and 99% confidence limits for the percentage of damaged laptops in the consignment.

Solution:

Sample size (n) = 400

No. of defective (x) = 20

Sample proportion(p) $= \frac{x}{n} = \frac{20}{400} = 0.05$ and $q = 1 - 0.05 = 0.95$

Level of confidence $(1 - \alpha)$ = 95% and 99%

(a) S.E.(p) = $\sqrt{\frac{pq}{n}} = \sqrt{\frac{0.05 \times 0.95}{400}} = 0.0109$

b)For $1- \alpha$ = 95%, Z_α = 1.96 from standard normal table.

C. I. for P $= p \pm Z_\alpha$ S.E.(p)

$= 0.05 \pm 1.96 \times 0.0109$

$= 0.05 \pm 0.0214$

$= [0.0286, 0.0714]$

Q7)A random sample of 300 first year students of ShankerDev Campus was selected from total 2500 students. The mean and standard deviation of marks in Statistics of these 300 students were found to be 50 and 10. Find 95% confidence interval for average marks of all students.

Population size(N) = 2500

Sample size (n) = 300

Sample mean($\bar{X}$) = 50

Standard Deviation (SD) = 10

Level of confidence $(1 - \alpha)$ = 95%

Level of significance (α) = 5%

Now, $SE(\bar{X}) = \frac{SD}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}} = \frac{10}{\sqrt{300}} \times \sqrt{\frac{2500-300}{2500-1}} = 0.5416$

For $1- \alpha$ = 95%, Z_α = 1.96 from standard normal table.

CI for μ $= \bar{X} \pm Z_\alpha$ S.E.($\bar{X}$)

$= 50 \pm 1.96 \times 0.5416$

$= [50 - 1.96 \times 0.5416, 50 + 1.96 \times 0.5416]$

$= [48.94, 51.06]$

Q8)The average score of 400 students taken as sample from the students appearing in CMAT examination of BBA was found to be 60 with standard deviation 10. Construct 99.73% confidence limit for the average score of all students appearing in that CMAT examination.

Given For $(1 - \alpha)$ = 99.73%, Z_α = 3.00 from standard normal table.

Solution
Sample size (n) = 400
Sample mean($\overline{X}$) = 60
Standard Deviation(SD) = 10
Level of confidence (1 - α) = 99.73%
Level of significance (α) = 1 – 99.73% = 0.27%

S.E.($\overline{X}$) = $\frac{SD}{\sqrt{n}} = \frac{10}{\sqrt{400}} = 0.5$
For (1 - α) = 99.73%, Z_α = 3.00 from standard normal table.
CI for μ = $\overline{X} \pm Z_\alpha$ S.E.($\overline{X}$)
= 60 $\pm$ 3.00 x 0.5
= [60 – 3.00x 0.5, 60 + 3.00x 0.5] = [58.5, 61.5]

Q9)The mean life of a sample of 10 electric bulbs was found to be 1456 hours with standard deviation of 423 hours. A second sample of 17 bulbs chosen from a different batch showed a mean life of 1280 hours and standard deviation 398 hours. Find 95%confidence interval for difference between two means.
Hints:
We have

Batch 1	Batch 2
$n_1 = 10$	$n_2 = 17$
$\overline{X}_1 = 1456$	$\overline{X}_2 = 1280$
$s_1 = 423$	$s_2 = 398$

$(S^2) = \frac{n_1 \times s_1^2 + n_2 \times s_2^2}{n_1 + n_2 - 2}$
$= \frac{10 \times 423^2 + 17 \times 398^2}{10 + 17 - 2}$
= 179286.32

$\text{S.E}(\overline{X}_1 - \overline{X}_2) = \sqrt{S^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

The tabulated value of the t test at α = 5% and df = 10+17-2 =25for two tailed test is 2.06.
$\therefore t_{tab} = 2.06$

Now, Confidence interval for difference between two means;
C. I. for(μ_1 - μ_2) = $(\overline{X}_1 - \overline{X}_2) \pm t_{\alpha, n1+n2-2}$ S.E.($\overline{X}$)
C. I. for(μ_1 - μ_2) = $(\overline{X}_1 - \overline{X}_2) \pm t_{\alpha, n1+n2-2} \sqrt{S^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

Q10)The following are the Brinell hardness values obtained for samples of two magnesium alloysFind 95%confidence interval for difference between two means.
.Alloy 1: 107, 148, 123, 165, 102, 119,
Alloy 2: 134, 115, 112, 151, 133, 129
Answer(-7.0707, 3.7367)

Q11)The mean and standard deviation calculated from a random sample of 100 units were found to be 980 and 150 respectively. Test at 5% level of significance, whether the mean of population is equal to 1000 or not.

Population mean (μ) = 1000 [test]
Sample size (n) = 100
Sample mean ($\overline{X}$) = 980
Sample SD (s) = 150
Level of significance (α) = 5%
[Key word: equal Two tailed]

Solution: Given
. Setting of hypothesis:
$H_0 : \mu = 1000$ [The mean of the population is equal to 1000.]
$H_1 : \mu \neq 1000$ [The mean of the population is not equal to 1000.]
Level of significance : It is given as α = 5%.
Test Statistic: The test statistic under H_0 is

$$Z = \frac{\overline{X} - \mu}{\sqrt{\frac{s^2}{n}}} = \frac{980 - 1000}{\sqrt{\frac{150^2}{100}}} = -1.33$$

$\therefore |Z_{cal}| = |-1.33| = 1.33$

Critical value : The tabulated value of the Z test for two tailed test at α = 5% is ± 1.96.

$\therefore |Z_{tab}| = |\pm 1.96| = 1.96$

Decision: Since calculated value of Z is less than tabulated value of Z, H_0 is accepted which means the mean of the population is equal to 1000.

Q12)Does an average box of cereal contain more than 368grams of cereal? A random sample of 25 boxes showed a sample mean of 372.5 grams. The company has specified σ to be 15 grams and the distribution to be normal. Test at α=0.05 level.

Solution:

Step 1.H_0 : μ =368[The average box of cereal contains 368 grams of cereal .]

Step 2.H_1 : μ >368[The average box of cereal contains more than 368 grams of cereal. (Right tailed test)

Step 3.Test Statistic: The test statistic under H_0 is

$$Z= \frac{\bar{X}-\mu}{\sqrt{\frac{S^2}{n}}} = \frac{372.5-368}{\sqrt{\frac{15}{25}}} = 1.5$$

Step 4:Critical value : The tabulated value of the Z test for left tailed test at α= 5% for one tailed test is 1.645.

$\therefore |Z_{tab}| = 1.645$

Step 5: Decision: Since calculated value of Z is less than tabulated value of Z, H_0 is accepted and H_1 is accepted . Therefore, we conclude that average box of cereal does not contain more than 368 grams of cereal.

Q13)A certain college conducts both morning and evening classes intended to be identical. A random sample of 200 morning class students yield examination result as average score of 72.4 with a standard deviation of 14.8. A random sample of 100 evening class students yield examination result as average score of 73.9 with a standard deviation of 17.9. Are the average score of morning and evening classes equal at 5% level of significance?

Solution: We have

Morning class	Evening class
$n_1 = 200$	$n_2 = 100$
$\bar{X}_1 = 72.4$	$\bar{X}_2 = 73.9$
$S_1 = 14.8$	$S_2 = 17.9$

[Key Word: Equal Two tailed]

Step 1.H_0 :$\mu_1 = \mu_2$ [The average score of morning class and evening class are equal.]

Step 2.H_1 :$\mu_1 \neq \mu_2$ [The average score of morning class and evening class are not equal.]

Step 3: Test Statistic:

The test statistic under H_0 is

$$Z= \frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} = \frac{72.4-73.9}{\sqrt{\frac{14.8^2}{200}+\frac{17.9^2}{100}}} = -0.72$$

Step 4.Critical value : The tabulated value of the Z test for two tailed test at α = 5% is ± 1.96.

$\therefore |Z_{tab}| = |\pm 1.96| = 1.96$

Step 5 Decision: Since calculated value of Z is less than tabulated value of Z, H_0 is accepted which means the average score of morning class and evening class are equal

.Q14)A consumer research organization selects several car models each year and evaluates their fuel efficiency. In this years study of two similar subcompact models from two different automakers, the average gas mileage for 40cars of brand A was 37.2 miles per gallon(mpg) and the standard deviation was 3.8 mpg. The 50 brand B cars that were tested average 32/mpg and the standard deviation was 4.3mpg. At α= 0.01 should it conclude that brand A cars have higher average gas mileage than that of brand B.

Solution:

Here

$n_1 = 40$ $n_2 = 50$

$\overline{X}_1 = 37.2$ mpg $\overline{X}_2 = 32.1$ mpg

$S_1 = 3.8$mpg $S_2 = 4.3$ mpg

H_0 : $\mu_1 = \mu_2$ [There is no significance difference between average mileage in brand A and brand B cars.]

H_1 : $\mu_1 > \mu_2$ [The brand A cars have higher average than that of brand B] Right tailed test

Step 3: Test Statistic:

The test statistic under H_0 is

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{37.2 - 32.1}{\sqrt{\frac{3.8^2}{40} + \frac{4.3^2}{50}}} = 5.97$$

Step 4.Critical value : The tabulated value of the Z test for one tailed test at α = 1% is 2.33

$\therefore |Z_{tab}| = 2.33$

Step 5 Decision: Since calculated value of Z is greater than tabulated value of Z, H_0 is rejected and null hypothesis is accepted

Q15)A sample of 600 persons selected randomly from a large city gives the result that males are 53%. Is there reason to doubt the hypothesis that males and females are equal number in the city.

Solution:

Number of peoples in a sample(n)=600

Sample proportion of males (p)=0.53

Analysis steps are:

Step 1: H_0 : P = 0.5 i.e. the males and females are in equal number in the city.

Step 2: H_1 : P ≠ 0.5 i.e.the males and females are not in equal number in the city.

Step 3: Test Statistics:

Under H_o the test statistics is given by

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.53 - 0.5}{\sqrt{\frac{0.5*0.5}{600}}} = 1.469$$

.

Step 4: Critical value:

Here we take we take as α=0.05 as not mentioned.

Tabulated value of Z at α=5% level of significance for two tailed test is 1.96.

Step 5: Decision:

Since $|Z_{tab}| < |Z_{cal}|$ then H_o is rejected and H_1 is accepted

Q16: It is claimed that both tea and coffee are equally popular in Ilam district. If in a random sample of 1200 persons 650 were regular consumers of tea. Is the claim justified at 5% level of significance.

Solution:

Sample size(n) =1200

No of success (X) =650

Sample proportion of success (p) = 650/1200 = 0.542

Level of significance(α) = 0.05

Population proportion (P)=50% (as not given)

Analysis steps are:

Step 1: H_0 : P = 0.5 i.e. the tea and coffee are equally popular in Ilam district.

Step 2: H_1 : P ≠ 0.5 i.e.the tea and coffee are not equally popular in Ilam district

Step 3: Test Statistics:

Under H_o the test statistics is given by

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.542 - 0.5}{\sqrt{\frac{0.5*0.5}{1200}}} = 2.91$$

Step 4: Critical value:Tabulated value of Z at α=5% level of significance for two tailed test is 1.96.

Step 5: Decision:Since $|Z_{tab}| < |Z_{cal}|$ then H_o is rejected and H_1 is accepted.

ImpQ17) Example: A coin is tossed 800 times and heads appear 480 times. Can you infer that the coin is unbiased at 1% level of significance.

Solution:

Number of trials, n=800

Observed number of heads, X=480

Sample proportion of heads (p) =x/n = 480/800 = 0.60

Population proportion of heads (P) =0.5(assumed)

Step 1: H_0 : P = 0.5 i.e. the coin is unbiased.

Step 2: H_1 : P ≠ 0.5 i.e.. the coin is not unbiased.

Step 3: Test Statistics:

Under H_o the test statistics is given by

$$Z = \frac{p-P}{\sqrt{\frac{PQ}{n}}} = \frac{0.60-0.50}{\sqrt{\frac{0.50*0.50}{800}}} = 5.65$$

Step 4: Critical value: Tabulated value of Z at α=1% level of significance for two tailed test is 2.578.

Step 5: Decision: Since $|Z_{tab}|<|Z_{cal}|$ then H_o is rejected and H_1 is accepted

. Q18)The mean and standard deviation calculated from a random sample of 10 units were found to be 980 and 150 respectively. Test at 5% level of significance, whether the mean of population is equal to 1000 or not.

Solution: We have

Population mean(μ) = 1000{ test}

Sample size (n) = 10

Sample mean($\overline{X}$) = 980

Sample SD(s) = 150 and

sample variance (s^2) = 22500

Level of significance(α) = 5%

[Key word : equal to = two tailed]

Setting of hypothesis:

Step 1.H_0 : μ =1000 [The mean of the population is equal to 1000.]

Step 2.H_1 : μ ≠1000 [The mean of the population is not equal to 1000.]

Step 3.Test Statistic:

The test statistic under H_0 is

$$t = \frac{\overline{X}-\mu}{\sqrt{\frac{s^2}{n-1}}} = \frac{980-1000}{\sqrt{\frac{22500}{10-1}}} = -0.40$$

∴$|t_{cal}|$ = | -0.40| = 0.40

and df = n – 1 = 10 – 1 = 9

Step 4. Critical value :Level of significance : It is given as α = 5%.

Degree of freedom =n-1=10-1=0

The tabulated value of the t test for two tailed test at α = 5% and df = 9 is 2.262.

∴t_{tab} = 2.262

Step5. Decision: Since calculated value of t is less than tabulated value of t, H_0 is accepted

Q19)A manufacture intends that has electric light bulbs have a life of 1000hrs. He tests a sample of 20 bulbs, drawn at random from a batch and discovers that the mean life of the sample bulb is 990hrs with a standard deviation of 22 hours. Does this signify the batches not up to standard?

Solution:

: We have

Population mean(μ) = 1000hrs{ test}

Sample size (n) = 20

Sample mean($\overline{X}$) = 99hrs

Sample SD(s) = 22hrs and sample variance (s^2) = 484

Level of significance(α) = 5%

[Key word :upto standard = one tailed]

Setting of hypothesis:

Step 1.H_0 : μ =1000hrs [The population mean of the life of electric light bulbs is equal to 1000hrs.

Step 2.H_1 : μ <1000 [The population mean of the life of electric light bulbs is less than 1000hrs.. IN other words the batch is not up to standard.]

Step 3.Test Statistic: The test statistic under H_0 is

$$t = \frac{\overline{X}-\mu}{\sqrt{\frac{s^2}{n-1}}} = \frac{990-1000}{\sqrt{\frac{484}{20-1}}} = -1.98$$

∴$|t_{cal}|$ = | -1.98| = 1.98 and

df = n – 1 = 20 – 1 =19

–NEC)

Step 4. Critical value :

Level of significance : It is given as α = 5%.

Degree of freedom =n-1=10-1=0

The tabulated value of the t test for left tailed tailed test at α = 5% and df = 19 is -1.729.

∴$|t_{tab}|$ = 1.729

Step5. Decision: Since calculated value of t is greater than tabulated value of t, H_0 is rejected and hence alternative hypothesis is accepted.There fore we have sufficient evidence to conclude that the batch is not up to standard.

Q20) A company tells that the mean life of its electric light bulbs is 28 months. A random sample of 10 bulbs has the following life in months: 24,26,32,28,20,20,23,34,30 and 43.
Test whether the mean life of the electric bulbs of the company is 28 or not at 5% level of significance

Solution: (Since data is given we should calculate S.)
Population mean(μ) = 28 [test]
Sample size (n) = 10
Calculation of sample mean ($\overline{X}$) and Sample variance(S^2):
Here, $\sum X=280$ and $\sum X^2 = 8294$ and $\bar{X} = 28$
We have,

Or, $S^2 = \frac{1}{n-1}\left[\sum X^2 - \frac{(\sum X)^2}{n}\right]$

Or, $S^2 = \frac{1}{10-1}\left[8294 - \frac{(280)^2}{10}\right]$

Or, $S^2 = 50.444$
S=7.102

Setting of hypothesis:
Step 1.H_0 : μ =28 [The population mean of the life of electric l bulbs is equal to 28months.]
Step 2.H_1 : $\mu \neq 28$ [The population mean of the life of electric light bulbs is not equal to 28 months..]

Step 3.Test Statistic: The test statistic under H_0 is

$$t = \frac{\overline{X}-\mu}{\sqrt{\frac{S^2}{n}}} = \frac{28-28}{\sqrt{\frac{50.44}{10}}} = 0$$

$\therefore |t_{cal}|$ = 0 and df = 10– 1 = 9

Step 4. Critical value :
Level of significance : It is given as α = 5%.
Degree of freedom =n-1=10-1=9
The tabulated value of the t test for two tailed tailed test at α = 5% and df = 9 is 2.262
$\therefore |t_{tab}| = 2.262$
Step5. Decision: Since calculated value of t is less than tabulated value of t, H_0 is accepted

Q21Sample of two types of electric light bulb were tested for length of life and following data were obtained:

	Type I	Type II
Sample no	8	7
Sample means	1234hrs	1036hrs
Sample S.D	36hrs	40hrs

Is the difference in the means sufficient to warrant that Type I is superior to type II regarding length of life:

Solution:
Given
$n_1 = 8$ $n_2 = 7$
$\overline{X}_1 = 1234$ $\overline{X}_2 = 1036$
$s_1 = 36$ $s_2 = 40$

$(S^2) = \frac{n_1 \times s_1^2 + n_2 \times s_2^2}{n_1 + n_2 - 2} = 1659.08$

[Key word: superior = one tailed]

H_0 : $\mu_1 = \mu_2$ [There is no significance difference between mean of two type of electric bulbs..]
H_1 : $\mu_1 > \mu_2$ [Mean life of Type I bulb is significantly greater than type II.

Test Statistic: The test statistic under H_0 is

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{1234-1036}{\sqrt{1659.08\left(\frac{1}{8}+\frac{1}{7}\right)}} = 198/21.08=9.39$$

Critical Value:
Level of significance: Since it is not given, we can assume α = 5%.
Degree of freedom(d.f) =$n_1 + n_2 - 2$ =8+7-2=13
The tabulated value of the t test at α = 5% and df = 13 for one tailed test is 1.771
$\therefore t_{tab} = 1.771$

Decision: Since calculated value of t is greater than tabulated value of t, H_0 is rejected and alternative hypothesis is accepted which means Mean life of Type I bulb is significantly greater than type II.

Q22)The following table gives marks in Statistics scored by 10 boys and 8 girls selected randomly from a college appearing in final examination of MBS first year.

Boys	65	60	70	75	55	45	52	50	48	59
Girls	60	75	66	58	49	50	67	70		

Can you conclude that the average score of boys is lower than that of girls at 1% level of significance?

Solution:

Calculation of sample means and standard deviations:

	Boys		Girls	
	X_1	X_1^2	X_2	X_2^2
	65	4225	60	3600
	60	3600	75	5625
	70	4900	66	4356
	75	5625	58	3364
	55	3025	49	2401
	45	2025	50	2500
	52	2704	67	4489
	50	2500	70	4900
	48	2304	-	-
	59	3481	-	-
Total	579	34389	495	31235

$n_1 = 10$

$\Sigma X_1 = 579$

$\Sigma X_1^2 = 34389$

$\overline{X}_1 = \frac{\Sigma X_1}{n_1} = \frac{579}{10} = 57.9$

$n_2 = 8$

$\Sigma X_2 = 495$

$\Sigma X_2^2 = 31235$

$\overline{X}_2 = \frac{\Sigma X_2}{n_2} = \frac{495}{8} = 61.88$

If data are given then

$S^2 = \frac{1}{n_1+n_2-2}[\sum X_1^2 - \frac{(\sum X_1)^2}{n_1} + \sum X_2^2 - \frac{(\sum X_2)^2}{n_2}]$

Putting value we get $S^2 = 91.98$

Key word: Lower = one tailed

Setting of Hypotheis:

H_0 : $\mu_1 = \mu_2$ [The average score of boys and girls are equal]

H_1 : $\mu_1 < \mu_2$ [the average score of boys is lower than that of girls.]

. Level of significance: It is given as α = 1%.

Test Statistic: The test statistic under H_0 is

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{57.9 - 61.88}{\sqrt{91.98\left(\frac{1}{10}+\frac{1}{8}\right)}} = -0.87$$

$\therefore |t_{cal}| = |-0.87| = 0.87$ and df $= n_1 + n_2 - 2 = 10 + 8 - 2 = 16$

Critical Value: The tabulated value of the t at α = 1% and df = 16 for left/one tailed test is 2.583.

$\therefore t_{tab} = 2.583$

. Decision: Since calculated value of t is less than tabulated value of t, H_0 is accepted which means the average score of boys is not lower than that of girls.

Q23)Two different types of drugs A and B were administered on certain patients for increasing weight at interval of one week time period. From the following observation can you conclude that second drug is more effective in increasing weight. Use 1% level of significance.

A	8	12	13	9	3	8	10	9
B	10	8	12	15	6	11	12	12

Solution:

Hints:

Key word: effective = one tailed

Setting of Hypotheis:

H_0 : $\mu_1 = \mu_2$ [Both drugs are equally effective]

H_1 : $\mu_1 < \mu_2$ [Second drug is more effective than first drug.]

. Level of significance: It is given as α = 1%.

Test Statistic: The test statistic under H_0 is

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

If data are given then

$S^2 = \frac{1}{n_1+n_2-2}[\sum X_1^2 - \frac{(\sum X_1)^2}{n_1} + \sum X_2^2 - \frac{(\sum X_2)^2}{n_2}]$

Do as above.

Q24)A random sample of nine students was selected to test for the effectiveness of a special course designed to improve memory. The following table gives the results of a memory test given to those students before and after this course.

Before	43	57	48	65	81	49	38	69	58
After	49	56	55	77	89	57	36	64	69

Test at the 1% level of significance whether this course makes any significant improvement in the average memory of all students.

Solution:

Calculation of $\bar{d}$ and (S^2)

	Before(X_1)	After(X_2)	$d = X_1 - X_2$	d^2
	43	49	-6	36
	57	56	1	1
	48	55	-7	49
	65	77	-12	144
	81	89	-8	64
	49	57	-8	64
	38	36	2	4
	69	64	5	25
	58	69	-11	121
Total			-44	508

n = 9, $\Sigma d = 44$ and $\Sigma d^2 = 508$

$$\bar{d} = \frac{\Sigma d}{n} = \frac{44}{9} = 4.89$$

$$S^2 = \frac{1}{n-1}[\Sigma d^2 - \frac{(\Sigma d)^2}{n}]$$

$$S^2 = \frac{1}{9-1}[508 - \frac{44^2}{9}] = 36.61$$

Key word: Improvement = one tailed

Setting of Hypothesis:

$H_0 : \mu_x = \mu_y$ [This course does not make significant improvement in the average memory of all students.]

$H_1 : \mu_x < \mu_y$ [This course makes significant improvement in the average memory of all students.]

. Level of significance: It is given as α = 1%.

Test Statistic: The test statistic is

$$t = \frac{\bar{d}}{\sqrt{\frac{S^2}{n}}} = \frac{4.89}{\sqrt{\frac{36.61}{9}}} = 2.42$$

$\therefore t_{cal}$ = 2.42 and df = n – 1 = 9 - 1 = 8

Critical Value: The tabulated value of t test at α = 1% and df = 8 for right/one tailed test is 2.896.

$\therefore t_{tab}$ = 2.896

Decision: Since calculated value of t is less than tabulated value t, H_0 is accepted which means this course does not make significant improvement in the average memory of all students.

Q25) An I.Q test was administeredto 5 persons before and after they were given the nourishing food Horlicks. The results are given below.

Candidates	I	II	III	IV	V
I.Q before Horlicks	110	120	123	132	125
I.Q after Horlicks	120	118	125	136	121

Test whether there is any change in I.Q after the Horlicks at 1% level of significance.

Solution:

Sample size(n)=5 ,

IQ before (X)	I.Q after (Y)	d=X-Y	d^2
110	120	-10	100

120	118	2	4
123	125	-2	4
132	136	-4	16
125	121	4	16
		$\sum d$=-10	$\sum d^2$=140

n = 5, Σd = 10 and Σd^2 = 140

$\bar{d} = \frac{\Sigma d}{n} = \frac{-10}{5}$ = -2

$S^2 = \frac{1}{n-1}[\sum d^2 - \frac{(\sum d)^2}{n}]$

$S^2 = \frac{1}{5-1}[140 - \frac{10^2}{5}]$= 30

Setting of Hypothesis:

H_0 : $\mu_x = \mu_y$ [There is no significant improvement in I.Q before and after Horlicks..]

H_1 : $\mu_x \neq \mu_y$ [.There is significant improvement in I.Q before and after Horlicks..]two tailed test

Test Statistic: The test statistic is

$t = \frac{\bar{d}}{\sqrt{\frac{S^2}{n}}} = \frac{-2}{\sqrt{\frac{30}{5}}}$ = -0.816

$\therefore |t_{cal}|$ = 0.816

Critical Value:

Level of significance: It is given as α = 1%.df = n – 1 = 5 - 1 = 4

The tabulated value of t test at α = 1% and df = 4for two tailed test is 4.6 $\therefore t_{tab}$ = 4.6

Decision: Since calculated value of t is less than tabulated value t, H_0 is accepted which means there is no significant improvement in I.Q after Horlicks was given to students.

Q26) Sales executed by Eight sales executive trainees in thousand of rupees before and after the training, in the same period are listed below,

Sales before training	23	20	19	21	18	20	18	17
Sales after training	25	25	24	24	22	23	25	21

Do these data indicate that the training has contributed to their performance.?

Solution

Hints:Setting of Hypothesis:

H_0 : $\mu_x = \mu_y$ [There is no significant difference in average sales before and after advertisement. This means the advertisement was not effective.]

H_1 : $\mu_x < \mu_y$ [There is significant difference in average sales before and after advertisement. This means the advertisement was effective.]

Do as above.

Q27)In a study of time and motion of factory, the supervisor estimates that the standard deviation to be 0.95 seconds. If you want to be 95% confident that the error will not exceed 0.10 second, what should be the size of the sample to estimate population mean?

Solution:

Standard deviation (SD) = 0.95

Sampling error(E) = 0.10

Level of Confidence (1 - α) = 95%

Sample size (n) = ?

For 1- α = 95%, Z_α = 1.96 from standard normal table.

Sample size(n) = $\left(\frac{Z_\alpha \times SD}{E}\right)^2$ = $\left(\frac{1.96\times0.95}{0.10}\right)^2$ = 346.70 ≅ 347

How large sample should be taken to keep the risk of error being ± 5 is 0.0456? It is provided that standard deviation is 20.

Solution

Standard deviation (SD) = 20

Risk (α) = 0.0456

Error (E) = 5

Level of confidence (1 - α) = 1 – 0.0456 = 0.9544 = 95.44%

For (1 - α) = 95.44%, Z_α = 2.00 from standard normal table

Sample size(n) = $\left(\frac{Z_\alpha \times SD}{E}\right)^2$ = $\left(\frac{2.00\times20}{5}\right)^2$ = 64

Q28)It is desired to estimate the proportion of the junior executives who change their first job within the first five years. This proportion is to be estimate within 3% of error and 0.95 degree of confidence is to be used. A study revealed that 30% of such junior executives changed their first job within 5 years.

(a) How large a sample is required to update the study?

(b) How large should be the sample if the no such previous estimates are available

Solution:

Level of confidence (1 - α) = 0.95 = 95%

Sampling Error (E) = 3% = 0.03

Previous estimate (P) = 30% = 0.30 and Q = 1 – P = 1 – 0.30 = 0.70

For 1- α = 95%, Z_α = 1.96 from standard normal table.

(a) Sample size(n) = $\left(\frac{Z_\alpha}{E}\right)^2$ x PQ = $\left(\frac{1.96}{0.03}\right)^2$ x 0.30 x 0.70 = 896.37 ≅ 896

(b) When previous estimate is not available, we assume P = Q = 0.5

Then

Sample size(n) = $\left(\frac{Z_\alpha}{E}\right)^2$ x PQ = $\left(\frac{1.96}{0.03}\right)^2$ x 0.50 x 0.50 = 1067.11 ≅ 1067

VimpQ29)In a random sample of 600 and 1000 men selected from two cities - Birgung and Bharatpur, 400 and 600 men were found to be literate. Do the data indicate at 1% level of significance that two cities are significantly different in the percentage of literacy?

Solution: We have

Birgunj	Bharatpur
n_1 = 600	n_2 = 1000
x_1 = 400	x_2 = 600
p_1 = 400/600 = 0.67	p_2 = 600/1000 = 0.60

Combined proportion $(\widehat{P}) = \frac{n_1\times p_1+n_2\times p_2}{n_1+n_2} = \frac{x_1+x_2}{n_1+n_2} = \frac{400+600}{600+1000}$ = 0.625

$\widehat{Q}$ = 1 - $\widehat{P}$ = 1 – 0.625 = 0.375

[Key word : different =Two tailed]

Setting of Hypothesis:

Step 1:H_0 : $P_1 = P_2$ [Two cities are not significantly different in the percentage of literacy.]

Step 2: H_1 : $P_1 \neq P_2$ [Two cities are significantly different in the percentage of literacy.]

Step 3: Test Statistic: The test statistic under H_0 is $Z = \frac{p_1-p_2}{\sqrt{\widehat{P}\widehat{Q}(\frac{1}{n_1}+\frac{1}{n_2})}} = \frac{0.67-0.60}{\sqrt{0.625\times0.375\times(\frac{1}{600}+\frac{1}{1000})}}$ = 2.68

$\therefore Z_{cal} = 2.68$

Step 4: Critical Value: The tabulated value of Z test at α = 1% for two tailed test is 2.57.

Step 5: Decision Since calculated value of Z is greater than tabulated value of Z, H H_0 is rejected which means two cities are significantly different in the percentage of literacy.

Q30 v imp: : One thousand articles from factory A were examined and 97% were found to be of good quality. Fifteen hundred similar articles from factory B were examined and 98% were found to be of good quality. Would you conclude that the products of factory B are superior to those of factory A?

Solution: We have

	Factory A	Factory B
	$n_1 = 1000$	$n_2 = 1500$
	$p_1 = 97\% = 0.97$	$p_2 = 98\% = 0.98$

Combined proportion $(\widehat{P}) = \frac{n_1 \times p_1 + n_2 \times p_2}{n_1 + n_2} = \frac{1000\times 0.97 + 1500\times 0.98}{1000+1500} = 0.976$

$\widehat{Q} = 1 - \widehat{P} = 1 - 0.976 = 0.024$

[Key word : Superior one tailed

Setting of Hypothesis:

Step 1: $H_0 : P_1 = P_2$ [The products of factory B and are not different.]

Step 2: $H_1 : P_1 < P_2$ [The products of factory B are superior to those of factory A.]

Step 3: Test Statistic: The test statistic under H_0 is

$$Z = \frac{p_1 - p_2}{\sqrt{\widehat{P}\widehat{Q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Do as above.

Q31 Random sample of 250 bolts manufactured by machine A and 200 bolts manufactured by machine B showed 24 and 10 defectives respectively. Test the hypothesis that the machines are showing different qualities of performance. Use 5% level of significance.

Hints: $n_1 = 250$, $n_2 = 200$, $x_1 = 24$ and $x_2 = 10$ nnUse Combined proportion ($\widehat{P} = \frac{x_1 + x_2}{n_1 + n_2}$ And $\widehat{Q} = 1 - \widehat{P}$

Step 1:$H_0 : P_1 = P_2$ [There is no significantly difference between the population proportion of defective bolts manufactured by two machines A and B respectively.

Step 2: $H_1 : P_1 \neq P_2$ [There is no significantly difference between the population proportion of defective bolts manufactured by two machines A and B respectively

#Shortcut value of Ztable for Z test at 1%. 5% and 10%

Critical Values	Level of significance		
	1 %	5%	10 %
Two tail test	2.575	1.96	1.645
Right Tail Test	2.33	1.645	1.28
Left Tail Test	-2.33	-1.645	-1.28