mood-book

UNIT-4

ESTIMATION

Estimate: An estimate is a statement made to find unknown population parameter

Estimator: It is a method or rule to determine an unknown population parameter.

Ex: sample mean ($\bar{x}$) is an parameter of population mean(μ) i.e sample mean is a method of determining population mean.

Point estimate: If an estimate of a population parameter is given by a single value is known as point estimate.

Ex: The weight of student is measured as 60kg. This measurement gives a point estimate.

Interval estimate: If an estimate of a population parameter is given by two different values between which the parameter may be considered to lie is known as Interval estimate.

Ex: weight is given as (60 ± 3)kg. Then the weight of the students lies in between 57kgs and 63kgs and this measurement gives interval estimate.

→ sample mean $\bar{x}$ is a point estimate of population mean μ

→ sample variance s^2 is a point estimate of population variances σ^2

→ A point estimate is denoted by 'θ'

→ A point estimator is a statistic for estimating population parameter θ will be denoted by "$\hat{\theta}$"

unbiased estimator

let $\hat{\theta}$ be an estimator of θ. Then $\hat{\theta}$ is said to be unbiased estimator of the parameter θ.

If $\boxed{E(\hat{\theta}) = \theta}$ otherwise the estimator is said to be biased.

prove that the sample mean is an unbiased estimator of population mean.

Let $x_1, x_2 \ldots x_n$ be a random samples of size n drawn from a population with mean μ and variance σ^2. Then

$$E(\bar{x}) = E\left(\sum_{i=1}^{n} \frac{x_i}{n}\right)$$

$$= \frac{1}{n} E\left(\sum_{i=1}^{n} x_i\right)$$

$$= \frac{1}{n} E\left(x_1 + x_2 + \ldots + x_n\right]$$

$$= \frac{1}{n}\left[E(x_1) + E(x_2) + \ldots + E(x_n)\right]$$

$= \frac{1}{n}[\mu + \mu + \cdots + \mu (n \text{ times})]$

$= \frac{n\mu}{n} = \mu$

$\therefore E(\bar{x}) = \mu$

$(1-\alpha) \times 100\%$ is confidence interval.

Here $(1-\alpha)$ is degrees of confidence or coefficient of confidence.

→ when $\alpha = 0.05$ we have 95% confidence limit then $Z_{\alpha/2} = 1.96$

$\alpha = 0.01$ we have 99% confidence then $Z_{\alpha/2} = 2.58$

$\alpha = 0.1$ we have 90% confidence then $Z_{\alpha/2} = 1.64$

for 99.73% confidence then $Z_{\alpha/2} = 3$

confidence limits for population mean(μ):

95% confidence limits are $\bar{x} \pm 1.96$ (S.E of $\bar{x}$)

99% confidence limits are $\bar{x} \pm 2.58$ (S.E of $\bar{x}$)

90% confidence limits are $\bar{x} \pm 1.64$ (S.E of $\bar{x}$)

99.73% confidence limits are $\bar{x} \pm 3$ (S.E of $\bar{x}$)

confidence limits for population proportion P:

95% confidence limits are $P \pm 1.96$ (S.E of P)

99% confidence limits are $P \pm 2.58$ (S.E of P)

90% confidence limits are $P \pm 1.64$ (S.E of P)

99.73% confidence limits are $P \pm 3$ (S.E of P)

→ IIIly for $\bar{x}_1 - \bar{x}_2$, $P_1 - P_2$ also same as above

Sample size for estimating population Mean:

Let $\bar{x}$ be the mean of random sample drawn from a population having mean μ and S.D σ

then $n = \left(\frac{Z\sigma}{E}\right)^2$

→ Sample size for estimating population proportion. $n = \frac{Z^2 PQ}{E^2}$ where $Q = 1-P$

confidence interval for μ, σ known

large scale:

If $\bar{x}$ is a mean of random sample size n from the population with variance σ^2 $(1-\alpha)\times 100\%$ confidence interval for μ is given by

$$\bar{x} - Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{x} + Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

i.e confidence interval

$$\left(\bar{x} - Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right), \bar{x} + Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)\right)$$

Maximum Error $E = Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$

$(\bar{x} \pm E)$

Case(ii): For small samples:

If $\bar{x}$ and s are the mean and S.D of random samples from a normal population with unknown variance σ^2. Then $(1-\alpha)\times 100\%$ confidence interval for μ is given by

$$\bar{x} - t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right) < \mu < \bar{x} + t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right)$$

i.e confidence interval is

$$\left(\bar{x} - t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right), \bar{x} + t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right)\right)$$

Maximum Error $E = t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right)$

Sample size for estimating population Mean is $n = \left(\frac{Z\sigma}{E}\right)^2$

→ sample size for estimating population proportion

$$n = \frac{Z^2 PQ}{E^2}$$

A sample of size 300 was taken when variance is 225 and mean 54. construct 95% confidence interval for the mean.

Since the sample size $n = 300$ is greater than 30 Hence the normal distribution is used as sampling distribution.

$n = 300$

$\bar{x} = 54$

$\sigma^2 = 225$

$\sigma = \sqrt{225} = 15$

S.E of $\bar{x} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{300}} = 0.866$

95% → $Z_{\alpha/2} = 1.96$

95% confidence interval for mean is

$$\left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

$(54-(1.96)(0.866), 54+(1.96)(0.866))$

$\therefore$ required confidence limits $(52.30, 55.697)$

A population random variable has mean 100 and S.D 16. What are the mean and S.D for the random sample of size 4, drawn with replacement

$\mu = 100$

$\sigma = 16$

$n = 4$

Since the sampling distribution is with replacement hence it is an infinite population

sample mean $\mu_{\bar{x}} = \mu$

$= 100$

$S.D.\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{16}{\sqrt{4}} = \frac{16}{2} = 8$

The mean and S.D of a population are 11795 and 14,054 respectively. If $n=50$ find 95% confidence interval for mean.

$n = 50$

$\bar{x} = 11795$

$\sigma = 14054$

$S.E \text{ of } \bar{x} = \frac{\sigma}{\sqrt{n}} = \frac{14054}{\sqrt{50}} = 1987.5357$

95% confidence interval for mean is

$$\left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

$(11795-(1.96)(1987.535), 11795+(1.96)(1987.535))$

∴ Required confidence interval is

(7899.4314, 15690.5686)

(7899, 15691)

A random sample of 400 items is found to have mean 82 and S.D 18. Find the max error of estimation and 95% confidence interval

$n = 400$

$\bar{x} = 82$

$\sigma = 18$

$E = ?$

95% confidence limits = ?

$Z_{\alpha/2} = 1.96$

$E = Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) = 1.96\left(\frac{18}{\sqrt{400}}\right)$

$= 1.764$

∴ The 95% confidence interval

$(\bar{x} - E, \bar{x} + E)$

$(82 - 1.764, 82 + 1.764)$

$(80.236, 83.764)$

A random sample of size 100 has a S.D 5 What can you say about the maximum error with 95% confidence

$n = 100$

$\sigma = 5$

$Z_{\alpha/2} = 1.96$

$E = Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$

$= 1.96\left(\frac{5}{\sqrt{100}}\right)$

$E = 0.98$

what is the size of the smallest sample required to estimate an unknown proportion to within a maximum error of 0.06 with 95% confidence.

$E = 0.06$

$Z_{\alpha/2} = 1.96$

$n = ?$

$P = 1/2,\ q = 1/2$

$n = \left(\frac{Z_{\alpha/2}}{E}\right)^2 PQ$

$= \left(\frac{1.96}{0.06}\right)^2 \left(\frac{1}{4}\right)$

$\simeq 266.7 \simeq 2\overset{6}{9}7$

$\boxed{n = 267}$

If we can assert 95% that the maximum error is 0.05 and $P = 0.2$ find the size of sample

$E = 0.05$

$Z_{\alpha/2} = 1.96$

$P = 0.2, Q = 0.8$

$$n = \left(\frac{Z\alpha/2}{E}\right)^2 PQ$$

$$= \left(\frac{1.96}{0.05}\right)^2 \times 0.2 \times 0.8$$

$$= 245.8 \simeq 246$$

$$\boxed{n = 246}$$

Assuming that $\sigma = 20$ How large a random sample be taken to assert with probability 0.95 that the sample mean will not differ from the true mean by more than 3 points

$\sigma = 20$

$Z\alpha/2 = 1.96$

$E = 3$

$n = ?$

$$n = \left(\frac{Z\alpha/2\,\sigma}{E}\right)^2$$

$$= \left(\frac{1.96 \times 20}{3}\right)^2$$

$$= 170.73 \simeq 171$$

$$\boxed{n = 171}$$

Find 95% confidence limits for the mean of a normality distributed population from which the following sample was taken 15, 17, 10, 18, 16, 9, 7, 11, 13, 14 population values are 15, 17, 10, 18, 16, 9, 7, 11, 13, 14

$n = 10$, $(n < 30)$

σ is unknown

95% confidence limits = ?

$$\bar{x} = \frac{15+17+10+18+16+9+7+11+13+14}{10}$$

$$\bar{x} = \frac{130}{10}$$

$$\boxed{\bar{x} = 13}$$

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

$$= \frac{1}{9}\left[(15-13)^2 + (17-13)^2 + (10-13)^2 + (18-13)^2 + (16-13)^2 + (9-13)^2 + (7-13)^2 + (11-13)^2 + (13-13)^2 + (14-13)^2\right]$$

$$s^2 = \frac{120}{9} = 13.33$$

$$s = \sqrt{13.33} = 3.65$$

$$E = t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

$$E = 2.262\left(\frac{3.65}{\sqrt{10}}\right)$$

$$E = 2.610$$

∴ The 95% confidence limits are

$$(\bar{x} - E, \bar{x} + E)$$

$$(13 - 2.610, 13 + 2.610)$$

population values are 15, 17, 10, 18, 16, 9, 7, 11, 13, 14

$n=10$, $(n<30)$

σ is unknown

95% confidence limits = ?

$$\bar{x} = \frac{15+17+10+18+16+9+7+11+13+14}{10}$$

$$\bar{x} = \frac{130}{10}$$

$$\boxed{\bar{x}=13}$$

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

$$= \frac{1}{9}\Big[(15-13)^2+(17-13)^2+(10-13)^2+(18-13)^2+ (16-13)^2+(9-13)^2+(7-13)^2+(11-13)^2+ (13-13)^2+(14-13)^2\Big]$$

$$s^2 = \frac{120}{9} = 13.33$$

$$s = \sqrt{13.33} = 3.65$$

$$E = t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

$$E = 2.262\left(\frac{3.65}{\sqrt{10}}\right)$$

$$E = 2.610$$

$\therefore$ The 95% confidence limits are

$(\bar{x}-E, \bar{x}+E)$

$(13-2.610, 13+2.610)$

$(10.390, 15.610)$

A random sample of 100 teachers in a large metropolitan area revealed a mean weekly salary of 487 with a S.D Rs 48. with what degree of confidence can we assert the avg weekly salary of all teachers in metropolitan area is between 470 to 502.

Given $n = 100$

$\mu = 487$

$\sigma = 48$

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{x} - 487}{\frac{48}{\sqrt{100}}} = \frac{\bar{x} - 487}{4.8}$$

standard normal variate for $\bar{x}_1 = 472$

when $\bar{x}_1 = 472$

$$Z = \frac{472 - 487}{4.8} = -3.125$$

$\bar{x}_2 = 502$

$$Z_2 = \frac{502 - 487}{4.8} = 3.125$$

Here $Z_1 < 0$ & $Z_2 > 0$

$$P(x_1 \le x \le x_2) = P(Z_1 \le Z \le Z_2) = A(Z_2) + A(Z_1)$$

$= A(3.125) + A(-3.125)$

$= 2A(3.125)$

$= 2 \times 0.4991$

$= 0.9982$

$= 0.9982 \times 100\%$

$= 99.82\%$

In a study of an automobile insurance a random sample of 80 body repair cost had a mean of Rs 472.36 and S.D of Rs 62.35 If $\bar{x}$ is used as point estimate to the true average repair cost with what confidence we can assert that the maximum error doesnot exceed Rs 10.

$n = 80$

$\bar{x} = 472.36$

$\sigma = 62.35$

$E = 10\,Rs$

$$E = \frac{Z_{\alpha/2}\,\sigma}{\sqrt{n}}$$

$$10 = \frac{Z_{\alpha/2}(62.35)}{\sqrt{80}}$$

$$\frac{10 \times \sqrt{80}}{62.35} = Z_{\alpha/2}$$

$$Z_{\alpha/2} = 1.434$$

The area of $Z_{\alpha/2} = 1.434$ is 0.4236 from normal variate

$\alpha/2 = 0.4236$

$\alpha = 2(0.4236)$

$\alpha = 0.8472$

$\therefore$ Degree of confidence limit $= 0.8472 \times 100\%$

$= 84.72\%$

It is desired to estimate the mean no. of hours of continous use until a certain comp will first required repairs if it can be assumed $\sigma = 48$ hours. How large a sample be needed so that one will be able to assert with 90% confidence that the sample mean is of by atmost 10hrs.

$$\sigma = 48$$

$$Z = 1.64$$

$$E = 10$$

$$n = \left(\frac{Z\sigma}{E}\right)^2$$

$$= \left(\frac{1.64 \times 48}{10}\right)^2$$

$$n = 61.96 \simeq 62$$

The mean of arandom sample is an unbia estimate of mean of the population 3, 6, 9, 15, 27 List all possible samples of size 3 that can be taken without replacement from the finite population.

(i) prove that $\bar{x}$ is an unbiased estimated of θ i.e $E(\bar{x}) = \theta$

Population values are 3, 6, 9, 15, 27

$N = 5$

i) The no. of samples of size 3 is drawn from finite population.

$NCn = 5C3 = 10$

The list of 10 samples are

(3,6,9) (3,6,15) (3,6,27) (3,9,15) (3,9,27) (3,15,27)

(6,9,15) (6,9,27) (6,15,27)

(9,15,27)

Mean of population $\theta = \frac{3+6+9+15+27}{5}$

$= \frac{60}{5} = 12$

$\boxed{\theta = 12}$

The means of above 10 samples

6 8 12 9 13 15

10 14 16

17

The probability of each one is 1/10

$\bar{x}$	6	8	12	9	13	15	10	14	16	17
$P(\bar{x})$	1/10	1/10	1/10	1/10	1/10	1/10	1/10	1/10	1/10	1/10

$E(\bar{x}) = \Sigma \bar{x} P(\bar{x})$

$= 6 \times \frac{1}{10} + \frac{8}{10} + \frac{12}{10} + \frac{9}{10} + \frac{13}{10} + \frac{15}{10} + \frac{10}{10} + \frac{14}{10} + \frac{16}{10} + \frac{17}{10}$

$E(\bar{x}) = 12, \ \theta = 12$

$E(\bar{x}) = \theta$

10 bearings made by a certain process have a mean diameter of 0.5060 cm with a S.D of 0.0040cm. Assuming that the data may be taken as a random sample from a normal distribution. construct 95% confidence interval for the actual average diameter of bearings.

$n = 10 \ (< 30)$

$t_{\alpha/2}$ at 9 degrees of freedom $= 2.26$

$E = t_{\alpha/2} \dfrac{s}{\sqrt{n}}$

$\bar{x} = 0.5060$ cm

$s = 0.0040$ cm

$E = 2.262 \times \dfrac{0.0040}{\sqrt{10}}$

$\therefore E = 0.00286$

$\therefore$ 95% confidence limits are $(\bar{x} - E, \bar{x} + E)$

$(0.5060 - 0.00286, \ 0.5060 + 0.00$

$(0.5032, \ 0.5085)$

Testing of Hypothesis:

A statistical Hypothesis is a statement about the unknown parameters of one or more populations. Testing of hypothesis is a process for deciding whether to accept or Reject the hypothesis.

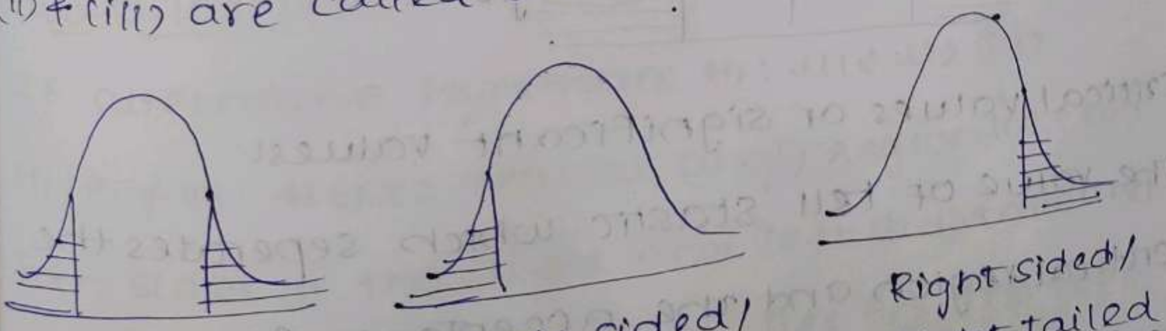
Types of Hypothesis:

1. Null Hypothesis: A NULL Hypothesis which asserts that there is no significant difference between the population parameter and statistic. It is denoted by H_0. It is in form of $H_0: \mu = \mu_0$

2. Alternative Hypothesis: The complementary of Null Hypothesis is called Alternative hypothesis. It is denoted by H_1. It is in form of

(i) $H_1: \mu \neq \mu_0$ (ii) $H_1: \mu > \mu_0$ (iii) $H_1: \mu < \mu_0$

In Alternative Hypothesis (i) is called two-tailed (ii) & (iii) are called one-tailed.



Two sided/ Two Tailed

left sided/ left tailed

Right sided/ Right tailed

Errors of sampling:

i) Type I - Error: Reject Ho when it is true

If the Null hypothesis is true but we reject the hypothesis then the error made is called Type-I error. It is denoted by 'α' (level of significance)

ii) Type II - Error: Accept Ho when it is wrong

If the Null Hypothesis is False but we accepted by test then the error made is called Type-II error. It is denoted by β error

Critical Region: A Region corresponding to statistic t in sample spaces which leads to the Rejection of Ho is called critical Region.

Those Region which leads to the acceptance of Acceptance Region. In general we take 5%, 1%, 10% area of two critical regions which cor

Critical values or significant values:

The value of test Stastic which seperates the critical region and the acceptance region is called critical value. This value is dependent

i) The level of significance used.

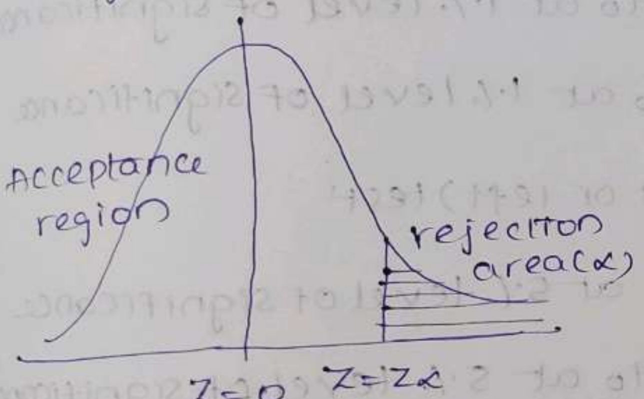
ii) The alternative hypothesis whether it is one-tailed / two tailed.

one Tailed test: For a test the critical region is represented at only one side (left or right) of the sampling distribution is called one tailed test

$H_1: \mu > \mu_0$ (Right tailed)

$H_1: \mu < \mu_0$ (left tailed)

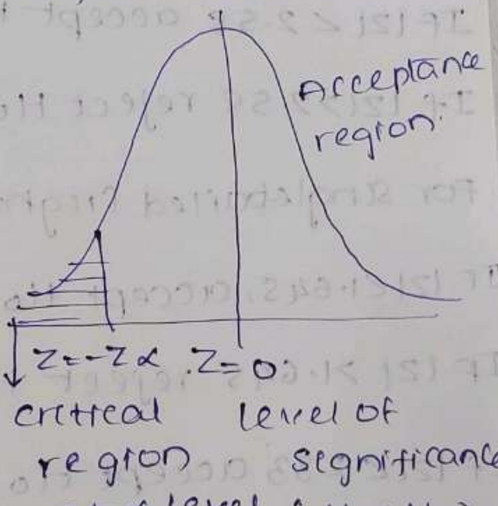
Right tailed test



(level of significance=α)

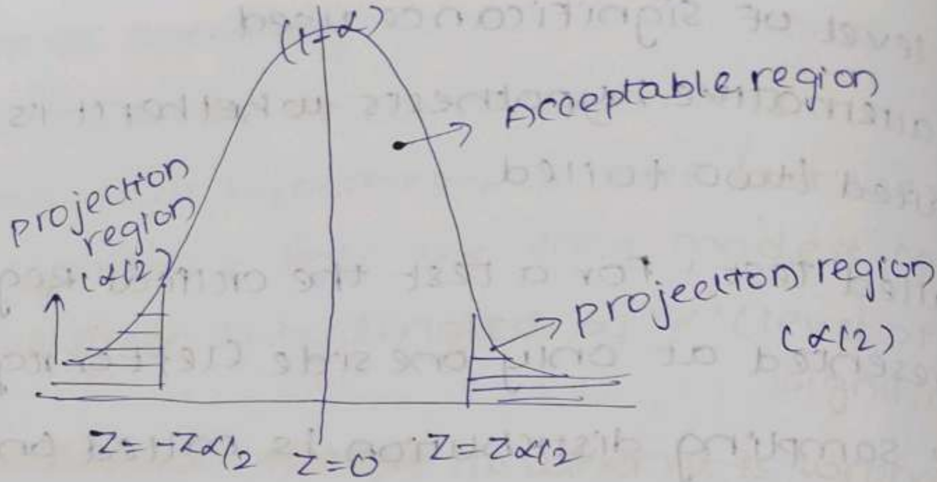
($\mu_1 > \mu_0$)

left tailed test



Two-tailed Test:

If alternative hypothesis $H_1: \mu_1 \neq \mu_2$ (or) $H_1: \sigma_1 \neq \sigma_2$ Hence critical (and) region lies on both sides of the right and left (ed) tails and the left critical region of area $\alpha/2$ lies on the left tail



For two tailed test:

If $|Z| < 1.96$ accept H_0 at 5% level of significance

If $|Z| > 1.96$ reject H_0 at 5% level of significance

If $|Z| < 2.58$ accept H_0 at 1% level of significance

If $|Z| > 2.58$ reject H_0 at 1% level of significance

For single tailed (right or left) test:

If $|Z| < 1.645$, accept H_0 at 5% level of significance

If $|Z| > 1.645$ reject H_0 at 5% level of significance

If $|Z| < 2.33$ accept H_0 at 1% level of significance

If $|Z| > 2.33$ reject H_0 at 1% level of significance.

critical values of Z

level of significance (α)	1%	5%	10%
two tailed test $\mu \neq \mu_0$	2.58	1.96	1.645
Right tailed test $\mu > \mu_0$	2.33	1.645	1.28
left tailed test $\mu < \mu_0$	-2.33	-1.645	-1.28

Test of significance for large samples:

To test the hypothesis that the probability of success in such a trail is P. Assuming it to be true. The mean $\mu = np$ and S.D $\sigma = \sqrt{npq}$ respectively.

If x be the observed no. of success in the sample and z is the standard normal variate is given by $z = \frac{x-\mu}{\sigma}$

A coin was tossed 960 times and return heads 183 times test the hypothesis (and return head) that the coin is unbiased use 0.05 level of significance.

$n = 960$

x = number of success = 183

p = probability of getting a head = $1/2$

$q = 1-p = 1-1/2 = 1/2$

$\mu = np = \frac{960}{2} = 480$

$\sigma = \sqrt{npq} = \sqrt{\frac{480}{2}} = \sqrt{240}$

$= 15.4919$

Null Hypothesis (H_0): The coin is unbiased

Alternative Hypothesis (H_1): The coin is unbiased

level of significance (α) = 0.05

Test statistic $|z| = \left| \frac{x-\mu}{\sigma} \right|$

$$z = \left| \frac{183-480}{15.4919} \right|$$

$$|z| = |-19.1713|$$

$$|z| = 19.1713$$

Conclusion: The calculated value $|z|$ is $= 19.1713$ is > the table value of z_α at 0.05 level of significance is 1.96

So, we reject H_0 i.e The coin is biased.

A dice is tossed 960 times It falls with 5 upwards 184 times is the die unbiased at level of significance of 0.01

$n = 960$

$x = 184$

p = probablity of getting a $5 = 1/6$

$$q = 1-p = 1-\frac{1}{6} = \frac{5}{6}$$

$$\mu = np = \frac{960}{6} = 160$$

$$\sigma = \sqrt{npq} = \sqrt{\frac{160 \times 5}{6}} = 11.547$$

Null hypothesis (H_0): The die is unbiased

Alternative hypothesis (H_1): The die is biased

level of significance (α) = 0.01

Test Stastistic : $|z| = \left| \frac{x-\mu}{\sigma} \right|$

$$= \left|\frac{184-160}{11.547}\right|$$

$$|Z| = 2.0784$$

conclusion: The calculated value $|Z| = 2.0784$ is the table value of Z_α at 0.01 level of significance is 2.58

so, we accept H_0 i.e The die is unbiased.

A coin was tossed 400 times and returned heads 216 times test the hypothesis that the coin is unbiased use 0.05 level of significance.

A die is tossed 256 times and it turns up with an even digit 150 times is the die biased

NOTE: when level of significance is not given by default we have to Take 0.05 level of significance

$n = 400$

$x = 216$

p = probability of getting heads $= 1/2$

$q = 1-p = 1-\frac{1}{2} = 1/2$

$\mu = np$

$= 400\left(\frac{1}{2}\right) = 200$

$\sigma = \sqrt{npq} = \sqrt{400 \times \frac{1}{4}} = 10$

Null Hypothesis: The coin is unbiased

Alternative Hypothesis: The coin is biased

level of significance $\alpha = 0.05$

Test statistic $|z| = \left|\frac{x-\mu}{\sigma}\right|$

$= \left|\frac{216-200}{10}\right|$

$|z| = 1.6$

Conclusion: The calculated value $|z| = 1.6$ is the table value of z_α at 1.96

so, we reject H_0 i.e the coin is baised

2) $n = 256$

$x = 150$

p = probability of getting even $= 1/2$

$q = 1-p \Rightarrow 1-\frac{1}{2} = \frac{1}{2}$

$\mu = np$

$= 256\left(\frac{1}{2}\right) = 128$

$\sigma = \sqrt{npq} = \sqrt{256 \times \frac{1}{4}} = 5.6568$

Null Hypothesis (H_0): The die is unbiased

Alternative Hypothesis (H_1): The die is biased

level of significance $\alpha = 0.01$

test statisitc : $|z| = \left|\frac{\bar{x}-\mu}{\sigma}\right|$

$= \left|\frac{150-128}{5.6568}\right|$

$|z| = 3.889$

conclusion: The calculated value $|z| = 3.889$. in table value of z_α is 1.96

so, we accept i.e the die is unbiased.

The four important test to the tests of significance are as follows.

1. Testing of significants for single mean.
2. Testing of significants for diff of mean.
3. Testing of significants for single proportion
4. Testing of significants for difference of proportion.

1. Testing of significants for single mean:

Let a random sample of size $n \geq 30$ has the sample mean $\bar{x}$. μ be the population mean also the population mean μ has the specified value μ_0.

working rule:

1. Null hypothesis (H_0): Their is no significant difference b/w the sample mean & population mean

2. Alternative Hypothesis (H_1):

i) $\mu \neq \mu_0$ (two tailed test)

ii) $\mu > \mu_0$ (Right tailed test)

iii) $\mu < \mu_0$ (left tailed test)

3. level of significance: set the level of significance α

4. Test statistic:

i) when the population S.D is known. In this case z is given by $|z| = \dfrac{|\bar{x} - \mu|}{\frac{\sigma}{\sqrt{n}}}$ where μ is sample mean

ii) when σ is unknown In this case $|z| = \left|\dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}\right|$

5. Decision:

If $|z| < z_\alpha$ we accept ~~we accept~~ H_0

If $|z| > z_\alpha$ we reject H_0

NOTE: we reject Null Hypothesis H_0 when $|z| > 3$ without mentioning any level of significance

- The confidence limits are also known as (fiducial) limits.

- The 95% or 5% fiducial limits are given by $\bar{x} \pm 1.96\left(\frac{\sigma}{\sqrt{n}}\right)$

Similarly, for 99% confidence limits are given by $\bar{x} \pm 2.58 \frac{\sigma}{\sqrt{n}}$

for 98% confidence limits are given by

$$\bar{x} \pm 2.33 \frac{\sigma}{\sqrt{n}}$$

A sample of 64 students have a mean weight of 70kg can this be regarded as a sample from the population with mean weight 56kg and S.D 25kg

$n = 64$

$\bar{x}$ = sample mean

= 70kg

μ = population mean = 56kg

σ = S.D = 25kg

Null Hypothesis H_0: The sample can be regarded is taken from population from with mean weight 56kg.

Alternative Hypothesis H_1: It can't be regarded that the sample is drawn from population with mean 56kgs

level of significance: $\alpha = 0.05$

$Z_{\alpha/2}$ at 5% = 1.96

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

$$= \left| \frac{70-56}{\frac{25}{\sqrt{64}}} \right| = 4.48$$

conclusion: The calculated value at 5% level of significance is greater than table value

$|Z| = 4.48 >$

$Z\alpha = 1.96$

$\therefore$ we reject H_0

Hence, the sample with mean weight 70kg is not drawn from this population.

An oceanographer wants to check whether the depth of the ocean in a certain region is 57.4 fathoms has had previously been recorded. what can he conclude at 0.05 level of significance if readings taken at 40 randomn locations In the given region yielding a mean of 59.1 fathoms with a S.D of 5.2 fathoms.

$n = 40$

$\bar{x} = 59.1$ fathoms

$\mu = 57.4$ fathoms

$\sigma = S.D = 5.2$ fathoms

Null Hypothesis (H_0): $\mu = 57.4$

Alternate Hypothesis (H_1): $\mu \neq 57.4$

level of significance: $\alpha = 5\% = 0.05$

The Z_α at 5% is 1.96

Test statistic:

$$|Z| = \left|\frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}\right| = \left|\frac{59.1 - 57.4}{\frac{5.2}{\sqrt{40}}}\right|$$

$$= 2.067$$

conclusion: The calculated value at 5% level of significance > table value

$|Z| > Z_\alpha$

$\therefore$ we reject H_0.

In a random sample of 900 members has a mean of 3.4 cm and S.D 2.61 cm is this sample has been taken from a large population of mean 3.25 cm and S.D 2.61 cm. If the population is normal and its mean is unknown find 95% fiducial limits of true mean.

$n = 900$

$\bar{x} = 3.4$ cms

$\mu = 3.25$ cms

$\sigma = S.D = 2.61$ cms

Null hypothesis (H_0) : $\mu = 3.25$ cms

Alternative hypothesis (H_1): $\mu \neq 3.25$ (Two tailed test)

level of significance: $\alpha = 95\%$ (or) $5\% = 0.05$

The Z_α at 5% is 1.96

Test statistic:

$$|z| = \left|\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| = \left|\frac{3.4-3.25}{\frac{2.61}{\sqrt{900}}}\right|$$

$$= 1.724$$

Conclusion: The calculated value at 5% level of significance is < table value

$\therefore |z| < z_\alpha$

$\therefore$ we accept H_0

The 95% fiducial limits are

$$\left(\bar{x}-1.96\left(\frac{\sigma}{\sqrt{n}}\right), \bar{x}+1.96\left(\frac{\sigma}{\sqrt{n}}\right)\right)$$

$$\left(3.4-1.96\left(\frac{2.61}{\sqrt{900}}\right), 3.4+1.96\left(\frac{2.61}{\sqrt{900}}\right)\right)$$

$$(3.2295, 3.57)$$

An ambulance service claims that it takes on the average less than 10min to reach its destination in emergency calls. A sample of 36 calls has a mean of 11min and the variance of 16min Test the claim at 0.05 level of significance.

$\mu = 10$

$\bar{x} = 11$

$n = 36$

$\sigma^2 = 16$

$\sigma = 4$

Null Hypothesis (H_0) = $\mu = 10$

Alternative hypothesis (H_1): $\mu < 10$

(left tailed test)

level of significance: $\alpha = 5\% = 0.05$

The $Z\alpha$ at 5% is 1.645

Test statistic

$$|Z| = \left| \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \right| = \left| \frac{11 - 10}{\frac{4}{\sqrt{36}}} \right| = 1.5$$

conclusion: The calculated value at 5% level of significance is < table value

$\therefore |Z| < Z\alpha$

we accept H_0

In a random sample of 60 workers. The average time taken by them to take the work is 33.8 men with S.D 6.1 men can be reject the null hypothesis $\mu = 32.6$ min in a favour of alternative null hypothesis $\mu > 32.6$. use $\alpha = 1\%$ level of significance

$n = 60$

$\bar{x} = 33.8$ min

$\mu = 32.6$ min

$\sigma = 6.1$ min

$H_0: \mu = 32.6$ min

$H_1: \mu > 32.6$ (Right tailed test)

L.O.S: $\alpha = 0.01$ The table value of

$Z_\alpha = 2.33$

Test statistic: $|Z| = \left| \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \right| = \left| \frac{33.8 - 32.6}{\frac{6.1}{\sqrt{60}}} \right|$

$= 1.523$

Conclusion: The calculated value $|Z| < Z_\alpha$ hence we accept H_0.

It is claimed that a random sample of 49 has a mean life of 15200 km. This sample was drawn from population whose mean is 15150 km has a SD of 1200 km. Test the significance at 0.05 level.

$n = 49$

$\mu = 15150$ km

$\sigma = 1200$ km

$H_0: \mu = 15150$ km

$H_1: \mu \neq 15150$ km (Two tailed test)

$\alpha = 0.05 \quad Z_\alpha = 1.96$

Test statistic: $|z| = \left| \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \right|$

$= \left| \dfrac{15200 - 15150}{\frac{1200}{\sqrt{49}}} \right|$

$= 0.2916$

conclusion: the calculated value of $|z| < z_\alpha$

here, we accept H_0

Test for equality of two means:-

Let $\bar{x}_1$ and $\bar{x}_2$ be the sample means of two independent large random samples of sizes n_1 and n_2 draw from two populations having means μ_1 and μ_2 and S.D σ_1 and σ_2

$H_0: \mu_1 = \mu_2$

H_1: i) $\mu_1 \neq \mu_2$ (Two tailed)

ii) $\mu_1 > \mu_2$ (Right tailed)

iii) $\mu_1 < \mu_2$ (left tailed)

Test statistic: when σ is known

Case i: $|z| = \left| \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \right|$

case ii) if $\sigma_1^2 = \sigma_2^2 = \sigma^2$:

$|z| = \left| \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right|$

case iii when σ is unknown

$$|Z| = \left| \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \right|$$

where s_1, s_2 are sample S.D's

The means of two large samples of sizes 1000 and 2000 members are 67.5 inches and 68 inches respectively. can the samples be regarded as drawn from the same population of S.D 2.5 inch

$n_1 = 1000$ $n_2 = 2000$

$\bar{x}_1 = 67.5$ inch $\bar{x}_2 = 68$ inches

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

$\sigma_1 = \sigma_2 = \sigma$ i.e The samples are drawn from sample population with S.D = 2.5 inch

H_0: $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$ (Two tailed test).

l.o.s: $\alpha = 0.05$ The table value of $Z_\alpha = 1.96$

Test statistic:

$$|Z| = \left| \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right| = \left| \frac{67.5 - 68}{\sqrt{(2.5)^2\left(\frac{1}{1000} + \frac{1}{2000}\right)}} \right|$$

$$= |-5.1639| = 5.1639$$

conclusion: The calculated value $|Z| > Z_\alpha$

Hence we reject H_0

A researcher wants to know the intelligence of students in school. He selected two groups of students. In the 1st group there are 150 students having mean IQ of 75, with S.D of 15 in the 2nd group there are 250 students having mean IQ of 70 with S.D of 20.

$n_1 = 150$ $n_2 = 250$

$\bar{x}_1 = 75$ $x_2 = 70$

$\sigma_1 = 15$ $\sigma_2 = 20$

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$ (Two tailed test)

L.O.S: $\alpha = 0.05$ The table value of $Z_\alpha = 1.96$

Test statistic: $\left| \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \right| = \left| \dfrac{75 - 70}{\sqrt{\dfrac{(15)^2}{150} + \dfrac{(20)^2}{250}}} \right|$

$= 2.839$

Conclusion: The calculated value $|Z| > Z_\alpha$

Hence, we reject H_0

Two types of new cars produced in USA are tested for petrol millage one sample is consisting of 42 cars having 15 kmph. while the other sample consisting of 80 cars having average 11.5 KMPN with population variances $\sigma_1^2 = 2$, $\sigma_2^2 = 1.5$ respectively. Test whether there is any significance difference in the petrol consumption of these types of cars using 1% L.O.S

A simple sample of height of 6400 englishmen has a mean of 67.85 inches and S.D of 2.56 inches while a sample of heights of 1600 australians has a mean of 68.55 inches and S.D of 2.52 inches do the data indicate the australians are on the average taller than englishmen use α as 0.01

$n_1 = 6400$ $\qquad$ $n_2 = 1600$

$\bar{x}_1 = 67.85$ inches $\qquad$ $\bar{x}_2 = 68.55$ inches

$\sigma_1 = 2.56$ $\qquad$ $\sigma_2 = 2.52$

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 < \mu_2$ (left tailed test)

$\alpha = 0.01$

The table value of $Z_\alpha = 2.33$

Test statistic $|Z| = \left| \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \right|$

$= \left| \dfrac{67.85 - 68.55}{\sqrt{\dfrac{(2.56)^2}{6400} + \dfrac{(2.52)^2}{1600}}} \right|$

$= 9.906$

The calculated value
so we reject H_0

we conclude that Australians are taller than englishmen.

The mean life of sample of 10 electric bulbs was found to be 1456 hrs with S.D of 423 hrs. a second sample of 17 bulbs choosen from a different batch showed a mean life of 1280 hrs with S.D of 398 hrs is their a significant difference between the means of 2 batches.

$n_1 = 10 \qquad n_2 = 17$

$\bar{x}_1 = 1456 \qquad \bar{x}_2 = 1280$

$\sigma_1 = 423 \qquad \sigma_2 = 398$

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$ (Two tailed test)

$\alpha = 0.05$ The table value of $Z_\alpha = 1.96$

Test statistic: $|Z| = \left| \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \right|$

$= \left| \dfrac{1456 - 1280}{\sqrt{\dfrac{(423)^2}{10} + \dfrac{(398)^2}{17}}} \right|$

$= 1.066$

Conclusion: The calculated $|Z| < Z_\alpha$

So, we accept H_0.

we can conclude that there is no difference between the life of electric bulbs of 2 batches

Test of significance for single proportion.

Suppose a large random sample of size n has a sample proportion p of members posessing a certain attribute (proportion of success).

→ To test the hypothesis that the proportion P in the population has a specified value P_0

$H_0: P = P_0$

$H_1: P \neq P_0$, $P > P_0$, $P < P_0$

(Two tailed) (Right tailed) (left tailed test)

Test statistic: $|z| = \left| \dfrac{p - P}{\sqrt{\frac{PQ}{n}}} \right|$

where p = sample proportion

P = population proportion

confidence interval for proportion P for large samples at α level of significance is

$$\left(p - z_{\alpha/2}\sqrt{\frac{PQ}{n}} ,\ p + z_{\alpha/2}\sqrt{\frac{PQ}{n}} \right)$$

At 99% → $z_{\alpha/2} = 3$

95% → $z_{\alpha/2} = 1.96$

90% → $z_{\alpha/2} = 1.645$

A manufacturer claimed that atleast 95% of the equipment which he supply to a factory confirmed to specifications. An examination of a sample of 200 pieces of equipment revealed that 18 were faulty. Test his claim at 5% l.o.s

$n = 200$

No. of faculty equipment $= 18$

$\therefore$ The no. of equipment upto the specifications are $x = 200 - 18 = 182$

$x = 182$

p = sample proportion $= \frac{x}{n} = \frac{182}{200} = 0.91$

P = population proportion $= 95\% = 0.95$

$H_0 : P = 0.95$

$H_1 : P < 0.95$ (left tailed test)

$\alpha = 0.05$

The table value of $Z\alpha = 1.645$

Test statistic : $|Z| = \left| \frac{p - P}{\sqrt{\frac{PQ}{n}}} \right|$ where $P = 0.95$

$Q = 1 - P$

$= 1 - 0.95$

$= 0.05$

$$= \left| \frac{0.91 - 0.95}{\sqrt{\frac{0.95 \times 0.05}{200}}} \right| = 2.59$$

Conclusion: The calculated $|Z| > Z\alpha$

So we reject H_0.

In a sample of 1000 people in Karnataka 540 are rice eaters and the rest are wheat eaters. Can it be assumed that both rice and wheat are equally popular in this state at 1% level of significance

$n = 1000$

x = Number of rice eaters $= 540$

p = sample population $= \frac{x}{n} = \frac{540}{1000} = 0.54$

P = population proportion $= 1/2$

$Q = 1 - P = 1 - 1/2 = 1/2$

$H_0: P = 1/2$

$H_1: P \neq 1/2$ (Two tailed test)

L.O.S $\alpha = 0.01$

The table value of $Z_\alpha = 2.58$

Test statistic: $|Z| = \left| \frac{p - P}{\sqrt{\frac{PQ}{n}}} \right|$

$= \left| \frac{0.54 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{1000}}} \right|$

$= 2.529$

Conclusion: The calculated value $|Z| = 2.529$

table value $Z_\alpha = 2.58$

Hence, we accept H_0

i.e both rice and wheat eaters

popular in karnataka

In a big city 325 men out of 600 men were found to be smokers does these information support the conclusion that the majority of men in this city are smokers.

$P = 1/2$

$x = 325$

$n = 600$

p = sample proportion $= \frac{325}{600} = 0.54$

P = population proportion $= 1/2$

$Q = 1 - P = 1 - 1/2 = 1/2$

$H_0 : P = 1/2$

$H_1 : P > 1/2$ (Right tailed test)

l.o.s: $\alpha = 0.05$ The table value of

$Z_\alpha = 1.645$

Test statistic : $|Z| = \left| \frac{p - P}{\sqrt{\frac{PQ}{n}}} \right| = \left| \frac{0.5416 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{600}}} \right|$

$= 2.03$

Conclusion: The calculated value > table value

Hence we reject Ho

we conclude that majority of men in the city are smokers.

In a random sample of 160 workers exposed to a certain amount of radiations 24 experienced some ill effects. construct a 99% confidence interval for the corresponding true proportion.

$x = 24$

$n = 160$

$p = \frac{x}{n} = \frac{24}{160} = 0.15$

$q = 1 - p$

$= 1 - 0.15$

$= 0.85$

The 99% confidence interval is

$$\left(P - 3\sqrt{\frac{PQ}{n}},\ P + 3\sqrt{\frac{PQ}{n}}\right)$$

$$\left(0.15 - 3\sqrt{\frac{0.15 \times 0.85}{160}},\ 0.15 + 3\sqrt{\frac{0.15 \times 0.85}{160}}\right)$$

$= (0.065, 0.234)$

A Random sample of 500 apples was taken from a large consinment of 60 were found to be bad construct 95% confidence limits for the % number of bad apples in the consinment

20 people were attacked by a disease and only 18 survive will you reject the hypothesis that the survival rate if attacked by this disease is 85% in favour of the hypothesis i.e more at 5% level.

$n = 20$

$x = 18$

$p = \frac{x}{n} = \frac{18}{20} = 0.9 \qquad P = 0.85$

$Q = 1 - P$

$= 1 - 0.85$

$Q = 0.15$

$H_0: P = 0.85$

$H_1: P \neq 0.85$

l.o.s $\alpha = 0.05$

$= 1.96$

Test statistics $|z| = \left| \frac{p - P}{\sqrt{\frac{PQ}{n}}} \right|$

$= \left| \frac{0.9 - 0.85}{\sqrt{\frac{0.85 \times 0.15}{20}}} \right|$

$= 0.626$

Conclusion: The calculated value $|z| = 0.625$

table value $Z_\alpha = 1.96$

Hence we accept H_0

Test for equality of two proportions:

Let $p_1 p_2$ be the sample proportions in two large random samples of sizes n_1 and n_2 drawn from two populations having proportions P_1 and P_2.

To test whether the two samples have been drawn from same population

$H_0: P_1 = P_2$

$H_1: P_1 \neq P_2$

ii) $P_1 > P_2$

iii) $P_1 < P_2$

Test statistic:

case(i) when population proportions P_1, P_2 are known

$$|z| = \left| \frac{p_1 - p_2}{\sqrt{\frac{P_1 Q_1}{n_1}} + \sqrt{\frac{P_2 Q_2}{n_2}}} \right|$$

case(ii) Method of cooling

$$|z| = \left| \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right|$$

where $p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$

$q = 1 - p$

A ra were Flyo 325 Test and sar

A random samples of 400men and 600women were asked whether they would like to have flyover near their residence. 200men and 325women were in favour of the proposal. Test the hypothesis that proportions of men and women in favour of the proposal are same at 5·/. level.

$n_1=400$ $\quad$ $n_2=600$

$x_1=200$ $\quad$ $x_2=325$

$p_1=\frac{x_1}{n_1}=\frac{200}{400}$ $\quad$ $p_2=\frac{325}{600}$

$=0.54$

$p_1=1/2$

$H_0: P_1=P_2$ There is no significant difference in the option of men & women as for as proposal of flyoveres concerned

$H_1: p_1 \neq P_2$ (Two tailed test)

$1.0.S: \alpha=0.05$

The table value of $Z_\alpha=1.96$

Test statistic: $|z|=\left|\frac{p_1-p_2}{\sqrt{pq\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}\right|$

$$p=\frac{n_1p_1+n_2p_2}{n_1+n_2}$$

$$=\frac{200+325}{400+600}$$

$$=\frac{525}{1000}=0.525$$

$p=0.525$

$q=1-p=0.475$

$$|z| = \left| \frac{0.5 - 0.54}{\sqrt{0.525 \times 0.475 \left(\frac{1}{400} + \frac{1}{600}\right)}} \right|$$

$$|z| = 1.25$$

Conclusion: The calculated value $|z| < z_\alpha$

Hence we accept H_0

A cigerrete manufacturing claims that its brand A line of ciggeretes outsells its brand B by 8%. If it is found that 42 out of sample of 200 smokers preferred brand A and 18 out of another sample of 100 smokers prefered brand B. Test whether the 8% difference is valid claim.

$n_1 = 200$ $\quad$ $n_2 = 100$

$x_1 = 42$ $\quad$ $x_2 = 18$

$p_1 = \frac{x_1}{n_1} = \frac{42}{200} = 0.21$ $\quad$ $p_2 = \frac{x_2}{n_2} = \frac{18}{100} = 0.18$

$H_0: P_1 - P_2 = 8\%$

Assume that 8% of difference in the sales of two brands of cigerretes is valid claim.

L.O.S: $\alpha = 0.05$

The table value of $z_\alpha = 1.96$

Test statistic: $|z| = \left| \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right|$

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$p = \frac{42 + 18}{300} = \frac{60}{300} = \frac{1}{5} = 0.2$$

$$P = 0.2$$

$$q = 0.8$$

$$|z| = \left| \frac{0.21 - 0.18}{\sqrt{0.2 \times 0.8 \left(\frac{1}{200} + \frac{1}{100}\right)}} \right|$$

$$|z| = 0.61$$

conclusion! The calculated value $|z| < z_\alpha$

Hence we accept H_0.

we conclude that the difference of 8% of sales of two brands of cigarettes is valid claim.

In two large populations there are 30% and 25% respecitvely of fair haired people is this difference likely to be hidden in samples of 1200 and 900 respectively from the two populations.

$n_1 = 1200$ $n_2 = 900$

P_1 = proportion of faired haired people in 1st population $= 30\% = 0.3$, $Q_1 = 0.7$

P_2 = proportion of faired haired people in 2nd population $= 25\% = 0.25$, $Q_2 = 0.75$

Null Hypothesis! H_0

The difference in proportion of populations is likely to be hidden in the sample

$H_0: P_1 = P_2$

$H_1: P_1 \neq P_2$ (Two tailed test)

l.o.s: 0.05

The table value of $Z_\alpha = 1.96$

$$|Z| = \left| \frac{p_1 - p_2}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}} \right| = \left| \frac{0.3 - 0.25}{\sqrt{\frac{0.3 \times 0.7}{1200} + \frac{0.25 \times 0.75}{900}}} \right|$$

$$|Z| = 2.55$$

Conclusion: The calculated value

$$|Z| > Z_\alpha$$

Here we reject is H_0

$\therefore$ The sample proportions are not equal.

In a city A 20% of random sample of 900 school boys has a certain slight physical defect. In another city B 18.5% of random sample of 1600 school boys has the same defect. Is the difference between the proportions significant at 0.05 level of significance

$n_1 = 900$ $\qquad$ $n_2 = 1600$

$x_1 = 20\%$ of 900 $\qquad$ $x_2 = 18.5\%$ of 1600

$= 0.2 \times 900$ $\qquad$ $= 296$

$x_1 = 180$

$P_1 = \frac{x_1}{n_1} = \frac{180}{900}$ $\qquad$ $P_2 = \frac{x_2}{n_2} = \frac{296}{1600}$

$= 0.2$ $\qquad$ $= 0.185$

$H_0: P_1 = P_2$

$H_1: P_1 \neq P_2$ (two tailed test)

l.o.s: $\alpha = 0.05$

The table value of $Z_\alpha = 1.96$

Test statistic: $P = \dfrac{x_1 + x_2}{n_1 + n_2}$

$= \dfrac{180 + 296}{1600 + 900} = \dfrac{476}{2500}$

$= 0.19$

$q = 0.81$

$$|z| = \left| \frac{0.2 - 0.185}{\sqrt{0.19 \times 0.81 \left(\frac{1}{900} + \frac{1}{1600} \right)}} \right|$$

$= |-0.918| = 0.918$

Conclusion: The calculated value

$|z| < Z_\alpha$

Hence we accept H_0 i.e there is no significant difference between proportions

In a random sample of 1000 persons from town A 400 are found to be wheat consumers In a sample of 800 from town B 400 are found to be wheat consumers. Do these data reveal a significant difference b/w town A and town B so far as the proportion of wheat consumers is concerned.

$n_1 = 1000$

$x_1 = 400$

$P_1 = \dfrac{x_1}{n_1} = \dfrac{400}{1000} = 0.4$

$n_2 = 800$

$x_2 = 400$

$P_2 = \dfrac{x_2}{n_2} = \dfrac{400}{800}$

$= 0.5$

$H_0: P_1 = P_2$

$H_1: P_1 \neq P_2$ (Two tailed test)

l.o.s 0.05

The table value of $Z_\alpha = 1.96$

$$|z| = \left| \frac{P_1 - P_2}{\sqrt{PQ\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \right|$$

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$$

$$= \frac{1000 \times 0.4 + 800 \times 0.5}{1000 + 800}$$

$$= \frac{400 + 400}{1800}$$

$$= \frac{800}{1800} = 0.44$$

$$Q = 1 - P = 1 - 0.44 = 0.56$$

$$|z| = \left| \frac{0.4 - 0.5}{\sqrt{0.44 \times 0.56\left(\frac{1}{1000}+\frac{1}{800}\right)}} \right|$$

$$= |-4.2471|$$

$$z = 4.247$$

Conclusion: The calculated value

$$z > z_\alpha$$

Hence we reject H_0

students-t-test for single Mean:

let a random sample of size ($n<30$) has a sample mean $\bar{x}$ to test the hypothesis that the population mean μ has specified value (μ_0) when σ is unknown.

$H_0: \mu=\mu_0$

$H_1: \mu\neq\mu_0$

Test statistic: $|t| = \left|\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n-1}}}\right|$ where s is the sample standard deviation follows t-distribution with $v=n-1$ degrees of freedom

NOTE: for a two tail test α is taken as $\alpha/2$

The confidence limits for the population mean is given by

$$\left(\bar{x}-t_\alpha\frac{s}{\sqrt{n-1}}, \bar{x}+t_\alpha\frac{s}{\sqrt{n-1}}\right)$$

The average breaking strength of steel rod is specified to be ~~18.5000~~ 1000 pounds to test this sample of 14 rods were tested. The mean and standard deviation were 17.85 and 1.955. Is the result of experiment significance

$n=14(<30)$

$\mu=18.5$

$\bar{x}$ = sample mean = 17.85

s = sample standard deviation = 1.955

$H_0: \mu = 18.5$ ie; The result of experiment is not significant

$H_1: \mu \neq 18.5$ (Two tailed test)

level of significance: $\alpha = 5\%$ $\Rightarrow \alpha/2 = \frac{0.05}{2}$

$\alpha = 0.05$ $= 0.025$

Test statistic: $|t| = \left| \dfrac{\bar{x} - \mu}{\frac{S}{\sqrt{n-1}}} \right|$

$= \left| \dfrac{17.85 - 18.5}{\frac{1.955}{\sqrt{13}}} \right| = 1.198$

Calculated value $|t| = 1.198$

The t distribution follows d.f is $\nu = n-1 = 14-1 = 13$

$\therefore t(0.025, 13) = 2.160$

Hence the calculated value < table value

we accept H_0

we conclude that the result of experiment is not significant.

A sample of 26 bulbs gives a mean life of 990 hours with a S.D of 20 hours. The manufacturer claims that the mean life of bulbs is 1000 hr. Is the sample not up to the standard.

$n = 26 (< 30)$

$\cancel{\mu = 990}$

$\mu = 1000$

$\bar{x}$ = sample mean = 990 hrs

S = sample standard deviation = 20hrs

$H_0: \mu = 1000$ i.e The sample is upto standard

$H_1: \mu < 1000$ (left failed Test)

level of significance: $\alpha = 5\% \Rightarrow \alpha = 0.05$

Test statistic: $|t| = \left| \dfrac{\bar{x} - \mu}{\frac{S}{\sqrt{n-1}}} \right|$

$= \left| \dfrac{990 - 1000}{\frac{20}{\sqrt{25}}} \right| = 2.5$

$\therefore$ calculated value $|t| = 2.5$

The t distribution follows d.f is $\nu = n-1$
$= 26 - 1 = 25$

$\therefore t(0.05, 25) = 1.708$

Hence the calculated value > table value

we reject H_0

The sample is not upto the standard.

A Machinist is making an engine parts with axle diameter of 0.700 inches. A random sample of 10 parts shows a mean diameter of 0.742 inches with S.D of 0.040 inches. Compute the statistic you would use to test whether the work is meeting the specification at 0.05 L.O.S.

When sample S.D is not given directly:

A random sample of 10 boys had the following IQ's 70 120 110 101 88 83 95 98 107 100

Do these data support the assumption of population mean IQ of 100?

(ii) Find a resonable range in which most of the mean IQ values of samples of 10 boys lie.

Given IQs are

70, 120, 110, 101, 88, 83, 95, 98, 107, 100

$n = 10 (n < 30)$

we have to find $\bar{x}$ & s

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{70+120+110+101+88+83+95+98+107+100}{10}$$

$\bar{x} = 97.2$

construct sample the S.D as follows.

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
70	-27.2	739.84
120	22.8	519.84
110	12.8	163.84
101	3.8	14.44
88	-9.2	84.64
83	-14.2	201.64
95	-2.2	4.84
98	0.8	0.64

107	9.8	96.04
100	2.8	7.84
		1833.60

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

$$= \frac{1}{9} [1833.60]$$

$$s = \sqrt{\frac{1833.60}{9}} = 14.27$$

H_0: $\mu = 100$

H_1: $\mu \neq 100$ (two tailed test)

L.O.S: $\alpha = 0.05 \Rightarrow \alpha/2 = 0.025$

The t follows $\nu = n-1$ d.f

$$\nu = 10-1$$
$$= 9 \text{ d.f}$$

Test statistic: $|t| = \left| \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} \right|$

$$= \left| \frac{97.2 - 100}{\frac{14.27}{3}} \right| = 0.5886$$

Conclusion

The $t_{(0.025, 9)} = 2.26$

Here calculated value < table value we accept H_0

we support the assumption of population mean IQ of 100

b) The 95% confidence limits of population mean

$$\left(\bar{x}-t_\alpha\frac{s}{\sqrt{n-1}},\ \bar{x}+t_\alpha\frac{s}{\sqrt{n-1}}\right)$$

$$\left(97.2-\frac{(2.26)(14.27)}{3},\ \frac{97.2+(2.26)(14.27)}{3}\right)$$

$$(86.44,\ 107.95)$$

A random sample of 10 bags of pesticides are taken whose weights are 50, 49, 52, 44, 45, 48, 46, 45, 49, 45 (kg). Test whether the average packing can be taken to 50kg. $\mu=50$

Students t test for difference of Means:

Let $\bar{x}$ and $\bar{y}$ be the means of two independent samples of sizes n_1 and n_2 (<30) drawn from two normal populaiton having means μ_1 and μ_2 - To test whether the two population means are equal.

$H_0: \mu_1=\mu_2$

$H_1: \mu_1\neq\mu_2,\ \mu_1>\mu_2,\ \mu_1<\mu_2$

Test statistic:

Case I: When s_1, s_2 are known

$$|t|=\left|\frac{\bar{x}-\bar{y}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right|$$

where $$S^2=\frac{n_1 s_1^2+n_2 s_2^2}{n_1+n_2-2}$$

t-dist follows n_1+n_2-2 d.f

case ii when S_1, S_2 are unknown

$$\bar{x} = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i$$

$$\bar{y} = \frac{1}{n_2}\sum_{j=1}^{n_2} y_j$$

$$S^2 = \frac{1}{n_1+n_2-2}\left[\sum(x_i-\bar{x})^2 + \sum(y_j-\bar{y})^2\right]$$

The confidence limits for difference of population means are given by $(\bar{x}-\bar{y}) \pm t_\alpha S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$.

Samples of two types of electric bulbs were tested for length of life and following data were obtain

type-I Type-II

x.

$n_2 = 7$

$\bar{y} = 1036$

$S_2 = 40$hrs

H_0: $\mu_1 \neq \mu_2$

H_1: $\mu_1 > \mu_2$ (Right tailed test)

L.O.S: $\alpha = 0.05$

Test statistic: $|t| = \left|\frac{\bar{x}-\bar{y}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right|$

$$S^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1+n_2-2}$$

$$= \frac{8\times(36)^2 + 7(40)^2}{8+7-2}$$

$$s^2 = 1659.08$$

$$s = \sqrt{1659.08}$$

$$= 40.31$$

$$|t| = \left| \frac{1234 - 1036}{40.731\sqrt{\frac{1}{8}+\frac{1}{7}}} \right|$$

$$= 9.39$$

Conclusion: The table value of $t(0.05, 13) = 1.771$

The calculated value is greater than table value

Hence we reject H_0 i.e we conclude that the mean of line of electric bulbs of type I is superior to type II.

To compare two kinds of bumper gaurds 6 of each kind were mounted on a car and then the car was run into a concrete wall. The following are the costs of repair:

gaurd 1	107	148	123	165	102	119
gaurd 2	134	115	112	151	133	129

use the 0.01 level of significance to test whether the difference b/w sample means is significant