mate thus constructed is such that the probability of the parameter lying in the interval can be determined. Accuracy of the estimate is indicated by the length of the interval. Thus **interval estimates** are intervals for which one can be $(1-\alpha)$ 100% confident that the parameter under investigation lies in this interval. Such an interval is known as **confidence interval** for the parameter with (having) $1-\alpha$ or $(1-\alpha)$ 100% degree of confidence. The two end points of the confidence interval are known as **confidence limits** or **fiducial limits** or **critical values** or confidence coefficients. **Confidence level** denoted by α is the percentage of confidence.

Consider a large random sample of size $n(\geq 30)$ from a population with *unknown* mean μ and *known* variance σ^2. Then the **large-sample confidence interval for** μ:

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

Note 1: Confidence interval is exact for random samples from normal populations.

Note 2: Confidence interval provides good approximation for large samples $(n \geq 30)$ from non-normal populations also.

Small-sample confidence interval for μ: (when $n < 30$, assuming sampling from normal population)

$$\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2}\frac{s}{n}.$$

Confidence level α	99.73%	99%	98%	96%	95.5%	95%	90%	80%	68.27%	50%
$Z_{\alpha/2}$	3.00	2.58	2.33	2.05	2.00	1.96	1.645	1.28	1.00	0.675

Worked Out Examples

Example: A random sample of 10 ball bearings produced by a company have a mean diameter of 0.5060 cm with s.d 0.004 cm. Find the maximum error estimate E and 95% confidence interval for the actual mean dia of ball bearings produced by this company assuming sampling from normal population.

Solution: Sample size $= n = 10 < 30$, so use t-distribution (small sampling).

Maximum error estimate at 95% confidence is

$$E = t_{\alpha/2}\frac{\sigma}{\sqrt{n}} = (2.262)\frac{(0.004)}{\sqrt{10}} = 0.00286$$

since $t_{0.025}$ with $n-1 = 10-1 = 9$ dof is 2.262. 95% confidence interval limits are $\bar{x} \pm t_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

$= 0.5060 \pm (2.262)\frac{(0.004)}{\sqrt{10}}$.

95% confidence interval is (0.5031, 0.5089).

Exercise

1. If on the average, the test strips painted across heavily travelled roads in 15 different locations, disappeared after they had been crossed by 146692 cars with s.d. 14380 cars, calculate 99% confidence intervals for the true average number of cars it takes to wear off the paint, assuming normal population.

   Hint: $n = 15 < 30$, $t_{0.005}$ with 14 dof is 2.977, $\bar{x} = 146692$, $\sigma = 14380$,

   C.I. $(146692 \pm \frac{(2.977)(14380)}{\sqrt{15}})$.

Ans: $135639 < \mu < 157745$

2. A random sample of 20 fuses subjected to overload has mean time for blow of 10.63 minutes with s.d. of 2.48 mt. What can we assert with 95% confidence about the maximum error if we use $\bar{x} = 10.63$ mts as a point estimate of true average it takes such fuses for blow when subjected to overload.

Ans: $E = t_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 2.093\frac{(2.48)}{\sqrt{20}} = 1.16$ mt

3. Construct a 99% confidence interval for the true mean weight loss if 16 persons on diet control after one month had a mean weight loss of 3.42 kgs with s.d. of 0.68 kgs.

   Hint: $n = 16$, $\bar{x} = 3.42$, $s = 0.68$, $t_{0.005} = 2.947$ for 15 dof

   $3.42 \pm 2.947\frac{(0.68)}{\sqrt{16}}$.

Ans: $2.92 < \mu < 3.92$

29.3 BAYESIAN ESTIMATION

Personal or subjective probability is the new concept introduced in Bayesian methods. Also, parameters are viewed as random variables in Bayesian methods.

To estimate the mean of a population, μ is treated as a random variable whose distribution is indicative of the "strong feelings" or assumption of a person about the possible value of μ. Let μ_0 and σ_0 be the mean and standard deviation of such a subjective "prior distribution".

Bayesian Estimation

Combining the prior feelings about the possible values of μ with direct sample evidence, the "posterior" distribution of μ in Bayesian estimation is approximated by normal distribution with

$$\mu_1 = \frac{n\bar{x}\sigma_0^2 + \mu_0\sigma^2}{n\sigma_0^2 + \sigma^2} \quad \text{and}$$

$$\sigma_1 = \sqrt{\frac{\sigma^2\sigma_0^2}{n\sigma_0^2 + \sigma^2}}$$

Here μ_1 and σ_1 are known as the mean and standard deviation of the posterior distribution. In the computation of μ_1 and σ_1, σ^2 is assumed to be *known*. When σ^2 is unknown, which is generally the case, σ^2 is replaced by sample variance s^2 provided $n \geq 30$ (large sample).

Bayesian interval for μ:

A $(1-\alpha)100\%$ Bayesian interval for μ is given by

$$\mu_1 - Z_{\alpha/2}\sigma_1 < \mu < \mu_1 + Z_{\alpha/2}\sigma_1.$$

Worked Out Examples

Example: A professor's feelings about the mean mark in the final examination in "probability" of a large group of students is expressed subjectively by normal distribution with $\mu_0 = 67.2$ and $\sigma_0 = 1.5$. (a) If the mean mark lies in the interval (65.0, 70.0), determine the prior probability the professor should assign to the mean mark. (b) Find the posterior mean μ_1 and posterior s.d. σ_1 if the examination is conducted on a random sample of 40 students yielding mean 74.9 and s.d. 7.4. Use $s = 7.4$ as an estimate of σ. (c) Determine the posterior probability which he will thus assign to the mean mark being in the interval (65.0, 70.0), using results obtained in (b)· (d) construct a 95% Bayesian interval for μ.

Solution:

a. Here $\mu_0 = 67.2$, $\sigma_0 = 1.5$, $n = 40$

standard variable corresponding to 65.0 is

$$Z_1 = \frac{65.0 - 67.2}{1.5} = -1.466,$$

Similarly, $$Z_2 = \frac{70 - 67.2}{1.5} = 1.866$$

Let X be the mean mark obtained in the final examination.

$$\text{Prior probability} = P(65 < X < 70)$$
$$= P(-1.47 < Z < 1.87)$$
$$= 0.4292 + 0.4693 = 0.8985$$

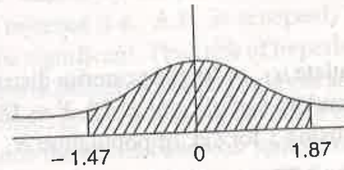


Fig. 29.1

b. Here $\bar{x} = 74.9$, $\sigma = s = 7.4$

$$\text{posterior mean } \mu_1 = \frac{n\bar{x}\sigma_0^2 + \mu_0\sigma^2}{n\sigma_0^2 + \sigma^2}$$

$$= \frac{40(74.9)(1.5)^2 + (67.2)(7.4)^2}{40(1.5)^2 + (7.4)^2} = 71.987 \approx 72$$

posterior s.d.

$$\sigma_1 = \sqrt{\frac{\sigma^2\sigma_0^2}{n\sigma_0^2 + \sigma^2}} = \sqrt{\frac{(7.4)^2(1.5)^2}{40(1.5)^2}} = 0.922568 \approx 0.923$$

c. Here $\mu_1 = 72$, $\sigma_1 = 0.923$.

Standard variable corresponding to 650 is

$$Z_1 = \frac{65 - 72}{0.923} = -7.5839,$$

similarly, corresponding to 70.0 is

$$Z_2 = \frac{70 - 72}{0.923} = -2.16684$$

Posterior probability $= P(65 < X < 70)$

$$= P(-7.584 < Z < -2.167)$$
$$= 0.5 - 0.4850 = 0.0150$$

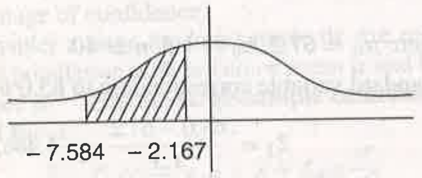


Fig. 29.2

d. 95% Bayesian interval limits are

$$\mu_1 \pm Z_{\alpha/2}\sigma_1 = 71.987 \pm (1.96)(0.922568)$$

Thus the Bayesian interval is

$$(70.17876, 72.909568).$$

EXERCISE

1. Calculate μ_1, σ_1 for the posterior distribution if the random sample size is 80, $\overline{x} = 18.85$, $s = 5.55$ using s for s.d. of population σ.

Ans: $\mu_1 = 18.77, \sigma_1 = 0.60$

2. An insurence agent feelings about the average monthly commission of insurance policies may be described by means of normal distribution with $\mu_0 =$ Rs. 3800 and $\sigma_0 =$ Rs. 260. (a) What probability is the agent thus assigning to the true average monthly commission being in the interval of Rs. 3,500 to Rs. 4000. (b) How does the probability in part (a) is affected if the mean commission is Rs. 3702 with s.d. Rs. 390 for 9 months? Use $s = 390$ as an estimate of σ.

Hint: $P(3500 < X < 4000) =$
$P(-1.154 < Z < 0.77) = 0.3749 + 0.2794.$

Ans: **a.** 0.6543

Hint: $\overline{x} = 3702$, $s = \sigma = 390$, $n = 9$,
$\mu_0 = 3800$, $\sigma_0 = 260$

$$\mu_1 = \frac{n\overline{x}\sigma_0^2 + \mu_0\sigma^2}{n\sigma_0^2 + \sigma^2} = 3721.6 \approx 3722,$$

$$\sigma_1 = \sqrt{\frac{\sigma^2\sigma_0^2}{n\sigma_0^2 + \sigma^2}} = 116.3 \approx 116$$

$P(3500 < X < 4000) =$
$P(-1.91 < Z < 2.40) = 0.4719 +$
$0.4918 = 0.9637$

Ans: **b.** 0.9637

3. The mean mark in mathematics in common entrance test will vary from year to year. If this variation of the mean mark is expressed subjectively by a normal distribution with mean $\mu_0 = 72$ and variance $\sigma_0^2 = 5.76$. (a) What probability can we assign to the actual mean mark being somewhere between 71.8 and 73.4 for the next year's test. (b) construct a 95% Bayesian interval for μ if the test is conducted for a random sample of 100 students from the next incoming class yielding a mean mark of 70 with s.d. of 8. (c) What posterior probability should we assign to the event of part (a).

Hint: $P(71.8 < X < 73.4) =$
$P(-0.083 < Z < 0.583)$
$= 0.0319 + 0.2190 = 0.2509$

Ans: **a.** 0.2509

Hint: $n = 100$, $\overline{x} = 70$, $\sigma_0 = 2.4$, $\mu_0 = 72$, $\sigma = 8$, so
$\mu_1 = 70.2, \sigma_1 = 0.7589$

C.I.: $\mu_1 \pm Z_{\alpha/2}(\sigma_1) = 70.2 \pm 1.96(0.7589)$

Ans: **b.** $68.71 < \mu < 71.69$

Hint: $P(71.8 < X < 73.4) =$
$P(2.105 < Z < 4.21) = 0.5 - 0.4821 =$
0.0179

since $Z_1 = \frac{71.8 - 70.2}{0.76} = 2.105,$

$$Z_2 = \frac{73.4 - 70.2}{0.76} = 4.2105$$

Ans: **c.** 0.0179

4. A producer of TV's believes from past experience that the mean length of life of TV's μ is a normal random variable with mean $\mu_0 = 800$ hours and standard deviation $\sigma_0 = 10$ hours. It is known that TV's have mean length of life that is approximately normally distributed with a standard deviation of 100 hours. Construct a 95% Bayesian interval for μ if a random sample of 25 TV's has an average life of 780 hours.

Hint: $\mu_1 = \frac{(25)(780)(10)^2 + (800)(100)^2}{25(10)^2 + (100)^2} = 796$

$$\sigma_1 = \sqrt{\frac{(10)^2(100)^2}{25(10)^2 + (100)^2}} = \sqrt{80},$$

B.I.: $780 \pm 1.96\left(\frac{100}{\sqrt{25}}\right)$

Here $Z_{\alpha/2} = 1.96$, $n = 25$, $\sigma = 100$,
$\sigma_0 = 10$, $\overline{x} = 780$.

Ans: $740.8 < \mu < 819.2$

29.4 TEST OF HYPOTHESIS

The principal objective of statistical inference is to draw inferences (or generalize) about the population on the basis of data collected by sampling from the population. Statistical inference consists of two major areas, estimation and tests of hypothesis. Estimation was discussed in Sections 29.1, 29.2 and 29.3. In tests of hypothesis, a postulate or conjecture or statement about a parameter of the population is tested for its validity or truthfulness.

Statistical decisions

Statistical decisions are decisions or conclusions about the population parameters on the basis of a random sample from the population.

Statistical Hypothesis

It is an assumption or conjecture or guess about the parameter(s) of population distribution(s). The statistical hypothesis is established before hand and may or may not be true. When more than one population is considered, statistical hypothesis consists of relationship between the parameters of the populations.

Null Hypothesis

It is (N.H.) denoted by H_0 is the statistical hypothesis which is to be actually tested for acceptance or rejection. N.H. is the hypothesis which is tested for possible rejection under the assumption that it is true (R.A. Fisher).

Alternative Hypothesis

(A.H.) denoted by H_1, is any hypothesis other than the null hypothesis. Neyman originated the concept of alternative hypothesis.

Test of Hypothesis

Test of hypothesis or test of significance or rules of decision is a procedure to decide whether to accept or reject the (null) hypothesis. This test determines whether observed samples differ significantly from expected results. Acceptance of hypothesis merely indicates that the data do not give sufficient evidence to refute the hypothesis. Whereas, rejection is a firm conclusion where the sample evidence refutes it.

When N.H. is accepted, result is said to be non-significant i.e., observed differences are due to 'chance' caused by the process of sampling. When N.H. is rejected (i.e., A.H. is accepted) the result is said to be significant. Thus test of hypothesis decides whether a statement concerning a parameter is true or false instead of estimating the value of the parameter. Since the test is based on sample observations, the decision of acceptance or rejection of the null hypothesis is always subjected to some error i.e., some amount of risk.

Types of errors in test of hypothesis:

	Accept H_0	Reject H_0
H_0 is true	correct decision	Type I error
H_0 is false	Type II error	correct decision

Type I error involves rejection of null hypothesis when it should be accepted (as true).

Type II error involves acceptance of the null hypothesis when it is false and should be rejected.

Level of significance

(L.O.S.) of a test denoted by α is the probability of committing type I error. Thus L.O.S. measures the amount of risk or error associated in taking decisions. It is costomary to fix α before sample information is collected and to choose (take) generally α as 0.05 or 0.01. L.O.S. $\alpha = 0.01$ is used for high precision and $\alpha = 0.05$ for moderate precision. L.O.S. is also expressed as percentage. Thus L.O.S. $\alpha = 5\%$ means there are 5 chances in 100 that N.H. is rejected when it is true or one is 95% confident that a right decision is made. L.O.S. is also known as the **size of the test**. Thus α = probability of committing type I error $= P$ (reject H_0/H_0) $= \alpha$ and β = prob (type II error) $= P$ (accept H_0/H_1) $= \beta$.
Power of the test is computed as $1 - \beta$.

Note 1: When the size of the sample is increased, the probability of committing both types of errors I and II i.e., α and β can be reduced simultaneously.

Note 2: α and β are known as producer's risk and consumer's risk respectively.

Note 3: When both α and β are small, the test procedure is good one giving good chance of making the correct decision.

Simple Hypothesis

It is a statistical hypothesis which completely specifies an exact parameter. Null hypothesis is always a simple hypothesis stated as an equality specifying an exact value of the parameter (includes any value not stated by A.H.).

Examples:

1. N.H. $= H_0 : \mu = \mu_0$

   i.e., population mean equals to a specified constant μ_0.

2. N.H. $= H_0 : \mu_1 - \mu_2 = \delta$

   i.e., the difference between the sample means equals to a constant δ.

Composite Hypothesis

It is stated in terms of several possible values i.e., by an inequality.

Alternative Hypothesis

It is a composite hypothesis involving statements expressed as inequalities such as $<$, $>$ or $\neq$.

Examples:

1. A.H.: $H_1; \mu > \mu_0$
2. A.H.: $H_1 : \mu < \mu_0$
3. A.H.: $H_1 : \mu \neq \mu_0$

Critical Region (C.R.)

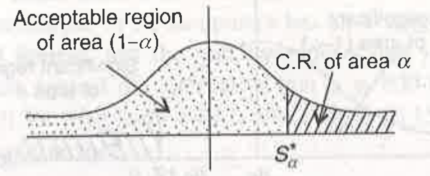
In any test of hypothesis, a test statistic S^*, calculated from the sample data, is used to accept or reject the null hypothesis of the test. Consider the area under the probability curve of the sampling distribution of the test statistic S^* which follows some known (given) distribution. This area under the probability curve is divided into to dichotomous regions, namely the region of rejection (significant region or critical region) where N.H. is rejected, and the region of acceptance (non-significant region or non-critical region) where N.H. is accepted. Thus **critical region** is the region of rejection of N.H. The area of the critical region equals to the level of significance α. Note that C.R. always lies on the tail (s) of the distribution. Depending on the nature of A.H., C.R. may lie on one side or both sides of the tails (s).

Critical value(s) or significant value(s)

It is (area) the value of the test statistic S^*_α (for given level of significance α) which divides (or separates) the area under the probability curve into critical (or rejection) region and non-critical (or acceptance) region.

One tailed test (O.T.T.) and two tailed test (T.T.T.)

Right one tailed test (R.O.T.T.): When the alternative hypothesis (A.H.): H_1 is of the greater than type i.e., $H_1 : \mu > \mu_0$ or $H_1 : \sigma_1^2 > \sigma_2^2$ etc., then the entire criticial region of area α lies on the right side tail of the probability density curve as shown shaded in the Fig. 29.3. In such case, the test of hypothesis (T.O.H.) is known as **right one tailed test**.



One-sided right tail test

Fig. 29.3

Left one tailed test (L.O.T.T.)

When the A.H.: H_1 is of the less than type i.e., H_1 $\mu_1 < \mu_0$ or $H_1 : \sigma_1^2 < \sigma_2^2$ etc. then the entire C.R. of area α lies on the left side tail of the curve as shown in Fig. 29.4.

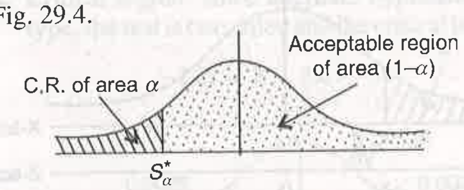


Fig. 29.4

Two tailed test (T.T.T.)

If A.H. is of the not equals type i.e., $H_1 : \mu_1 \neq \mu_2$ or $H_1 : \sigma_1 \neq \sigma_2$ etc. then the C.R. lies on both sides of the right and left tails of the curve such that the C.R. of area $\frac{\alpha}{2}$ lies on the right tail and C.R. of area $\frac{\alpha}{2}$ lies on the left tail, as shown in Fig. 29.5.

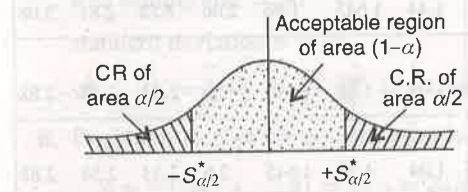


Fig. 29.5

Thus the test of hypothesis or test of significance or rule of decision consists of the following six steps.

1. Formulate N.H.: H_0.
2. Formulate A.H.: H_1.
3. Choose L.O.S.: α.
4. C.R.: is determined by the critical value S^*_α and the kind of A.H. (based on which the test is R.O.T.T. or L.O.T.T. or T.T.T.).
5. Compute the test statistic S^* using the sample data.
6. Decision: Accept or reject N.H. depending on the relation between S^* and S^*_α.

P-Value

In tests of hypothesis, preselection of a significance level α does not account for values of test statistics that are "close" to the critical region. Thus a test statistic value that is non-significant say for $\alpha = 0.05$ may become significant for $\alpha = 0.01$. In applied statistics, P-value approach is designed to give the user an alternative (in terms of probability) to a mere "reject" or "do not reject" conclusion.

***P*-Value** is the lowest level (of significance) at which the observed value of the test statistic is significant.

In the significance testing by P-Value approch, α is not pre-determined but the conclusion is based on the size of the P-Value which is computed using the value of test statistic.

29.5 TEST OF HYPOTHESIS CONCERNING SINGLE POPULATION MEAN μ: (WITH KNOWN VARIANCE σ^2: LARGE SAMPLE)

Let μ and σ^2 be the mean and variance of a population from which a random sample of size n is drawn. Let $\overline{x}$ be the mean of the sample. Then for large samples ($n \geq 30$), from central theorem it follows that the sampling distribution of $\overline{x}$ is approximately normally distributed with mean $\mu_{\overline{x}} = \mu$ and $\sigma^2_{\overline{x}} = \frac{\sigma^2}{n}$.

The test statistic for single mean with known variance is $Z = \frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}$.

To test whether the population mean μ equals to a specified constant μ_0 or not, formulate the test of hypothesis as follows:

1. N.H.: $\mu = \mu_0$.
2. A.H.: $\mu \neq \mu_0$.
3. L.O.S.: α.
4. C.R.: Since the A.H. is a not equal to type, a T.T.T. is considered. For given α, critical values

$-Z_{\alpha/2}$ and $+Z_{\alpha/2}$ are determined from normal table since normal distribution is assumed. For example, for $\alpha = 5\%$ or 0.05 from normal table $-Z_{0.025} = -1.96$ and $Z_{0.025} = 1.96$. Thus the critical region consists of the two shaded regions in Fig. 29.6 i.e., reject N.H.: H_0 if $Z < -Z_{\alpha/2}$ or $Z > Z_{\alpha/2}$.

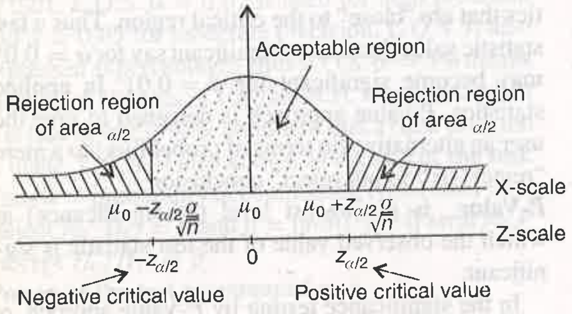


Fig. 29.6

5. Compute the test statistic Z, denoted by Z_{cal} or simply Z by

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

Here $\bar{x}$, the mean of the sample of size n, is calculated from the sample data.

6. Conclusion: Reject H_0 if Z_{cal} or Z falls in the critical region i.e., observed sample statistic is probably significant or highly significant at α level. Otherwise accept H_0 (if $-Z_{\alpha/2} < Z < Z_{\alpha/2}$).

Note 1: Suppose the A.H. is $H_1 : \mu > \mu_0$. Then the critical region is given by $Z > Z_\alpha$ since we consider a right one tail test is this case, i.e., reject H_0 if $Z > Z_\alpha$ otherwise accept H_0 (if $Z < Z_\alpha$) (see Fig. 29.7).

Note 2: If A.H. is $H_1 : \mu < \mu_0$ then consider a left one tail test with C.R. given by $Z < -Z_\alpha$ as shown in Fig. 29.8, i.e., reject H_0 is $Z < -Z_\alpha$ otherwise accept H_0 (if $Z > -Z_\alpha$).

Note 3: Reference table of critical values for a given L.O.S. α for T.T.T., R.O.T.T. and L.O.T.T.

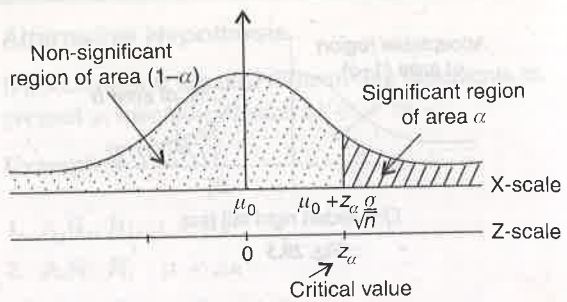


Fig. 29.7

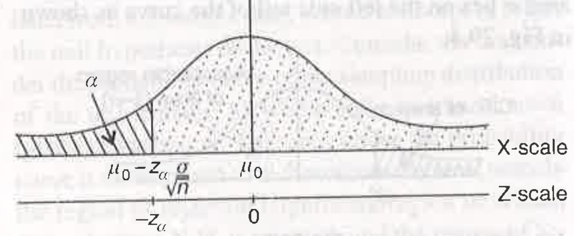


Fig. 29.8

$\alpha\%$ α	15% 0.15	10% 0.1	5% 0.05	4% 0.04	1% 0.01	.5% 0.005	.2% 0.002
$-Z_{\alpha/2}$ and $+Z_{\alpha/2}$ for T.T.T.	−1.44 and 1.44	−1.645 and 1.645	−1.96 and 1.96	−2.06 and 2.06	−2.58 and 2.58	−2.81 and 2.81	−3.08 and 3.08
$-Z_\alpha$ for L.O.T.T.	−1.04	−1.28	−1.645	−2.6	−2.33	−2.58	−2.88
Z_α for R.O.T.T.	1.04	1.28	1.645	2.6	2.33	2.58	2.88

Note 4: For large sample $n \geq 30$, even if σ is *unknown*, σ can be replaced by sample variances (which can be computed from sample information).

WORKED OUT EXAMPLES

Test of hypothesis: For one mean

Example 1: The length of life X of certain computers is approximately normally distributed with mean

800 hours and standard deviation 40 hours. If a random sample of 30 computers has an average life of 788 hourse, test the null hypothesis that $\mu = 800$ hours against the alternative that $\mu \neq 800$ hours at (a) 0.5% (b) 1% (c) 4% (d) 5% (e) 10% (f) 15% level of significance.

Solution:

Case a:

1. Null hypothesis: $\mu = 800$ hours
2. Alternate hypothesis : $\mu \neq 800$ hours
3. α level of significance $= .5\% = .005$ (case (a))
4. Critical region: since alternate hypothesis is $\neq$ type, the test is two tailed and the critical region is

$$-2.81 < Z < 2.81$$

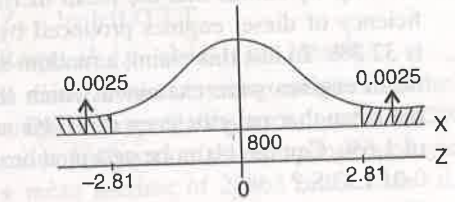


Fig. 29.9

5. Calculation of statistic:

Here $\bar{x}$ = mean of the sample = 788

n = sample size = 30

standard deviation $\sigma = 40$,

so $Z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{788-800}{40/\sqrt{30}} = -1.643$

6. Decision: Accept the null hypothesis H_0 since

$$Z = -1.643 > -2.81 = Z_{\alpha/2} = Z_{0.0025}$$

Case b: α = level of significance = 1% = 0.01 critical region

$$-2.58 < Z < 2.58$$

Desision: Accept H_0 since
$Z = -1.643 > -2.58 = Z_{\alpha/2} = Z_{0.005}$ (Fig. 29.10)

Case c: $\alpha = 4\% = 0.04$, C.R.: $-2.06 < Z < 2.06$. Accept H_0 since $Z = -1.643 > -2.06$ (Fig. 29.11)

Case d: $\alpha = 5\% = 0.05$, C.R.: $-1.96 < Z < 1.96$. Accept H_0 since $Z = -1.643 > -1.96$ (Fig. 29.12)

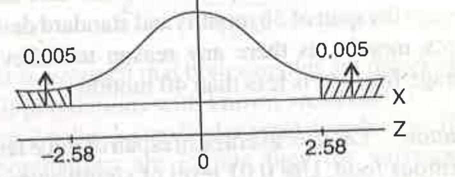


Fig. 29.10

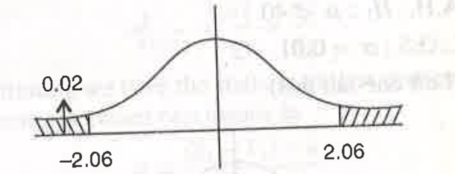


Fig. 29.11

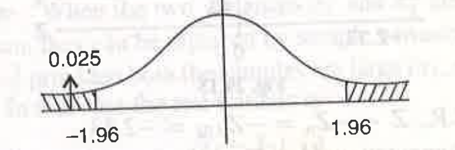


Fig. 29.12

Case e: $\alpha = 10\% = 0.10$, C.R.: $-1.645 < Z < 1.645$. Accept H_0 since $Z = -1.643 > -1.645$

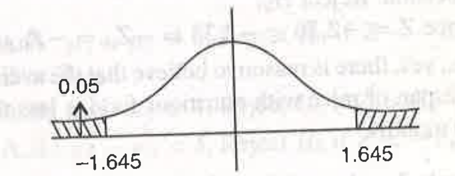


Fig. 29.13

Case f: $\alpha = 15\% = 0.15$, C.R.: $-1.44 < Z < 1.44$. Reject H_0 since $Z = -1.643 < -1.44$

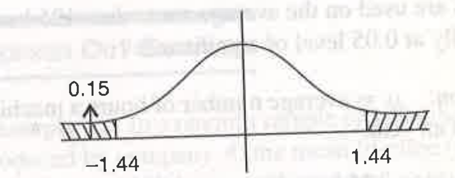


Fig. 29.14

Test is significant.

Example 2: Mice with an average lifespan of 32 months will live upto 40 months when fed by a certain nutrious food. If 64 mice fed on this diet have an average lifespan of 38 months and standard deviation of 5.8 months, is there any reason to believe that average lifespan is less than 40 months.

Solution: Let μ = average lifespan of mice fed with nutritious food. Use 0.01 level of significance

1. N.H.: $H_0 : \mu = 40$ months
2. A.H.: $H_1 : \mu < 40$
3. L.O.S.: $\alpha = 0.01$
(Left one-tail test)

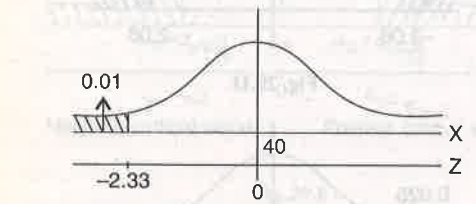


Fig. 29.15

4. C.R.: $Z < -Z_\alpha = -Z_{0.01} = -2.33$
5. Computation: Here $\overline{x} = 38$, $\sigma = 5.8$, $n = 64$

$$Z = \frac{38 - 40}{5.8/\sqrt{64}} = -2.76$$

6. Decision: Reject H_0,
since $Z = -2.76 < -2.33 = -Z_\alpha = -Z_{0.01}$
i.e., yes, there is reason to believe that the average lifespan of mice with nutritious food is less than 40 months.

Example 3: A machine runs on an average of 125 hours/year. A random sample of 49 machines has an annual average use of 126.9 hours with standard deviation 8.4 hours. Does this suggest to believe that machines are used on the average more than 125 hours annually at 0.05 level of significance?

Solution: μ = average number of hours a machine runs in an year.

1. $H_0 \cdot \mu = 125$ hours/year
2. $H_1 \cdot \mu > 125$
3. L.O.S.: $\alpha = 0.05$
4. C.R.: $Z > Z_\alpha = Z_{0.05} = 1.64$
5. Calculation: $Z = \frac{\overline{x}-\mu}{\sigma/\sqrt{n}} = \frac{126.9-125}{8.4/\sqrt{49}} = 1.58$

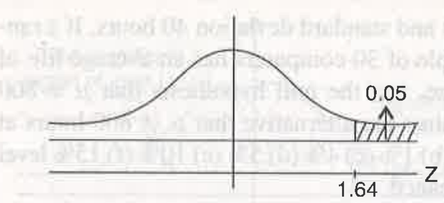


Fig. 29.16

6. Decision: Accept H_0 since $Z = 1.58 < 1.64 = Z_{0.05}$ i.e., can not believe that machine works more than 125 hours in an year.

EXERCISE

Test of hypothesis: For one mean

1. A company claims that the mean thermal efficiency of diesel engines produced by them is 32.3%. To test this claim, a random sample of 40 engines were examined which showed the mean thermal efficiency of 31.4% and s.d of 1.6%. Can the claim be accepted or not, at 0.01 L.O.S.?

   Hint: $Z = \frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{31.4-32.3}{1.6/\sqrt{40}} = -3.56 < -2.81 = Z_\alpha = Z_{0.01}$

Ans: Reject N.H.: $H_0 \cdot \mu_0 = 32.3$ from a T.T.T. against A.H.: $H_1 : \mu_0 \neq 32.3$

2. It has previously been recorded that the average depth of ocean at a particular region is 67.4 fathoms. Is there reason to believe this at 0.01 L.O.S. if the readings at 40 random locations in that particular region showed a mean of 69.3 with s.d. of 5.4 fathoms?

   Hint: $Z = \frac{69.3-67.4}{5.4/\sqrt{40}} = 2.23 < 2.58 = Z_{\alpha/2} = Z_{0.005}$

Ans: Accept $H_0 : \mu_0 = 67.4$ against A.H.: $H_1; \mu_0 \neq 67.4$ by a T.T.T.

3. To determine whether the mean breaking strength of synthetic fibre produced by a certain company is 8 kg or not, a random sample of 50 fibres were tested yielding a mean breaking strength of 7.8 kg. If s.d. is 0.5 kg, test at 0.01 L.O.S.

   Hint: $Z = \frac{78-8}{0.5/\sqrt{50}} = -2.83 > -2.575 = Z_{\alpha/2} = Z_{0.005}$

Ans: Reject $H_0 : \mu = 8$, accept A.H.: $H_1 : \mu \neq 8$ by a T.T.T.

4. Can it be concluded that the average lifespan of Indian is more than 70 years if a random sample of 100 Indians has an average lifespan of 71.8 years with a s.d. of 8.9 years.

   Hint: $Z = \frac{71.8-70}{8.9/\sqrt{100}} = 2.02 > 1.645 = Z_\alpha = Z_{0.05}$ by right O.T.T.

Ans: Yes, average lifespan is more than 70 years.

5. A company producing computers states that the mean lifetime of its computers is 1600 hours. Test this claim at 0.01 L.O.S. against the A.H.: $\mu < 1600$ hours if 100 computers produced by this company has mean lifetime of 1570 hours with s.d. of 120 h.

   Hint: $Z = \frac{(1570-1600)}{120/\sqrt{100}} = -2.50 < -2.33 = Z_{0.01}$ by left O.T.T.

Ans: Reject H_0 i.e., claim is not tenable

6. A manufacturer of tyres guarantees that the average lifetime of its tyres is more than 28000 miles. If 40 tyres of this company tested, yields a mean lifetime of 27463 miles with s.d. of 1348 miles, can the guarantee be accepted at 0.01 L.O.S.?

   Hint: $Z = \frac{27463-28000}{1348/\sqrt{40}} = -2.52 < -2.33 = Z_{0.01}$ by left O.T.T.

Ans: No, tyres run for < 28000 miles.

29.6 TEST OF HYPOTHESIS CONCERNING TWO MEANS

When variances σ_1 and σ_2 are known or large samples

Let $\overline{x}_1$ be the mean of a random sample of size n_1 drawn from a population with mean μ_1 and variance σ_1^2. Let $\overline{x}_2$ be the mean of an independent random sample of size n_2 drawn from another population with mean μ_2 and variance σ_2^2. To test the hypothesis for difference of means, consider the null hypothesis $\mu_1 - \mu_2 = \delta$ = given constant. So when $\delta = 0$, there is no difference between the means i.e., the two populations have the same means. If $\delta \neq 0$, the means of the two populations are different. In these cases, the test statistic will depend on the difference between the sample means $\overline{x}_1 - \overline{x}_2$ and is given by

$$Z = \frac{(\overline{x}_1 - \overline{x}_2) - \delta}{\sigma_{\overline{x}_1 - \overline{x}_2}}$$

Here it is assumed that both samples are drawn from normal populations with known variances.

Here Z follows standard normal distribution. If the two populations are infinite then the variance of the sampling distribution of the difference between the sample means is

$$\sigma^2_{\overline{x}_1 - \overline{x}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

Substituting we have the statistic for test concerning difference between two means as

$$Z = \frac{(\overline{x}_1 - \overline{x}_2) - \delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Note: When the two variances σ_1^2 and σ_2^2 are unknown, they can be replaced by sample variances s_1^2 and s_2^2 provided both the samples are large ($n_1, n_2 \geq 30$). In this case the test statistic is

$$Z = \frac{(\overline{x}_1 - \overline{x}_2) - \delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

The critical regions for testing $\mu_1 - \mu_2 = \delta$ are:

1. A.H.: $\mu_1 - \mu_2 \neq \delta$, Reject H_0 if $Z < -Z_{\alpha/2}$ or $Z > Z_{\alpha/2}$
2. A.H.: $\mu_1 - \mu_2 > \delta$, Reject H_0 if $Z > Z_\alpha$
3. A.H.: $\mu_1 - \mu_2 < \delta$, Reject H_0 if $Z < -Z_\alpha$.

The A.H. 2 and 3 are used to determine whether one product (population) is better than (superior to) the other product.

WORKED OUT EXAMPLES

Example 1: In a random sample of 100 tube lights produced by company A, the mean lifetime (mlt) of tube light is 1190 hours with standard deviation of 90 hours. Also in a random sample of 75 tube lights from company B the mean lifetime is 1230 hours with standard deviation of 120 hours. Is there a difference between the mean lifetimes of the two brands of tube lights at a significance level of (a) 0.05 (b) 0.01?

Solution: Let X_A, X_B denote the lifetime (in hours) of tube lights produced by company A and B respectively. It is given that the mean lifetime of tube lights of company A is $\overline{X}_A = 1190$, standard deviation for tube lights of A is $s_A = 90$. Similarly $\overline{X}_B = 1230, s_B = 120, n_A =$ sample size of tube lights from $A = 100$, $n_B =$ sample size from $B = 75$

1. Null hypothesis: $H_0 : \mu_1 - \mu_2 = \delta = 0$ i.e., no difference.
2. Alternate hypothesis: $H_1 : \mu_1 - \mu_2 \neq 0$ i.e., there is difference.
3. L.O.S.: α : (a) 0.05 (b) 0.01.
4. Critical region: two tailed test.

If

a. $-1.96 < Z < 1.96$ Accept N.H.

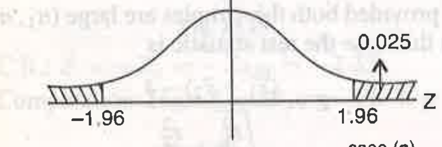


Fig. 29.17

b. $-2.57 < Z < 2.57$ Accept N.H.

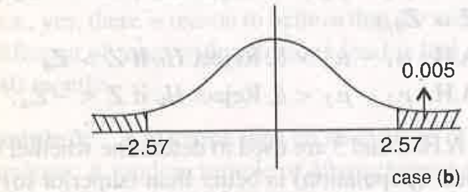


Fig. 29.18

5. Computation: $\mu_{\overline{X}_A - \overline{X}_B} = \mu_{\overline{X}_A} - \mu_{\overline{X}_B} = \mu_A - \mu_B = 0$

$$\sigma_{\overline{X}_A - \overline{X}_B} = \sqrt{\sigma^2_{\overline{X}_A} + \sigma^2_{\overline{X}_B}} = \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}$$

$$= \sqrt{\frac{(90)^2}{100} + \frac{(120)^2}{75}} = 16.5227$$

Test statistic:

$$Z = \frac{(\overline{X}_A - \overline{X}_B) - (\mu_A - \mu_B)}{\sigma_{\overline{X}_A - \overline{X}_B}} = \frac{1190 - 1230}{16.5227} = -2.421$$

6. Decision:

a. For $\alpha = 0.05$

Reject N.H. since $Z = -2.421 < -1.96 = Z_{\alpha/2} = Z_{0.025}$ i.e., yes, there is difference between the mean lifetimes of the tube lights produced by A and B.

b. For $\alpha = 0.01$

Accept N.H. since $Z = -2.421$ lies in the acceptable region $-2.57 < Z < 2.57$ i.e., no, there is no difference between $\overline{X}_A$ and $\overline{X}_B$.

Example 2: To test the effects a new pesticide on rice production, a farm land was divided into 60 units of equal areas, all portions having identical qualities as to soil, exposure to sunlight etc. The new pesticide is applied to 30 units while old pesticide to the remaining 30. Is there reason to believe that the new pesticide is better than the old pesticide if the mean number of kgs of rice harvested / unit using new pesticide (N.P.) is 496.31 with s.d. of 17.18 kgs while for old pesticide (O.P.) is 485.41 kgs and 14.73 kgs. Test at a level of significance (a) $\alpha = 0.05$ (b) 0.01.

Solution: Let the subscripts N and O denote respectively the quantities related the new pesticide and old pesticide.

1. N.H.: $H_0 : \mu_N - \mu_O = \delta = 0$ i.e., no difference.
2. A.H.: $H_1 : \mu_N - \mu_O = \delta > 0$ i.e., new pesticide is superior to (better than) old pesticide.
3. L.O.S.: (a) $\alpha = 0.05$ (b) 0.01.
4. Critical region: Right one tailed test.

Case a: $\alpha = 0.05$

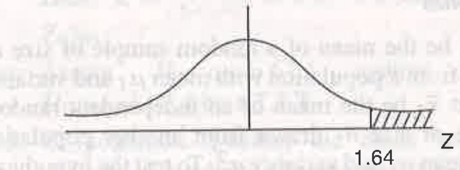


Fig. 29.19

Accept N.H. if $Z < Z_\alpha = Z_{0.05} = 1.64$

Case b: $\alpha = 0.01$

Accept N.H. if $Z < Z_\alpha = Z_{0.01} = 2.33$

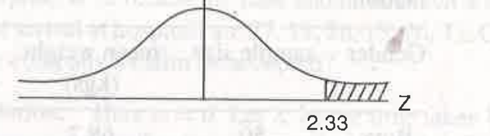


Fig. 29.20

5. Computation: Given data is

$\overline{X}_N = 496.31, \overline{X}_O = 485.41, s_N = 17.18, s_O = 14.73, n_N = 30, n_O = 30$

Test statistic is

$$Z = \frac{(\overline{X}_N - \overline{X}_O) - (\mu_N - \mu_O)}{\sigma_{\overline{X}_N - \overline{X}_O}} = \frac{(\overline{X}_N - \overline{X}_O) - 0}{\sqrt{\frac{(s_N)^2}{n_N} + \frac{(s_O)^2}{n_O}}}$$

$$= \frac{(496.31 - 485.41) - 0}{\sqrt{\frac{(17.18)^2}{30} + \frac{(14.73)^2}{30}}} = 2.63814$$

6. Decision

Case a: $\alpha = 0.05$

Reject N.H. since $Z = 2.638 > Z_\alpha = Z_{0.05} = 1.64$ i.e., accept A.H. or new pesticide is superior to old pesticide.

Case b: $\alpha = 0.01$

Reject N.H. since $Z = 2.638 > Z_\alpha = Z_{0.01} = 2.33$ i.e., Accept A.H. or new pesticide is better than the old pesticide.

EXERCISE

1. A random sample of 40 'geyers' produced by company A have a mean lifetime (mlt) of 647 hours of continuous use with a s.d. of 27 hours, while a sample 40 produced by another company B have mlt of 638 hours with s.d. 31 hours. Does this substantiate the claim of company A that their 'geyers' are superior to those produced by company B at (a) 0.05 (b) 0.01 L.O.S.

Hint: N.H.: $\mu_A = \mu_B = 0$, A.H.: $\mu_A - \mu_B > 0$, reject N.H. if $Z > 1.645$.

$$Z_{\text{cal}} = \frac{(647 - 638) - 0}{\sqrt{\frac{(27)^2}{40} + \frac{(31)^2}{40}}} = 1.38, \text{ accept N.H.}$$

Ans: **a.** No, there is no difference between 'geyers' produced by the two companies A and B.

b. Accept N.H. since $1.38 < Z_\alpha = 2.33$

2. Test at 0.05 L.O.S. a manufacturer's claim that the mean tensile strength (mts) of a thread A exceeds the mts of thread B by at least 12 kgs. if 50 pieces of each type of thread are tested under similar conditions yielding the following data:

	sample size	mts (kgs)	s.d. (kgs)
Type A	50	86.7	6.28
Type B	50	77.8	5.61

Hint: $H_0 \cdot \mu_A - \mu_B \geq 12$, $H_1 \cdot \mu_A - \mu_B < 12$, reject H_0 if $Z < Z_\alpha = -1.64$

$Z = \frac{(86.7 - 77.8) - 12}{\sqrt{\frac{(6.28)^2}{50} + \frac{(5.61)^2}{50}}} = -2.60$, reject H_0,

Accept $\mu_1 - \mu_2 < 12$.

Ans: Claim not tenable.

3. Test the N.H.: $\mu_A - \mu_B = 0$ against the A.H.: $\mu_A - \mu_B \neq 0$ at 0.01 L.O.S. for the following data:

	sample size	mts (kgs)	s.d. (kgs)
Type A	40	247.3	15.2
Type B	30	254.1	18.7

Hint: $Z = \frac{(247.3 - 254.1) - 0}{\sqrt{\frac{(15.2)^2}{40} + \frac{(18.7)^2}{30}}} = -1.62866$,

Acceptable region: $-2.58 < Z < 2.58$, Accept N.H.

Ans: Accept N.H. i.e., no difference between type A and B

4. If random sample data show that 42 men earn on the average $\overline{x}_1 = 744.85$ with s.d. $s_1 = 397.7$ while 32 women earn on the average $\overline{x}_2 = 516.78$ with s.d. $s_2 = 162.523$, test at 0.05 level of significance whether the average

income for men and women is same or not.

Hint: $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$,

$$Z = \frac{(744.85-516.78)-0}{\sqrt{\frac{158165.43}{42}+\frac{26413.61}{32}}} = 3.36$$

Since $Z = 3.36 > 1.96 = Z_\alpha = Z_{0.05}$, Reject N.H. H_0.

Ans: Not same.

5. A company claims that alloying reduces resistance of electric wire by more than 0.050 ohm. To test this claim samples of standard wire and alloyed wire are tested yielding the following results:

Type of wire	sample size	mean resistance (ohms)	s.d. (ohms)
Standard	32	0.136	0.004
Alloyed	32	0.083	0.005

Can the claim be substantiated at 0.05 L.O.S.

Hint: $H_0 : \mu_1 - \mu_2 = 0.05$, $H_1 : \mu_1 - \mu_2 > 0.05$, $Z_\alpha = Z_{0.05} = 1.645$

$$Z = \frac{(0.136-0.083)-(0.05)}{\sqrt{\frac{(0.004)^2}{32}+\frac{(0.005)^2}{32}}} = 2.65,$$

Reject N.H. since $Z = 2.65 > 1.645 = Z_{0.05}$

Ans: Data substantiate the claim.

6. To test the claim that men are taller than women, a survey was conducted resulting in the following data:

Gender	sample size	mean height (cm)	s.d. (cm)
Men	1600	172	6.3
Women	6400	170	6.4

Is the claim tenable at 0.01 L.O.S.

Hint: $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 > \mu_2$,

$$Z = \frac{172-170}{\sqrt{\frac{(6.3)^2}{1600}+\frac{(6.4)^2}{6400}}} = 11.32$$

Reject H_0 since $Z = 11.32 > 2.33 = Z_\alpha = Z_{0.01}$.

Ans: Yes, men are taller than women.

7. Test the claim that teen-age boys are heavier than teen-age girls given the following information:

Gender	sample size	mean weight (kgs)	s.d. (kgs)
Boys	50	68.2	2.5
Girls	50	67.5	2.8

Use L.O.S. (a) 0.05 (b) 0.01

Hint: $Z = \frac{\overline{X}_1-\overline{X}_2}{\sigma_{\overline{X}_1-\overline{X}_2}} = \frac{68.2-67.5}{0.53} = 1.32,$

$$\sigma_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{(2.5)^2}{50}+\frac{(2.8)^2}{50}} = 0.53$$

Ans: **a.** Accept N.H. i.e., no difference between mean weights.

b. Reject N.H. i.e., boys are heavier than girls.

29.7 TEST FOR ONE MEAN (SMALL SAMPLE: t-DISTRIBUTION)

For "expensive" populations such as satelites, aeroplanes, nuclear reactors, super computers, etc. the investigation of characteristics of large samples ($n \geq 30$) is uneconomical, impractiable and time consuming. In all such cases, the size of the sample, drawn is small (i.e., $n < 30$). For σ unknown and for small sample size, the test statistic cannot be used. Then the decision criterion is based on the t-distribution with $\nu = n - 1$ degrees of freedom. Thus the test statistic for small sample test (with σ unknown) concerning one mean is

$$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}}$$

This is also known as "one-sample t-test". So the test procedure for small samples is similar to the procedure for large samples except that 't' values are used in place of Z values and σ is replaced by s. For example, when A.H. is $\mu \neq \mu_0$ then the C.R. is $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$ etc.

WORKED OUT EXAMPLES

Examples: An ambulance service company claims that on an average it takes 20 minutes between a call for an ambulance and the patient's arrival at the hospital. If in 6 calls the time taken (between a call and arrival at hospital) are 27, 18, 26, 15, 20, 32. Can the company's claim be accepted?

Solution: Here $n = 6$. Let X be the time taken between a call and patient's arrival at hospital. From given data $\overline{X}$ = average time taken

$$\overline{X} = \frac{27+18+26+15+20+32}{6} = \frac{138}{6} = 23$$

standard deviation: $s = \sqrt{\frac{\sum(X_i-\overline{X})^2}{n-1}}$.

$$s^2 = \frac{(27-23)^2+(18-23)^2+(26-23)^2+(15-23)^2+(20-23)^2+(32-23)^2}{6-1}$$

$s^2 = 40.8$, $s = 6.38748$

1. N.H.: $X = 20$ minutes

2. A.H.: $X > 20$

3. L.O.S.: $\alpha = 0.10$

4. Critical region:

Reject N.H. if $t > t_\alpha = 1.476$ where $t_{0.10}$ with $\nu = n - 1 = 6 - 1 = 5$ degrees of freedom.

5. Calculation: $t = \frac{\overline{X}-\mu}{s/\sqrt{n}} = \frac{23-20}{6.39/\sqrt{6}} = 1.15$

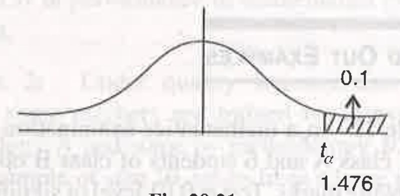


Fig. 29.21

6. Decision: Accept H_0 since $t = 1.15 < 1.476 = t_{0.1}$ with 5 dof i.e., accept the claim of the company.

EXERCISE

1. Mean lifetime (mlt) of computers manufactured by a company is 1120 hours with standard deviation of 125 hours. (a) Test the hypothesis that mean lifetime of computers has not changed if a sample of 8 computers has a mlt of 1070 hours (b) Is there decrease in mlt? Use (i) 0.05 (ii) 0.01 L.O.S.

Hint: N.H.: $\mu = 1120$, A.H.: $\mu \neq 1120$,

$t = \frac{1070-1120}{125/\sqrt{8}} = -1.1313$, $t_{0.005}$ with 7 dof is ± 3.499, $t_{0.025}$ with 7 dof is ± 2.365:

Accept H_0 in both cases.

Ans: **a.** Two-tailed test indicates that there is no reason at either level to belive that mlt has changed.

Hint: A.H.: $\mu < 1120$ (i) $t_{0.05}$ with 7 dof is -1.895 (ii) $t_{0.01}$ with 7 dof is -2.998.

b. One-tail test indicates no decrease in mlt at either of the L.O.S.

2. Producer of 'gutkha', claims that the nicotine content in his 'gutkha' on the average is 1.83 mg. Can this claim be accepted if a random sample of 8 'gutkhas' of this type have the nicotine contents of 2.0, 1.7, 2.1, 1.9, 2.2, 2.1, 2.0, 1.6 mg?

Hint: $\overline{x} = \frac{15.6}{8} = 1.95$, $s = \frac{\sqrt{0.3}}{7} = 0.20702$,

$t = \frac{1.95-1.83}{.207/\sqrt{8}} = 1.6395$, $t_\alpha = t_{0.05}$ with 7 dof is 1.895.

N.H.: $\mu = 1.83$, A.H.: $\mu > 1.83$

Ans: Yes, the producer's claim can be accepted with 95% confidence.

3. In 1950 in India the mean life expectancy was 50 years. If the life expectancies from a random sample of 11 persons are 58.2, 56.6, 54.2, 50.4, 44.2, 61.9, 57.5, 53.4, 49.7, 55.4, 57.0, does it confirm the expected view.

Hint: H_0: $\mu = 50$, $H_1 : \mu \neq 50$, $\overline{x} = \frac{598.5}{11} = 54.41$, $s = 4.859$, $t = 3.01$, reject H_0 since $t = 3.01 > 2.228 = t_{0.0025}$ with 10 dof.

Ans: No, the life expectancy is more than 50 years.

4. An auditor claims that he takes on an average 10.5 days to file income tax returns (I.T. returns). Can this claim be accepted if a random sample shows that he took 13, 19, 15, 10, 12, 11, 14, 18 days to file I.T. returns? Use (a) 0.01 (b) 0.05 L.O.S.

Hint: N.H.: $\mu = 10.5$, A.H.: $\mu > 10.5$,

$\overline{x} = \frac{112}{8} = 14$, $s = \sqrt{\frac{72}{7}} = 3.207$,

$t = \frac{14-10.5}{3.207/\sqrt{8}} = 3.0869$, (a) $t_{0.01}$ with 7 dof is 2.998 (b) $t_{0.05}$ with 7 dof is 1.895.

Ans: Reject the claim, i.e., it takes more than 10.5 days to file I.T. returns.

5. If 5 pieces of certain ribbon selected at random have mean breaking strength of 169.5 pounds with s.d. of 5.7, do they confirm to the specification mean breaking strength of 180 pounds.

   Hint: H_0: $\mu = 180$, $H_1 : \mu < 180$, L.O.S.: $\alpha = 0.01$, $t = -4.12$,

   $t_\alpha = t_{0.01}$ with 4 dof is -3747, so reject N.H.

Ans: They do not confirm to specification i.e., mbs is below.

6. In a random sample of 10 bolts produced by a machine the mean length of bolt is 0.53 mm and standard deviation 0.03 mm. Can we claim from this that the machine is in proper working order if in the past it produced bolts of length 0.50 mm? Use (a) 0.05 (b) 0.01 L.O.S.

   Hint: $H_0 : \mu = 0.50$, $H_1 : \mu \neq 0.50$,

   $t = \frac{\bar{x}-\mu}{s}\sqrt{n-1} = \frac{0.53-0.50}{0.03}\sqrt{10-1} = 3.0$

   Acceptable region $-2.26 < t < 2.26$

Ans: **a.** At 0.05 L.O.S., by a T.T.T., reject H_0.

b. At 0.01 L.O.S., by T.T.T., accept H_0

Acceptable region $-3.25 < t < 3.25$.

29.8 SMALL-SAMPLE TEST CONCERNING DIFFERENCE BETWEEN TWO MEANS

Suppose the two sample sizes n_1, n_2 or both are small ($n < 30$) and two samples are drawn from two normal populations with population variances σ_1^2 and σ_2^2 *unknown* but equal (i.e., $\sigma_1 = \sigma_2 = \sigma$). Then the pooling varience σ^2 is given by

$$\sigma^2 = \frac{\sum(x_{i1} - \bar{x}_1)^2 + \sum(x_{2i} - \bar{x}_2)^2}{n_1 + n_2 - 2}$$

$$= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

where $\bar{x}, s_1^2$ and $\bar{x}_2, s_2^2$ are the mean and variance of two samples of size n_1 and n_2 respectively. In a test concerning the difference between the means for small samples, the t-test statistic is

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - \delta}{\sigma_{\bar{x}_1 - \bar{x}_2}}$$

with $n_1 + n_2 - 2$ degrees of freedom. This test is also known as two-sample pooled t-test. Rewriting

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - \delta}{\sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}}\sqrt{\frac{n_1 n_2(n_1 + n_2 - 2)}{n_1 + n_2}}$$

with $n_1 + n_2 - 2$ dof.

The critical regions with this t-distribution can be obtained in a similar way. For example when A.H. is $\mu_1 - \mu_2 \neq \delta$, then the critical region (Reject H_0) is

$$t < -t_{\alpha/2}, n_1 + n_2 - 2 \quad \text{or} \quad t > t_{\alpha/2}, n_1 + n_2 - 2.$$

Note 1: The critical values are given when $n_1 + n_2 - 2 \geq 30$ (although n_1 and n_2 are small).

Note 2: The two-sample t-test can not be used if $\sigma_1 \neq \sigma_2$.

Note 3: The two-sample t-test can not be used for "before and after" kind of data, where the data is naturally paired. In other words the samples must be "independent" for two sample t-test.

WORKED OUT EXAMPLES

Example 1: In a mathematics examination 9 students of class A and 6 students of class B obtained the following marks. Test at 0.01 level of significance whether the performance in mathematics is same or not for the two classes A and B. Assume that the samples are drawn from normal populations having same variance.

A	44	71	63	59	68	46	69	54	48
B	52	70	41	62	36	50			

Solution: Let X_A and X_B be the marks obtained in mathematics of class A and class B. Then from the given data $\bar{X}_A = \frac{\sum x_i}{n_1} = \frac{522}{9} = 58$, $\bar{X}_B = \frac{311}{6} = 51.83$

$$s_A^2 = \frac{\sum(X_i - \bar{X})^2}{n_1 - 1} = \frac{872}{8} = 109, s_A = 10.44,$$

$$s_B^2 = \frac{804.8334}{5}, \quad s_B = 12.687$$

Here n_A = sample size from (population) class A = 9

n_B = sample size from (population) class B = 6

1. N.H.: $\mu_1 - \mu_2 = 0$ i.e., no difference in performance.
2. A.H.: $\mu_1 - \mu_2 \neq 0$ i.e., there is difference.
3. L.O.S.: $\alpha = 0.01$.
4. Critical region: Two-tailed test. Reject N.H. if $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$ where $t_{\alpha/2} = t_{0.005}$ with $n_1 + n_2 - 2 = 9 + 6 - 2 = 13$ degrees of freedom. From table, $t_{0.005}$ is 3.012.
5. Computation: Test statistic

$$t = \frac{(\bar{X}_A - \bar{X}_B) - (\mu_1 - \mu_2)}{\sqrt{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}}\sqrt{\frac{n_A n_B(n_A + n_B - 2)}{n_A + n_B}}$$

$$t = \frac{(58-51.83)-0}{\sqrt{(9-1)(109)+(6-1)(160.96)}}\sqrt{\frac{(9)(6)(13)}{9+6}} = 1.030.$$

6. Decision: Accept N.H. since

$$t = 1.03 < 3.012 = t_{\alpha/2} = t_{0.005}$$

i.e., there is no difference between the two classes A and B in performance in mathematics examination.

Example 2: Under quality improvement programme some teachers are trained by instruction methodology A and some by methodology B. In a random sample of size 10, taken from a large group of teachers exposed to each of these two methods, the following marks are obtained in an appropriate achievement test

Method A	65	69	73	71	75	66	71	68	68	74
Method B	78	69	72	77	84	70	73	77	75	65

Assuming that populations sampled are approximately normally distributed having same variance, test the claim that method B is more effective at 0.05 level of significance.

Solution: Let subscripts A and B denote data pertaining to methodology A and B respectively. Then from the given data, $n_A = n_B = 10$,

$\bar{X}_A$ = average marks obtained in appropriate achievement test by teachers trained under methodology A is $\frac{700}{10} = 70$. Similarly, $\bar{X}_B = \frac{740}{10} = 74$

$$s_A^2 = \frac{102}{9} = 11.33, \quad s_A = 3.366,$$

$$s_B^2 = \frac{262}{9} = 29.11, \quad s_B = 5.3954$$

1. N.H.: $H_0 : \mu_1 - \mu_2 = 0$ i.e., no difference in teaching methodologies
2. A.H.: $H_1 : \mu_1 - \mu_2 < 0$ i.e., method B is more effective (superior) than method A
3. L.O.S.: $\alpha = 0.05$
4. Critical region (left one tailed test)

   Reject H_0 if $t < -t_\alpha = -t_{0.05}$ with $n_A + n_B - 2 = 10 + 10 - 2 = 18$ degrees of freedom. From table $t_{0.05} = -1.734$
5. Computation

$$t = \frac{(\bar{X}_A - \bar{X}_B) - (\mu_1 - \mu_2)}{\sqrt{(n_A - 1)s_A^2 + (n_{B-1})s_B^2}}\sqrt{\frac{n_A n_B(n_A + n_B - 2)}{n_A + n_B}}$$

$$= \frac{(70 - 74) - 0}{\sqrt{9(11.33) + 9(29.11)}}\sqrt{\frac{(10)(10)(18)}{10 + 10}} = -1.989$$

6. Decision: Reject N.H. since $t = -1.989 < -1.734 = t_{0.05}$ i.e., accept the claim that method B is more effective (better) than the method A.

Example 3: Out of a random sample of 9 mice, suffering with a disease, 5 mice were treated with a new serum while the remaining were not treated. From the time of commencement of experiment, the following are the survival times:

Treatment	2.1	5.3	1.4	4.6	0.9
No treatment	1.9	0.5	2.8	3.1	

Test whether the serum treatment is effective in curing the disease at 0.05 L.O.S., assuming that the two distributions are normally distributed with equal variances.

Solution: Let μ_T and μ_{NT} be the mean survival times of the mice treated and not treated with serum respectively.

1. N.H.: $H_0 : \mu_T - \mu_{NT} = 0$ i.e., not effective

2. A.H.: $H_1 : \mu_T - \mu_{NT} > 0$ i.e., serum is effective

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

4. Critical region: Reject N.H. if $z > t_{0.95,7} = 1.90$ since the dof is $\nu = n_T + n_{NT} - 2 = 5 + 4 - 2 = 7$.

5. Computation: $n_T = 5, \overline{X}_T = \frac{14.3}{5} = 2.86$

$$s_T^2 = \frac{15.532}{4} = 3.883, s_T = 1.9705,$$

$$n_{NT} = 4, \overline{X}_{NT} = \frac{8.3}{4} = 2.075,$$

$$s_{NT}^2 = \frac{4.0875}{3} = 1.3625,$$

$$s_{NT} = 1.16726,$$

$$S_p^2 = \frac{(n_T - 1)s_T^2 + (n_{NT} - 1)s_{NT}^2}{n_T + n_{NT} - 2} =$$

$$S_p^2 = \frac{(5-1)(1.9705)^2 + (4-1)(1.16726)^2}{5+4-2}$$

$$= 2.802, S_p = 1.674$$

$$t = \frac{(2.86 - 2.075) - 0}{1.674\left(\frac{1}{5} + \frac{1}{4}\right)} = 0.6990 \approx 0.7$$

6. Decision: Accept N.H. since $t = 0.7 < 1.9 = t_{0.95,7}$ i.e., serum treatment is not effective.

Exercise

1. Random samples of specimens of coal from two mines A and B are drawn and their heat-producing capacity (in millions of calories/ton) were measured yielding the following results:

Mine A: 8350, 8070, 8340, 8130, 8260

Mine B: 7900, 8140, 7920, 7840, 7890, 7950

Is there significant difference between the means of these two samples at 0.01 L.O.S.

Hint: N.H.: $\mu_1 - \mu_2 = 0$, A.H.: $\mu_1 - \mu_2 \neq 0$, $t_{0.005}$ with $5 + 6 - 2 = 9$ dof is 3.250

Accept if $-3.250 < t < 3.250$, $\overline{x}_1 = \frac{41150}{5} = 8230$, $\overline{x}_2 = \frac{47640}{6} = 7940$

$s_1^2 = \frac{63000}{4} = 15750$, $s_2^2 = \frac{54600}{5} = 10920$,

$t = \frac{8230-7940}{\sqrt{63000+54600}}\sqrt{\frac{5.6.9}{11}} = 4.19$

Reject N.H. since $t = 4.19 > t_{0.005} = 3.250$.

Ans. Yes, there is significant difference.

2. To test the claim that substrate concentration (S.C.) causes an increase in the mean velocity (M.V.) of a chemical reaction by more than 0.5 m/l/30 minutes a study is conducted resulting in the following data:

Reaction with S.C. of	No. of runs	Mean velocity	Sample s.d.
1.5 moles/litre	15	7.5	1.5
2.0 moles/litre	12	8.8	1.2

Is the claim tenable at 0.01 L.O.S. assuming that the populations are normally distributed with equal variances.

Hint: $H_0 : \mu_1 - \mu_2 = \delta = 0.5$, $H_1 : \mu_1 - \mu_2 > \delta = 0.5$, $t_{0.01,25} = 2.485$

$$S_p^2 = \frac{14(1.5)^2 + 11(1.2)^2}{15 + 2 - 2} = 1.8936,$$

$$t = \frac{(8.8 - 7.5) - (0.5)}{1.376\sqrt{\frac{1}{15} + \frac{1}{12}}} = 1.50$$

Reject N.H. since $t = 1.50 < 2.48$ (Right O.T.T.)

Ans. No, claim not tenable.

3. A study is conducted to determine whether the wear of material A exceeds that of B by more than 2 units. If test of 12 pieces of material A yielded a mean wear of 85 units and s.d. of 4 while test of 10 pieces of material B yielded a mean of 81 and s.d. 5, what conclusion can be drawn at 0.05 L.O.S. Assume that populations are approximately normally distributed with equal variances.

Hint: $H_0 : \mu_1 - \mu_2 = 2$, $H_1 : \mu_1 - \mu_2 > 2$, $t_{0.05}$ with 20 dof is 1.725

$$S_p^2 = \frac{11(16) + 9(25)}{12 + 10 - 2} = 20.052,$$

$$t = \frac{(85 - 81) - 2}{4.478\sqrt{\frac{1}{12} + \frac{1}{10}}} = 1.04$$

Accept H_0 since $t = 1.04 < 1.725 = t_{0.05,20}$.

Ans. Wear of A does not exceed that of B by 2 units.

4. To determine whether vegetarian and non-vegetarian diets effects significantly on increase in weight a study was conducted yielding the following data of gain in weight.

Vegetarian: 34, 24, 14, 32, 25, 32, 30, 24, 30, 31, 35, 25

Non-vegetarian: 22, 10, 47, 31, 44, 34, 22, 40, 30, 32, 35, 18, 21, 35, 29

Can we claim that the two diets differ pertaining to weight gain, assuming that samples are drawn from normal populations with same variance.

Hint: $\overline{X}_V = \frac{336}{12} = 28$, $n_V = 12$, $\overline{X}_{NV} = \frac{450}{15} = 30$, $n_{NV} = 15$

$$S^2 = 71.6, t = \frac{\overline{X}_V - \overline{X}_{NV}}{\sqrt{S^2\left(\frac{1}{n_V} + \frac{1}{n_{NV}}\right)}} = \frac{28-30}{\sqrt{71.6\left(\frac{1}{12} + \frac{1}{15}\right)}}$$

$$= -0.609$$

Accept N.H.: $\mu_V = \mu_{NV}$ against A.H.: $\mu_V \neq \mu_{NV}$ at 0.05 L.O.S.

Since $-0.609 > -2.06$ and $t_{0.05}$ at $12 + 15 - 2 = 25$ dof is 2.06

Ans. Vegetarian and non-vegetarian diets do not differ significantly as far as their effect on increase in weight.

5. In a study on the influence of habitation, the intelligent quotients (IQs) of 16 students from urban area was found to have a mean of 107 and s.d. of 10, while the IQs of 14 students from a rural area showed a mean of 112 and s.d. 8. Determine whether the IQs differ significantly at (a) 0.01 (b) 0.05 levels.

Hint:

$$S_p^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} = \frac{16(10)^2 + 14(8)^2}{16 + 14 - 2} = 89.1136$$

$$t = \frac{112 - 107}{9.44\sqrt{\frac{1}{16} + \frac{1}{14}}} = 1.45$$

a. Accept $H_0 : \mu_U = \mu_R$, Acceptable region $(-2.76, 2.76)$

b. Accept H_0, since 1.45 lies in the acceptable region $(-2.05, 2.05)$

Ans. No, habitation has influence on IQs.

6. To test the claim that application of pesticide increases production of rice, a study was conducted as follows. Out of 24 plots of equal areas, equal soil conditions, same exposure to sunlight, 12 plots were treated with pesticides while the remaining 12 plots were left untreated. The mean increase in rice production was 4.8 kgs and s.d. of 0.40 kgs for treated plots and 5.1 kgs mean and 0.36 s.d. for untreated plots. Is there significant increase in rice production due to pesticide application at (a) 0.01 (b) 0.05 L.O.S.

Hint:

$$S^2 = \frac{12(0.40)^2 + 12(0.36)^2}{12 + 12 - 2} = 0.157609$$

$$t = \frac{5.1 - 4.8}{0.397\sqrt{\frac{1}{12} + \frac{1}{12}}} = 1.85$$

$t_{0.99}$ at 22 dof, right O.T.T. is 2.51

$t_{0.95}$ at 22 dof, right O.T.T. is 1.72

Ans. (a) Accept $H_0 : \mu_1 = \mu_2$ (b) Reject H_0, significant.

29.9 PAIRED-SAMPLE *t*-TEST

Paired observations arise in a very special experimental situation where each homogeneous experimental unit receives both population conditions. As a result, each experimental unit has a pair of observations, one for each population. Thus the paired observations are on the same unit or matching units.

Examples: To test the effectiveness of "insulin" some 10 diabetic patients sugar level in blood is measured "before" and "after" the insulin is injected. Here the individual diabetic patient is the experimental unit and the two populations are blood sugar level "before" and "after" the insulin is injected.

So for each observation is one sample, there is a corresponding observation in the other sample pertaining to the same character. Thus the two samples are *not* independent. Paired t-test is applied for n paired observations (which are dependent) by taking the (signed) differences $d_1, d_2, \ldots, d_n$ of the paired data. To test whether the differences d form a random sample from a population with $\mu_D = d_O$ use large sample test (on Page 766) or one-sample t-test (on Page 773) when sample is small (the one sample t-test in this case is known as the paired-sample t-test). The test statistic is

$$\frac{\bar{d} - \mu_d}{S_d/\sqrt{n}}$$

with $\nu = n - 1$ dof and $\bar{d}$ and S_d^2 are the mean and variance of the differences $d_1, d_2, \ldots, d_n$.

Worked Out Examples

Examples: In a study of usefulness of yoga in weight reduction, a random sample of 16 persons undergoing yoga were examined of their weight before (without) and after (with) yoga with the following results:

Weight before	209	178	169	212	180	192	158	180	170	153	183	165	201	179	243	144
Weight after	196	171	170	207	177	190	159	180	164	152	179	162	199	173	231	140

Test whether yoga is useful in weight reduction at 0.01 level of significance.

Solution: Let μ be the mean of population of differences,

1. N.H.: $\mu = 0$ i.e., not useful.
2. A.H.: $\mu > 0$ i.e., yoga is useful in weight reduction.
3. L.O.S.: $\alpha = 0.01$.
4. Critical region: Right one tailed test. Reject N.H. if $t > t_{0.01}$ with $16 - 1 = 15$ degrees of freedom. From table $t_{0.01} = 2.602$.
5. Calculation: differences d_i's are

$13, 7, -1, 5, 3, 2, -1, 0, 6, 1, 4, 3, 2, 6, 12, 4$

$\bar{x} = \text{mean of differences of sampled data} = \frac{66}{16} = 4.125$

$s^2 = \frac{247.73}{15} = 16.516, s = 4.064$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{4.125 - 0}{4.064/\sqrt{16}} = 4.06$$

6. Decision: Reject N.H. since $t = 4.06 > 2.602 = t_{0.01}$ i.e., yoga is useful in weight reduction.

Exercise

1. Use paired sample test at 0.05 level of significance to test from the following data whether

Weight in gms	Scale I	11.23	14.36	8.33	10.50	23.42	9.15	13.47	6.47	12.40	19.38
	Scale II	11.27	14.41	8.35	10.52	23.41	9.17	13.52	6.46	12.45	19.35

the differences of the means of the weights obtained by two different scales (weighting machines) is significant.

Hint: $\bar{x} = -\frac{0.2}{10} = -0.02$, $s = 0.028674$, $n = 10$, $t = \frac{-0.02-0}{0.028/\sqrt{10}} = -2.21$, $t_\alpha = t_{0.05}$ with 9 dof is 1.833.

Ans. No significant difference in the two scales.

2. The average weekly losses of man-hours due to strikes in an institute before and after a disciplinary program was implemented are as follows:

Before	45	73	46	124	33	57	83	34	26	17
After	36	60	44	119	35	51	77	29	24	11

Is there reason to believe that the disciplinary program is effective at 0.05 level of significance?

Hint: $\bar{x} = 5.2, s = 4.08, n = 10, t = 4.03$, $t_{0.05}$ with 9 dof is 1.833.

Ans. Yes, program is effective.

3. The pulsality index (P.I.) of 11 patients before and after contracting a disease are given below. Test at 0.05 level of significance whether there is a significant increase of the mean of P.I. values.

Before	0.4	0.45	0.44	0.54	0.48	0.62	0.48	0.60	0.45	0.46	0.35
After	0.5	0.60	0.57	0.65	0.63	0.78	0.63	0.80	0.69	0.62	0.68

Hint: $\bar{x} = \frac{188}{11} = 0.171, s = 0.065, n = 11$, $t = 8.72$, $t_{0.05}$ with 10 dof is 1.812.

Ans. Yes, there is significant increase in P.I. values.

4. The following data gives the amount of androgen present in blood of 15 deers before and 30 minutes after a certain drug in injected to them.

Before	2.76	5.18	2.68	3.05	4.10	7.05	6.60	4.79	7.39	7.30	11.78	3.9	26	67.48	17.04
After	7.02	3.1	5.44	3.99	5.21	10.26	13.91	18.53	7.91	4.85	11.1	3.74	94.03	94.03	41.7

Test at 0.05 L.O.S. whether there is significant change in the concentration levels of androgen in blood.

Hint: $\bar{x} = 9.848, s = 18.474, t = 2.06$, critical region: $t < -2.145$ and $t > 2.145$ (with 14 dof).

Ans. Yes, there is difference in mean circulating levels of androgen in the blood of deer.

5. The blood pressure (B.P.) of 5 women before and after intake of a certain drug are given below:

Before	110	120	125	132	125
After	120	118	125	136	121

Test at 0.01 L.O.S. whether there is significant change in B.P.

Hint: $\bar{x} = \frac{10}{5} = 2, s = 5.477, t = 0.817$, $t_{0.01}$ with 4 dof is 3.747.

Ans. No significant change in B.P.

6. Marks obtained in mathematics by 11 students before and after intensive coaching are given below:

Before	24	17	18	20	19	23	16	18	21	20	19
After	24	20	22	20	17	24	20	20	18	19	22

Test at 0.05 L.O.S. whether the intensive coaching is useful?

Hint: $\bar{x} = -1, s = 2.296, t = -1.38$, $t_{0.05}$ with 10 dof is 1.812.

Ans. Not useful.

29.10 TEST OF HYPOTHESIS: ONE PROPORTION: SMALL SAMPLES

The quality control engineer wants to know the proportion of defective products (items) in his industry, a university the percentage of first classes and a electronic component manufacturer the probability that a component works for a certain period and so on. In these cases, the observations on various items or objects are classified into two mutually exclusive (dichotomus) classes (forming a binomial population).

Let X, the number of successes be a binomial random variable. Let p be the parameter of the binomial distribution. Then the test of hypothesis concerning one proportion for *small samples* is as follows:

1. N.H.: $H_0 : p = p_0$ i.e., a proportion (percentage or probability) equals some given constant p_0.
2. A.H.: $H_1 : p \neq p_0$ (or $p < p_0$ or $p > p_0$).
3. L.O.S.: α
4. Test statistic: Binomial variable X with $p = p_0$.
5. Computation: Let x be the number of successes in a sample of size n.

   Compute p-value:

   a. A.H.: $p \neq p_0$:
   $P = 2P(X \leq x \text{ when } p = p_0)$ if $x < np_0$.
   $P = 2P(X \geq x \text{ when } p = p_0)$ if $x > np_0$

   b. A.H.: $p < p_0$: $P = P(X \leq x \text{ when } p = p_0)$

   c. A.H.: $p > p_0$: $P = P(X \geq x \text{ when } p = p_0)$
6. Decision: Reject N.H.: H_0 if $P \leq \alpha$.

$\chi^2 = 1.53 < 9.488$ where $\chi^2_{0.05}$ with $\nu = (3-1)(3-1) = 4$ dof is 9.488.

2. To determine the effectiveness of drugs against "aids", three types of medicines, allopathic, homeopathic and ayurvedic were tested on 50 persons with the following results.

Effectiveness	Drug			
	Allopathy	Homeo-pathy	Ayurved	Total
No relief	11	13	9	33
Some relief	32	28	27	87
Total relief	7	9	14	30
Total	50	50	50	150

Hint: $e_{11} = 11, e_{12} = 11, e_{13} = 11, e_{21} = 29, e_{22} = 29, e_{23} = 29, e_{31} = 10, e_{32} = 10, e_{33} = 10$.

Ans. Accept N.H. since $\chi^2 = 3.8100313 < 9.488 = \chi^2_{0.05}$ with $\nu = (3-1)(3-1) = 4$ dof i.e., three drugs are equally effective or the drugs are homogeneous.

3. The following table shows the opinions of voters before and after a presidential election.

	Before	After	Total
For ruling party	79	91	170
For opposition	84	66	150
Undecided	37	43	80
Total	200	200	400

Test the claim at 0.05 L.O.S. whether there has been a change of opinion of voters.

Hint: $e_{11} = e_{12} = 85, e_{21} = e_{22} = 75, e_{31} = e_{32} = 40$

Ans. No, there is no change in opinion of voters since $\chi^2 = 3.46 < 5.991 = \chi^2_{0.05}$ at $(3-1)(2-1) = 2$ dof (i.e., accept N.H.)

29.15 GOODNESS OF FIT TEST

To determine if a population follows a specified known theoretical distribution such as normal distribution, binomial distribution or Poisson distribution, the χ^2 (chi-square) test is used to assertion how closely the actual distribution approximate the assumed theoretical distributions. This test, which is based on how good a fit is there between the observed frequencies (o_i from the sample) and the expected frequencies (e_i from the theoretical distribution) is known as **"goodness-of-fit-test"**. This test judges whether the sample is drawn from a certain hypothetical distribution i.e., whether the observed frequencies follow a postulated distribution or not.

The statistic χ^2 is a measure of the discrepancy existing between the observed and expected (or theoretical) frequencies.

$$\left.\begin{array}{l}\text{Statistic for test of}\\ \text{"goodness of fit"}\end{array}\right\}\ \chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \qquad (1)$$

Here O_i and e_i are the observed and expected frequencies of the ith cell (or class interval), such that $\sum O_i = \sum e_i = N$ = Total frequency.

k is the number of cells or class intervals, in the given frequency distribution.

Here χ^2 is a random variable which is very closely approximated with ν degrees of freedom.

Degrees of Freedom (dof) for χ^2-Distribution

Let k be the number of terms in the formula (1) for χ^2. Then the dof for χ^2 is:

a. $\nu = k - 1$ if e_i can be calculated *without* having to estimate population parameters from sample statistics.

b. $\nu = k - 1 - m$ if e_i can be calculated *only by* estimating m number of population parameters from sample statistics.

Examples:

i. B.D.: p is the parameter, $m = 1$, $\nu = k - 1 - m = k - 1 - 1 = k - 2$

ii. P.D.: λ is the parameter, $m = 1$, $\nu = k - 2$

iii. N.D.: μ, σ are two parameters, $m = 2$, $\nu = k - 1 - 2 = k - 3$

Test for Goodness-of-Fit

1. N.H.: good-fit exists between the theoretical distribution and given data (of observed frequencies).
2. N.H.: no good fit.
3. L.O.S.: α (prescribed).
4. Critical region: Reject N.H. if $\chi^2 > \chi^2_\alpha$ with ν dof, i.e., theoretical distribution is a poor fit.
5. Compute χ^2 from (1).
6. Decision: Accept N.H. if $\chi^2 < \chi^2_\alpha$. i.e., the theoretical distribution is a good fit to the data.

Solution: Mean of the given frequency distribution is

$$\lambda = \frac{0\times52+1\times151+2\times130+3\times102+4\times45+5\times12+6\times5+7\times1+8\times2}{52+151+130+102+45+12+5+1+2}$$

$$\lambda = \frac{1010}{500} = 2.02$$

The Poisson distribution that fits to the data with this parameter $\lambda = 2.02$ is

$$P(x,\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-2.02}(2.02)^x}{x!}\ x = 0, 1, 2, \ldots, 8$$

x:	0	1	2	3	4	5	6	7	8
$P(x)$	0.1326	0.26796	0.2706	0.1822	0.092	0.037	0.0125	0.0036	0.00091
Expected frequency $=500\times P(x)$	66.32 $\approx$ 66	133.94 $\approx$ 134	135.32 $\approx$ 135	91.116 $\approx$ 91	46.01 $\approx$ 46	18.5896 $\approx$ 19	6.25 $\approx$ 6	1.806 $\approx$ 2	0.456 $\approx$ 0

Note 1: If $\chi^2 = 0$ then O_i and e_i agree exactly.

Note 2: when $\chi^2 > 0$.

a. χ^2 small: O_i are close to e_i, indicating "good" fit.

b. χ^2 large: O_i differ considerably from e_i indicating "poor" fit.

Conditions for Validity of χ^2-Test

1. Sample size n should be large (i.e., $n \geq 50$).
2. If individual frequencies O_i (or e_i) are small say $O_i < 10$ then combine neighbouring frequencies so that combined frequency O_i (or e_i) is ≥ 10.
3. The number of classes k should be neither too small nor too large. In general $4 \leq k \leq 16$.

WORKED OUT EXAMPLES

Example 1: Test for goodness of fit of a poisson distribution at 0.05 L.O.S. to the following frequency distribution:

Number of patients arriving/hour: (x)	0	1	2	3	4	5	6	7	8
Frequency	52	151	130	102	45	12	5	1	2

1. N.H.: H_0: R.V. x has Poisson distribution with $\lambda = 2.02$
2. A.H.: H_1: R.V. does *not* have P.D.
3. L.O.S. $\alpha = 0.05$
4. Critical region: Reject N.H. if $\chi^2 > 14.067$ where $\chi^2_{0.05}$ with $k - 1 - m = 9 - 1 - 1 = 9 - 2 = 7$ dof is 14.067 (since only one parameter λ is needed to calculate expected frequencies)
5. Calculation:

$$\chi^2 = \sum_i \frac{(O_i - e_i)^2}{e_i}$$

$$= \frac{(52-66)^2}{66} + \frac{(151-134)^2}{134} + \frac{(130-135)^2}{135} + \frac{(102-91)^2}{91} + \frac{(45-46)^2}{46} + \frac{(12-9)^2}{19} + \frac{[(5+1+2)-(6+2+0)]^2}{8}$$

$$\chi^2 = 9.2419$$

6. Decision: Accept N.H. i.e., Poisson distribution with $\lambda = 2.02$ is a good fit to the given frequency distribution since $\chi^2 = 9.2419 < 14.067 = \chi^2_{0.05}$ with 7 dof.

Example 2: Use 0.05 L.O.S. to test that the following given data may be treated as a random sample from a normal population

Class	Frequency
5.0–8.9	3
9.0–12.9	10
13.0–16.9	14
17.0–20.9	25
21.0–24.9	17
25.0–28.9	9
29.0–32.9	2
Total	80

Solution: A.M. $= \bar{x} = 18.85$, $\sigma =$ s.d. $= 5.5$.
The normal distribution with these two parameters is given below:

X Class Mark	$z = \frac{X-\bar{X}}{\sigma} = \frac{X-18.85}{5.5}$	Area under N.C.	Probability of a class	Expected frequency = probx 80
4.95	−2.52	.4941	0.03	2.4
8.95	−1.8	.4641	.1064	8.512
12.95	−1.072	.3577	.221	17.67
16.95	−0.345	.1368	.285	22.8
20.95	0.3818	.1480	.216	17.304
24.95	1.109	.3643	.105	8.4
28.95	1.836	.4693		
32.95	2.563	.4948	.0255	2.04

$$\chi^2 = \frac{(3-2.4)^2}{2.4} + \frac{(10-8.5)^2}{8.5} + \frac{(14-17.7)^2}{17.7} + \frac{(25-22.8)^2}{22.8} + \frac{(17-17.3)^2}{17.3} + \frac{(9-8.4)^2}{8.4} + \frac{(2-2.04)^2}{2.04} = 1.4524$$

Cannot reject H_0 i.e., accept H_0 since $\chi^2 = 1.4524 < 9.488 = \chi^2_{0.05}$ with $k - 1 - 2 = 7 - 3 = 4$ dof (since 2 parameters $\bar{X}, \sigma$ are needed).

EXERCISE

1. Test for goodness of fit of a Poisson distribution at 0.01 L.O.S. to the following observed data of e-mails received:

No. of e-mails	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Observed frequency	3	15	47	76	68	74	46	39	15	9	5	2	0	1

Hint: $\lambda = \frac{1814}{400} = 4.535 \approx 4.6$
$e_0 = 4.0,\ e_1 = 18.4,\ e_2 = 42.8,\ e_3 = 65.2,\ e_4 = 74.8,\ e_5 = 69.2,\ e_6 = 52.8,\ e_7 = 34.8,\ e_8 = 20,\ e_9 = 10,\ e_{10} = 4.8,\ e_{11} = 2.0,\ e_{12} = 0.8,\ e_{13} = 0.4$

$$\chi^2 = \frac{(18-22.4)^2}{22.4} + \frac{(47-42.8)^2}{42.8} + \cdots + \frac{(8-8.0)^2}{8.0} = 6.749.$$

Ans. P.D. with $\lambda = 4.6$ provides a good fit since $\chi^2 = 6.749 < 16.919 = \chi^2_{0.01}$ with 9 dof.

2. Test for goodness of fit of a uniform distribution to the following data obtained when a die is tossed 120 times.

Face	1	2	3	4	5	6
Observed	20	22	17	18	19	24
Expected	20	20	20	20	20	20

Use 0.05 L.O.S.

Hint: $\chi^2 = \frac{(20-20)^2}{20} + \frac{(22-20)^2}{20} + \frac{(17-20)^2}{20} + \frac{(18-20)^2}{20} + \frac{(19-20)^2}{20} + \frac{(24-20)^2}{20} = 1.7.$

Ans. Uniform distribution is a good fit to the data. Accept N.H. that die is balanced since $\chi^2 = 1.7 < 11.070 = \chi^2_{0.05}$ with $6 - 1 = 5$ dof.

3. Test for goodness of fit of normal distribution to the following frequency table:

Class	1.45–1.95	1.95–2.45	2.45–2.95	2.95–3.45	3.45–3.95	3.95–4.45	4.45–4.95
Frequency O_i	2	1	4	15	10	5	3

Hint: $e_1 = 0.5 + 2.1 + 5.9 = 8.5$, $e_2 = 10.3$, $e_4 = 7.0 + 3.5 = 10.5$, $e_3 = 10.7$

$$\chi^2 = \frac{(7-8.5)^2}{8.5} + \frac{(15-10.3)^2}{10.3} + \frac{(10-10.7)^2}{10.7} + \frac{(8-10.5)^2}{10.5} = 3.05.$$

Ans. Normal distribution with $\bar{x} = 3.5$ and $\sigma = 0.7$ is a good fit since $\chi^2 = 3.05 < 7.815 = \chi^2_{0.05}$ for 3 dof (Here first three classes are clubbed to have 7 frequency and last two classes to have 8 frequency, so $k = 4$ classes $- 1 = 3$.

4. Test for goodness of fit of a binomial distribution to the data given below:

X_i	0	1	2	3	4	5	6
O_i	5	18	28	12	7	6	4

Hint: $\mu = 2.4 = 6p$, $p = 2.4$
e_i: 4 15 25 22 11 3 0
Clubbing: e_i 19 25 22 14

Ans. Reject N.H. i.e., B.D. is not a good fit since $\chi^2 = 6.39 > 5.99 = \chi^2_{0.05}$ with $4 - 2 = 2$ dof.

5. Test for goodness of fit of a Poisson distribution to the following data:

X	0	1	2	3	4	5
O_i	275	138	75	7	4	1

Hint: $\lambda = \frac{330}{500} = 0.66$, $e_1 = 258$, $e_2 = 171$, $e_3 = 56$, $e_4 = 12.4$, $e_5 = 2.05$, $e_6 = 0.25$,

$$\chi^2 = \frac{(275-258)^2}{258} + \frac{(138-171)^2}{171} + \frac{(75-56)^2}{56} + \frac{(12-15)^2}{15} = 14.534.$$

Ans. P.D. is not a good fit since $\chi^2 = 14.534 > 5.991 = \chi^2_{0.05}$ with 2 dof.

6. Test for goodness of fit of normal distribution to the following data:

Class	0–100	100–250	250–500	500–750	750–1000	1000–1250	1250–1500
Frequency	7	9	19	12	8	5	4

Hint: $e_1 = 7.62$, $e_2 = 6.32$, $e_3 = 15.26$, $e_4 = 16.17$, $e_5 = 11.52$, $e_6 = 5.24$, $e_7 = 1.88$, club the last two classes.

Ans. Normal distribution with $\bar{X} = 541.8$ and $\sigma = 375.46$ is a good fit to the given data since $\chi^2 = 4.751 < 7.81 = \chi^2_{0.05}$ with $6 - 2 - 1 = 3$ dof.

29.16 ESTIMATION OF PROPORTIONS

Engineering problems dealing with proportions, percentages or probabilities are one and the same because a proportion multiplied by 100 is percentage and proportion with number of trials very large is interpreted as probability.

Examples: Percentage of engineering students getting first class, proportion of students gaining useful employment or the probability that a first class graduate is employed.

Sample proportion $= \frac{X}{n}$ where X is the number of times an event occurs in n trials.

Sample proportion is an unbiased estimator of the of true proportion, the binomial parameter p.

Large Sample Confidence Interval for p

When n is large, normal approximation is used for binomial distribution to construct confidence interval for p from the inequality

$$-z_{\alpha/2} < \frac{X - np}{\sqrt{np(1-p)}} < z_{\alpha/2}$$

by replacing $\frac{x}{n}$ by p. Thus the confidence interval for p, when n is large, is

$$\frac{x}{n} - z_{\alpha/2}\sqrt{\frac{\frac{x}{n}\left(1-\frac{x}{n}\right)}{n}} < p < \frac{x}{n} + z_{\alpha/2}\sqrt{\frac{\frac{x}{n}\left(1-\frac{x}{n}\right)}{n}}.$$

Maximum Error of Estimate

The magnitude of error committed in using sample proportion $\frac{X}{n}$ for true proportion p is given by the maximum error of estimate E, where

$$E = z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

Sample size:

a. when p is known, the sample size n is given by

$$n = p(1-p)\left[\frac{z_{\alpha/2}}{E}\right]^2$$

b. when p is unknown, then

$$n = \frac{1}{4}\left[\frac{z_{\alpha/2}}{E}\right]^2.$$