day

Testing of Hypothesis

A statistical hypothesis test is a method of making decisions using data whether from a controled experement or an observational study (not controled Hypothesis testing is sometimes called conformatory data analysis) In contrast to explenatory data analysis.

In frequency probability these decision are almost always made using null hypothesis test. (H₀)

A statistical hypothesis is some statement or an assertion about a population or equivallently about the probability distribution charactering a popⁿ which we want to varify on the basis of information available from a sample.

Simple and Composite Hypothesis

A Simple hypothesis is a hypothesis which Specifies the pop^n distribution Completly

eg: $H_0 : X \sim B(100, 1/2)$, n & P are specified

$H_0 : X \sim N(5, 200)$, μ & σ^2 are specified

Composite hypothesis

A composite hypothesis is the hypothesis does not specifying the pop^n distribution completly.

$X \sim B(100, P)$ and $H_1 : P > 0.5$

$X \sim N(\mu, \sigma^2)$ & $H_1 : \sigma^2$ unspecified

Null & Alternative Hypothesis

The Null hypothesis, H_0, represents a theory that has been put forward either because it is belived to be true or because it is to be used as basis for arguments, but has not been proved.

3 eg: In a clinical trial of a new drug, H_0 might be that the new drug is not better on ~~the~~ average, than the current drug. We would write H_0: their is no difference b/w the two drugs on average.

The alternative hypothesis, H_1, is a statement of what a statistical hypothesis test is setup to establish.

eg: In a clinical trial of a new drug H_1 might be that the new drug ~~is~~ have a different effect than the average, than the current drug. we would write H_1: the two drugs have a different effects on average.

Type I and Type II errors

Action taken based on a sample	State of nature	
	H_0 is true	H_0 is false
Reject H_0	Type I error	No error
Accept H_0	No error	Type II error

In a hypothesis test Type I error occurs when the H_0 is rejected when it is true. ie, H_0 is wrongly rejected.

$P(\text{type I error}) = \text{Significant level} = \alpha$

It is also be reffered to as an error of the 1st kind

In a hypothesis test a type II error occurs when H_0 is accept when it is false

$P(\text{Type II error}) = \beta$

It is also be reffered to as an error of the II^{nd} kind

For any given set of data Type I & Type II errors are inversly related, the smaller risk of one, the higher the risk of other.

Test Statistic

1-11 Tuesday

A test statistic is a quantity calculated from our sample of data its value is used to decide whether or not H_0 should be rejected in our hypothesis test.

Critical Region

In a statistical test procedure the range of variation of the test statistic is divided into two regions; the acceptness region & rejection region. the rejection region is also called critical region.

The region of rejection of H_0 when H_0 is true. is that region of outcome set where H_0 is rejected if the sample point false in that region. It is called critical region.

Significance level or α or
Size of the test

The probability of test statistic falling in the critical region when the hypothesis is true is called significand level or size of the test. this is the probability of rejecting the hypothesis when it is true or probability of the type I error.

Significance level =

$$\alpha = P(\text{rejecting } H_0 / H_0)$$

usually the significance level is choosen to 5% or 1%.

1

Power of the test

The power of statistical hypothesis test measures the test's ability to reject when it is actually false. ie, to make a correct decision.

Probability of rejecting the hypothesis when the null hypothesis is false is called power of the test. [H_1 is true]

$$\text{power} = P\left(\text{Rejecting } H_0 / H_1\right)$$

$$= 1 - P\left(\text{Accepting } H_0 / H_1\right)$$

$$= 1 - P(\text{type II error})$$

$$= \underline{\underline{1-\beta}}$$

(The range of power of a test is 0 to 1. ideally we want a test to have high power, close to 1

p value or probability value

The p value of a statistical hypothesis test is the probability of getting a value of test statistics as extreme as or more extreme than observed by chance alone, if the null hypothesis H_0 is true.

Remarks

In a quality control terminology α & β are termed as producer's risk and consumers risk respectivly.

One tailed & two tailed tests

A test of any statistical hypothesis where the alternative hypothesis is one tailed right tailed or left tailed is called one tailed test

eg: A test for testing mean of the popⁿ.
null hypothesis

$H_0: \mu = \mu_0$

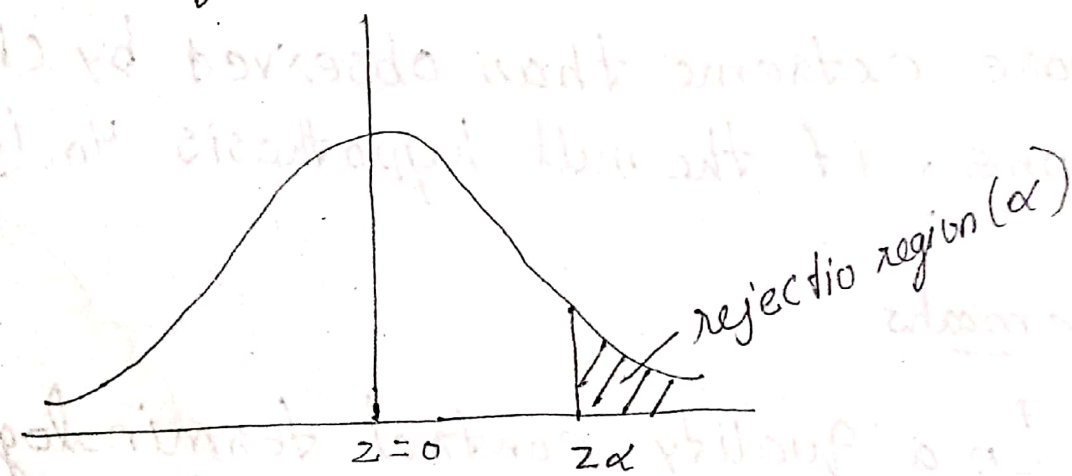
a

against alternative hypothesis

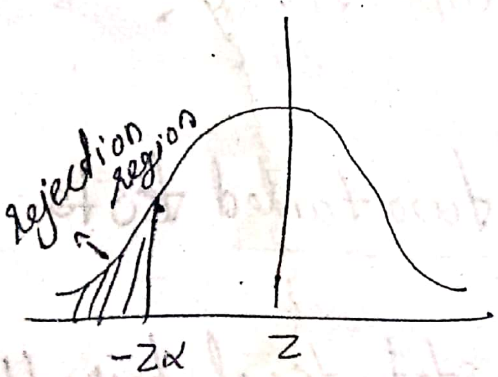
$H_1: \mu > \mu_0$ (right tailed)

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

For right tailed test $P(z > z\alpha) = \alpha$

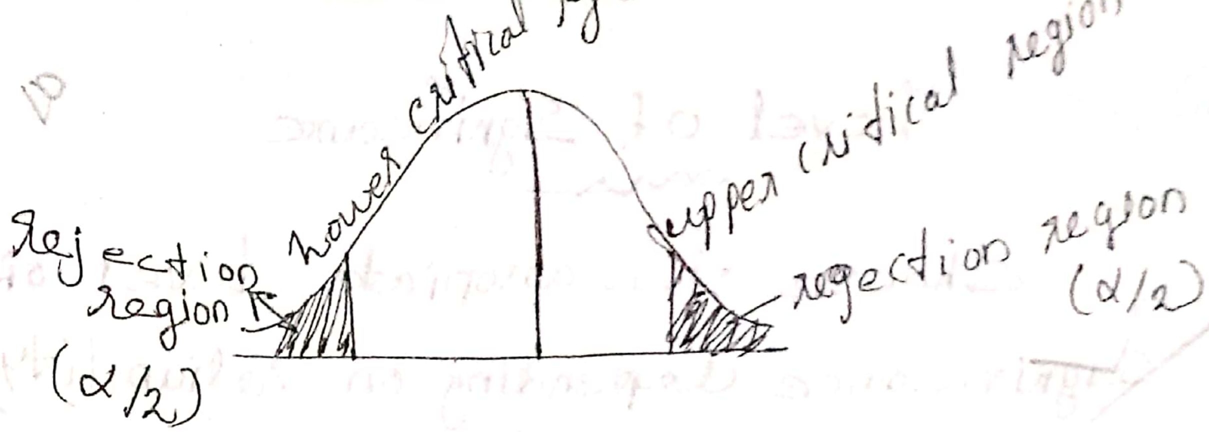


For left tailed test $P(z < -z\alpha) = \alpha$



A test of statistical hypothesis where the alternative hypothesis is two tailed such as $\mu = \mu_0$ against H_1 $\mu \neq \mu_0$ $(\mu > \mu_0$ & $\mu < \mu_0)$ is called two tailed test.

10



Critical Values of 'z'

Critical Value	Level of Significance		
	1%	5%	10%
Two tailed test	$\lvert z_\alpha \rvert = 2.58$	$\lvert z_\alpha \rvert = 1.96$	$\lvert z_\alpha \rvert = 1.645$
Right tailed test	$z_\alpha = 2.33$	$z_\alpha = 1.645$	$z_\alpha = 1.28$
Left tailed test	$z_\alpha = -2.33$	$z_\alpha = -1.645$	$z_\alpha = -1.28$

Procedure for testing of hypothesis

1) Null hypothesis.
Set up the Null hypothesis H_0.

11) Alternative hypothesis H_1 this will enable us to decide whether we have to use a single tailed test or two tailed test.

iii) Level of Significance

Choose the apropiate level of significance depending on reliability the esstimates and permisible risk.

iv) test statisstic or test criteria.

Compute the test statisstic $Z = \frac{t - E(t)}{SE(t)}$ under the null hypothesis

v Conclusion. We compare Z, the computed value of Z in step for with significant value (table value) $Z\alpha$ at the given level of significant α.

If $|Z| > Z\alpha$ ie, the calculated value is greater than table value. We say it is significant.

Testing of hypotheses

($n \geq 30$) → Large sample test
($n < 30$) → Small sample test.

Large sample test

Z-test

- Test of Significance for Single Mean

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

- Test of significance for difference of Mean

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (\sigma \text{ known})$$

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (\sigma \text{ unknown})$$

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

(σ unknown but not equal)

- Testi of singnificance for single proportion.

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

- Test of significance for difference of proportion

$$z = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\bar{p}_1 = \frac{x_1}{n_1} \quad \bar{p}_2 = \frac{x_2}{n_2}$$

$$p = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2}$$

$$q = 1 - p$$

χ^2-test
(non-Parametric test)

$$\chi^2 = \sum_{i=1}^{n} \frac{(f_i - e_i)^2}{e_i}$$

- Goodness of Fit test
- Independence of Attributes (contingency table)

a	b
c	d

$$\chi^2 = \frac{N(ad - bc)^2}{(a+c)\times(b+d)\times(a+b)\times(c+d)}$$

$$N = a+b+c+d$$

Small sample test.

t-test

- σ known → Follow z test
- σ unknown →

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}}$$

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

(Variance) F-test

$$F = \frac{\frac{n_1 s_1^2}{n_1 - 1}}{\frac{n_2 s_2^2}{n_2 - 1}}$$

14·8·3. Test of Significance for Single Mean. We have proved that if x_i, $(i = 1, 2, \ldots, n)$ is a random sample of size n from a normal population with mean μ and variance σ^2, then the sample mean is distributed normally with mean μ and variance σ^2/n, *i.e.*, $\bar{x} \sim N(\mu, \sigma^2/n)$. However, this result holds, *i.e.*, $\bar{x} \sim N(\mu, \sigma^2/n)$, even in random sampling from non-normal population provided the sample size n is large [*c.f.* Central Limit Theorem]. Thus for large samples, the *standard normal variate* corresponding to $\bar{x}$ is :

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

Under the *null hypothesis* H_0, that the sample has been drawn from a population with mean μ and variance σ^2, *i.e.*, there is no significant difference between the sample mean ($\bar{x}$) and population mean (μ), the test statistic (for large samples), is :

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1) \qquad \ldots(14\cdot9a)$$

Remarks 1. If the population *s.d.* σ is unknown then we use its estimate provided by the sample variance given by [See (14·8*b*)]. $\hat{\sigma}^2 = s^2 \quad \Rightarrow \quad \hat{\sigma} = s$ (for large samples).

2. *Confidence limits for* μ. 95% confidence interval for μ is given by :

$$|Z| \le 1\cdot96, \ \text{i.e.,} \ \left| \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \right| \le 1\cdot96 \ \Rightarrow \ \bar{x} - 1\cdot96\,(\sigma/\sqrt{n}) \le \mu \le \bar{x} + 1\cdot96\,(\sigma/\sqrt{n}) \qquad \ldots(14\cdot10)$$

and $\bar{x} \pm 1\cdot96\sigma/\sqrt{n}$ are known as 95% confidence limits for μ. Similarly, 99% confidence limits for μ are $\bar{x} \pm 2\cdot58\sigma/\sqrt{n}$ and 98% confidence limits for μ are $\bar{x} \pm 2\cdot33\sigma/\sqrt{n}$.

However, in sampling from a finite population of size N, the corresponding 95% and 99% confidence limits for μ are respectively

$$\bar{x} \pm 1\cdot96 \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \quad \text{and} \quad \bar{x} \pm 2\cdot58 \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \qquad \ldots(14\cdot10a)$$

3. The confidence limits for any parameter (P, μ, etc.) are also known as its *fiducial limits.*

Example 14·17. *A sample of 900 members has a mean 3·4 cms. and s.d. 261 cms. Is the sample from a large population of mean 3·25 cms. and s.d. 2·61 cms. ?*

If the population is normal and its mean is unknown, find the 95% and 98% fiducial limits of true mean.

Solution. *Null Hypothesis, (H_0) :* The sample has been drawn from the population with mean $\mu = 3{\cdot}25$ cms. and S.D. $\sigma = 2{\cdot}61$ cms.

Alternative Hypothesis, $H_1 : \mu \neq 3{\cdot}25$ (Two-tailed).

Test Statistic. Under H_0, the test statistic is : $Z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$, (Since n is large.)

Here, we are given : $\bar{x} = 3{\cdot}4$ cms., $n = 900$ cms., $\mu = 3{\cdot}25$ cms. and $\sigma = 2{\cdot}61$ cms.

$$\therefore \qquad Z = \frac{3{\cdot}40 - 3{\cdot}25}{2{\cdot}61/\sqrt{900}} = \frac{0{\cdot}15 \times 30}{2{\cdot}61} = 1{\cdot}73$$

Since $|Z| < 1{\cdot}96$, we conclude that the data don't provide us any evidence against the null hypothesis (H_0) which may, therefore, be accepted at 5% level of significance.

95% *fiducial limits* for the population mean μ are :

$$\bar{x} \pm 1{\cdot}96\,(\sigma/\sqrt{n}\,) = 3{\cdot}40 \pm 1{\cdot}96\,(2{\cdot}61/\sqrt{900}) = 3{\cdot}40 \pm 0{\cdot}1705,\ i.e.,\ 3{\cdot}5705 \text{ and } 3{\cdot}2295$$

98% *fiducial limits* for μ are given by :

$$\bar{x} \pm 2{\cdot}33\frac{\sigma}{\sqrt{n}} = 3{\cdot}40 \pm 2{\cdot}33 \times \frac{2{\cdot}61}{30} = 3{\cdot}40 \pm 0{\cdot}2027,\ i.e.,\ 3{\cdot}6027 \text{ and } 3{\cdot}1973$$

Remark. 2·33 is the value z_1 of Z from standard normal probability integrals, such that

$$P(|Z| > z_1) = 0{\cdot}98 \quad \Rightarrow \quad P(Z > z_1) = 0{\cdot}49$$

14·8·4. Test of Significance for Difference of Means. Let $\bar{x}_1$ be the mean of a sample of size n_1 from a population with mean μ_1 and variance σ_1^2 and let $\bar{x}_2$ be the mean of an independent random sample of size n_2 from another population with mean μ_2 and variance σ_2^2. Then, since sample sizes are large,

$$\bar{x}_1 \sim N(\mu_1, \sigma_1^2/n_1) \quad \text{and} \quad \bar{x}_2 \sim N(\mu_2, \sigma_2^2/n_2)$$

Also $\bar{x}_1 - \bar{x}_2$, being the difference of two independent normal variates is also a normal variate. The value of Z (S.N.V.) corresponding to $\bar{x}_1 - \bar{x}_2$ is given by :

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - E(\bar{x}_1 - \bar{x}_2)}{S.E.\ (\bar{x}_1 - \bar{x}_2)} \sim N(0, 1)$$

Under the null hypothesis, $H_0 : \mu_1 = \mu_2$, *i.e.*, there is no significant difference between the sample means, we get

$$E(\bar{x}_1 - \bar{x}_2) = E(\bar{x}_1) - E(\bar{x}_2) = \mu_1 - \mu_2 = 0;\ V(\bar{x}_1 - \bar{x}_2) = V(\bar{x}_1) + V(\bar{x}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2},$$

the covariance term vanishes, since the sample means $\bar{x}_1$ and $\bar{x}_2$ are independent.

Thus under $H_0 : \mu_1 = \mu_2$, the test statistic becomes (for large samples),

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} \sim N(0, 1) \qquad \ldots(14{\cdot}11)$$

Remarks 1. If $\sigma_1^2 = \sigma_2^2 = \sigma^2$, *i.e.*, if the samples have been drawn from the populations with common S.D. σ, then under $H_0 : \mu_1 = \mu_2$,

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\{(1/n_1) + (1/n_2)\}}} \sim N(0, 1) \qquad \ldots[14{\cdot}11a]$$

2. If in (14·11*a*), σ is not known, then its estimate based on the sample variances is used. If the sample sizes are not sufficiently large, then an unbiased estimate of σ^2 is given by :

$$\hat{\sigma}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 + n_2 - 2)}, \text{ since}$$

$$E(\hat{\sigma}^2) = \frac{1}{n_1 + n_2 - 2}\{(n_1 - 1)\,E(S_1^2) + (n_2 - 1)\,E(S_2^2)\} = \frac{1}{n_1 + n_2 - 2}\left[(n_1 - 1)\,\sigma^2 + (n_2 - 1)\,\sigma^2\right] = \sigma^2$$

But since sample sizes are large, $S_1^2 \simeq s_1^2$, $S_2^2 \simeq s_2^2$, $n_1 - 1 \simeq n_1$, $n_2 - 1 \simeq n_2$. Therefore in practice, for large samples, the following estimate of σ^2 without any serious error is used :

$$\hat{\sigma}^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2} \qquad \ldots[14{\cdot}11b]$$

However, if sample sizes are small, then an exact sample test, *t*-test for difference of means (*c.f.* Chapter 16) is to be used.

Example 14·25. *The means of two single large samples of 1,000 and 2,000 members are 67·5 inches and 68·0 inches respectively. Can the samples be regarded as drawn from the same population of standard deviation 2·5 inches ? (Test at 5% level of significance.)*

Solution. In usual notations, we are given :

$$n_1 = 1{,}000,\ n_2 = 2{,}000,\ \bar{x}_1 = 67{\cdot}5 \text{ inches},\ \bar{x}_2 = 68{\cdot}0 \text{ inches}.$$

Null hypothesis, $H_0 : \mu_1 = \mu_2$ and $\sigma = 2{\cdot}5$ inches, *i.e.,* the samples have been drawn from the same population of standard deviation 2·5 inches.

Alternative Hypothesis, $H_1 : \mu_1 \neq \mu_2$ (Two-tailed)

Test Statistic. Under H_0, the test statistic is :

$$Z = \frac{\bar{x} - \bar{x}_2}{\sqrt{\left\{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right\}}} \sim N(0, 1) \qquad \text{(since samples are large)}$$

Now $$Z = \frac{67{\cdot}5 - 68{\cdot}0}{2{\cdot}5 \times \sqrt{\left(\frac{1}{1000} + \frac{1}{2000}\right)}} = \frac{-0{\cdot}5}{2{\cdot}5 \times 0{\cdot}0387} = -5{\cdot}1.$$

Conclusion. Since $|Z| > 3$, the value is highly significant and we reject the null hypothesis and conclude that samples are certainly not from the same population with standard deviation 2·5.

14·7·1. Test of Significance for Single Proportion. If X is the number of successes in n independent trials with constant probability P of success for each trial, then

$$E(X) = nP \quad \text{and} \quad V(X) = nPQ, \quad \text{where} \quad Q = 1 - P, \text{ is the probability of failure.}$$

It has been proved that for large n, the binomial distribution tends to normal distribution. Hence for large n, $X \sim N(nP, nPQ)$, *i.e.*,

$$Z = \frac{X - E(X)}{\sqrt{V(X)}} = \frac{X - nP}{\sqrt{nPQ}} \sim N(0, 1) \qquad \ldots (14\cdot4)$$

and we can apply the normal test.

Remarks 1. In a sample of size n, let X be the number of persons possessing the given attribute. Then

Observed proportion of successes $= X/n = p$, (say).

$$\therefore \qquad E(p) = E(X/n) = \frac{1}{n} E(X) = \frac{1}{n} nP = P \qquad \ldots (14\cdot4a)$$

Thus the sample proportion 'p' gives an unbiased estimate of the population proportion P.

Also $V(p) = V\left(\frac{X}{n}\right) = \frac{1}{n^2} V(X) = \frac{1}{n^2} nPQ = \frac{PQ}{n} \quad \Rightarrow \quad \text{S.E. }(p) = \sqrt{\frac{PQ}{n}}$ ... (14·4b)

Since X and consequently X/n is asymptotically normal for large n, the normal test for the proportion of successes becomes :

$$Z = \frac{p - E(p)}{\text{S. E. }(p)} = \frac{p - P}{\sqrt{PQ/n}} \sim N(0, 1) \quad \text{... (14·4c)}$$

2. If we have sampling from a finite population of size N, then

$$\text{S.E. }(p) = \sqrt{\left(\frac{N-n}{N-1}\right) \cdot \frac{PQ}{n}} \quad \text{... (14·4d)}$$

3. Since the probable limits for a normal variate X are $E(X) \pm 3\sqrt{V(X)}$, the probable limits for the observed proportion of successes are :

$$E(p) \pm 3\,\text{S.E .}(p), \quad i.e., \quad P \pm 3\sqrt{PQ/n}.$$

If P is not known then taking p (the sample proportion) as an estimate of P, the probable limits for the proportion in the population are : $p \pm 3\sqrt{pq/n}$. ... (14·4e)

However, the limits for P at level of significance α are given by : $p \pm z_\alpha \sqrt{pq/n}$, ... (14·4f)

where z_α is the significant value of Z at level of significance α.

In particular : 95% confidence limits for P are given by : $p \pm 1{\cdot}96\sqrt{pq/n}$, ... (14·4g)

and 99% confidence limits for P are given by : $p \pm 2{\cdot}58\sqrt{pq/n}$. ... (14·4h)

Example 14·1. *A die is thrown 9,000 times and a throw of 3 or 4 is observed 3,240 times. Show that the die cannot be regarded as an unbiased one and find the limits between which the probability of a throw of 3 or 4 lies.*

Solution. If the coming of 3 or 4 is called a success, then in usual notations :

$$n = 9{,}000 \; ; \; X = \text{Number of successes} = 3{,}240$$

Under the null hypothesis (H_0) *that the die is an unbiased one*, we get

$$P = \text{Probability of success} = \text{Probability of getting a 3 or 4} = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

Alternative hypothesis, $H_1 : p \neq \frac{1}{3}$, (*i.e.*, die is biased).

We have $\quad Z = \dfrac{X - nP}{\sqrt{nQP}} \sim N(0, 1)$, since n is large.

Now $\quad Z = \dfrac{3240 - 9000 \times (1/3)}{\sqrt{9000 \times (1/3) \times (2/3)}} = \dfrac{240}{\sqrt{2000}} = \dfrac{240}{44{\cdot}73} = 5{\cdot}36$

Since $|Z| > 3$, H_0 is rejected and we conclude that the die is almost certainly biased.

Since die is not unbiased, $P \neq \frac{1}{3}$. The probable limits for 'P' are given by :

$\hat{P} \pm 3\sqrt{\hat{P}\hat{Q}/n} = p \pm 3\sqrt{pq/n}$, where $\hat{P} = p = \dfrac{3{,}240}{9{,}000}\, 0{\cdot}36$ and $\hat{Q} = q = 1 - p = 0{\cdot}64$.

Probable limits for population proportion of successes may be taken as :

$$\hat{P} \pm 3\sqrt{\hat{P}\hat{Q}/n} = 0{\cdot}36 \pm 3\sqrt{\frac{0{\cdot}36 \times 0{\cdot}64}{9000}} = 0{\cdot}36 \pm 3 \times \frac{0{\cdot}6 \times 0{\cdot}8}{30\sqrt{10}} = 0{\cdot}345 \text{ and } 0{\cdot}375.$$

Hence the probability of getting 3 or 4 almost certainly lies between 0·345 and 0·375.

Example 14·5. *Twenty people were attacked by a disease and only 18 survived. Will you reject the hypothesis that the survival rate, if attacked by this disease, is 85% in favour of the hypothesis that it is more, at 5% level. (Use Large Sample Test.)*

Solution. In the usual notations, we are given : $n = 20$.

$$X = \text{Number of persons who survived after attack by a disease} = 18$$

$$p = \text{Proportion of persons survived in the sample} = \frac{18}{20} = 0{\cdot}90$$

Null Hypothesis, $H_0 : P = 0{\cdot}85$, *i.e.*, the proportion of persons survived after attack by a disease in the lot is 85%.

Alternative Hypothesis, $H_1 : P > 0{\cdot}85$ (Right-tailed alternative).

Test Statistic. Under H_0, the test statistic is :

$$Z = \frac{p - P}{\sqrt{PQ/n}} \sim N(0,1), \text{ (since sample is large).}$$

Now $$Z = \frac{0{\cdot}90 - 0{\cdot}85}{\sqrt{0{\cdot}85 \times 0{\cdot}15/20}} = \frac{0{\cdot}05}{0{\cdot}079} = 0{\cdot}633$$

Conclusion. Since the alternative hypothesis is one-sided (right-tailed), we shall apply right-tailed test for testing significance of Z. The significant value of Z at 5% level of significance for right-tailed test is $+ 1{\cdot}645$. Since computed value of $Z = 0{\cdot}633$ is less than $1{\cdot}645$, it is not significant and we may accept the null hypothesis at 5% level of significance.

14·7·2. Test of Significance for Difference of Proportions. Suppose we want to compare two distinct populations with respect to the prevalence of a certain attribute, say A, among their members. Let X_1, X_2 be the number of persons possessing the given attribute A in random samples of sizes n_1 and n_2 from the two populations respectively. Then sample proportions are given by : $p_1 = X_1/n_1$ and $p_2 = X_2/n_2$.

If P_1 and P_2 are population proportions, then

$$E(p_1) = P_1,\ E(p_2) = P_2 \qquad \text{[c.f. Equation (14·4a)]}$$

and
$$V(p_1) = \frac{P_1 Q_1}{n_1} \quad \text{and} \quad V(p_2) = \frac{P_2 Q_2}{n_2}$$

Since for large samples, p_1 and p_2 are independently and asymptotically normally distributed, $(p_1 - p_2)$ is also normally distributed. Then the standard variable corresponding to the difference $(p_1 - p_2)$ is given by :

$$Z = \frac{(p_1 - p_2) - E(p_1 - p_2)}{\sqrt{V(p_1 - p_2)}} \sim N(0, 1) \qquad \ldots (*)$$

Under the *null hypothesis*, $H_0 : P_1 = P_2$, *i.e.*, there is no significant difference between the sample proportions, we have

$$E(p_1 - p_2) = E(p_1) - E(p_2) = P_1 - P_2 = 0 \qquad \text{(Under } H_0)$$

Also $\quad V(p_1 - p_2) = V(p_1) + V(p_2)$,

the covariance term Cov (p_1, p_2) vanishes, since sample proportions are independent.

$$\Rightarrow \qquad V(p_1 - p_2) = \frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2} = PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right),$$

[$\because$ under $H_0 : P_1 = P_2 = P$ (say), and $Q_1 = Q_2 = Q$].

Hence, under $H_0 : P_1 = P_2$, the test statistic for the difference of proportions becomes :

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0, 1) \qquad \ldots (14·5)$$

In general, we do not have any information as to the proportion of A's in the populations from which the samples have been taken. Under $H_0 : P_1 = P_2 = P$ (say), an unbiased estimate of the population proportion P, based on both the samples is given by :

$$\hat{P} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{X_1 + X_2}{n_1 + n_2} \qquad \ldots (14·5a)$$

The estimate is unbiased, since

$$E(\hat{P}) = \frac{1}{n_1 + n_2} E\left[n_1 p_1 + n_2 p_2\right] = \frac{1}{n_1 + n_2}\left[n_1 E(p_1) + n_2 E(p_2)\right]$$

$$= \frac{1}{n_1 + n_2}(n_1 P_1 + n_2 P_2) = P \qquad [\because P_1 = P_2 = P, \text{ under } H_0]$$

Thus (14·5) along with (14·5a) gives the required test statistic.

Remarks 1. Suppose we want to test the significance of the difference between p_1 and p, where $p = \dfrac{(n_1 p_1 + n_2 p_2)}{(n_1 + n_2)}$ gives a pooled estimate of the population proportion on the basis of both the samples. We have $\quad V(p_1 - p) = V(p_1) + V(p) - 2\,\text{Cov}(p_1, p)$.

Since p_1 and p are not independent, Cov $(p_1, p) \neq 0$. $\qquad \ldots (*)$

$$= \frac{n_1}{n_1 + n_2} \cdot \frac{pq}{n_1} = \frac{pq}{n_1 + n_2} \cdot \quad [\because \; P = p \text{ and } Q = q]$$

$$\text{Var}(p) = \text{Var}\left[\frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}\right] = \frac{1}{(n_1 + n_2)^2} \text{Var}(n_1 p_1 + n_2 p_2)$$

$$= \frac{1}{(n_1 + n_2)^2} \left[n_1^2 \text{Var}(p_1) + n_2^2 \text{Var}(p_2)\right],$$

covariance term vanishes since p_1 and p_2 are independent.

$$\therefore \quad \text{Var}(p) = \frac{1}{(n_1 + n_2)^2} \left(n_1^2 \cdot \frac{pq}{n_1} + n_2^2 \cdot \frac{pq}{n_2}\right) = \frac{pq}{n_1 + n_2}$$

Substituting (*) and simplifying, we shall get

$$V(p_1 - p) = \frac{pq}{n_1} + \frac{pq}{n_1 + n_2} - 2 \frac{pq}{n_1 + n_2} = pq \left[\frac{n_2}{n_1 (n_1 + n_2)}\right]$$

Also $E(p_1 - p) = E(p_1) - E(p) = P - P = 0$

Thus, the test statistic in this case becomes :

$$Z = \frac{(p_1 - p) - E(p_1 - p)}{\text{S.E.}(p_1 - p)} = \frac{p_1 - p}{\sqrt{\left\{\frac{n_2}{(n_1 + n_2)} \cdot \frac{pq}{n_1}\right\}}} \sim N(0, 1) \qquad \text{... (14·5b)}$$

2. Suppose the population proportions P_1 and P_2 are given to be distinctly different, *i.e.*, $P_1 \neq P_2$ and we want to test if the difference $(P_1 - P_2)$ in population proportions is likely to be hidden in simple samples of sizes n_1 and n_2 from the two populations respectively.

We have seen that in the usual notations,

$$Z = \frac{(p_1 - p_2) - E(p_1 - p_2)}{\text{S.E.}(p_1 - p_2)} = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\left(\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}\right)}} \sim N(0, 1)$$

Here sample proportions are not given. If we set up the *null hypothesis* $H_0 : p_1 = p_2$, *i.e.*, the samples will not reveal the difference in the population proportions or in other words the difference in population proportions is likely to be hidden in sampling, the test statistic becomes :

$$| Z | = \frac{| P_1 - P_2 |}{\sqrt{\left(\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}\right)}} \sim N(0, 1) \qquad \text{... (14·5c)}$$

Example 14·7. *Random samples of 400 men and 600 women were asked whether they would like to have a flyover near their residence. 200 men and 325 women were in favour of the proposal. Test the hypothesis that proportions of men and women in favour of the proposal, are same against that they are not, at 5% level.*

Solution. *Null Hypothesis* $H_0 : P_1 = P_2 = P$ (say), *i.e.*, there is no significant difference between the opinions of men and women as far as proposal of flyover is concerned.

Alternative Hypothesis, $H_1 : P_1 \neq P_2$ (two-tailed).

We are given :

$$n_1 = 400, X_1 = \text{Number of men favouring the proposal} = 200$$

$$n_2 = 600, X_2 = \text{Number of women favouring the proposal} = 325$$

$\therefore\ p_1$ = Proportion of men favouring the proposal in the sample $= \dfrac{X_1}{n_1} = \dfrac{200}{400} = 0{\cdot}5$

p_2 = Proportion of women favouring the proposal in the sample $= \dfrac{X_2}{n_2} = \dfrac{325}{600} = 0{\cdot}541$

Test Statistic. Since samples are large, the test statistic under the Null Hypothesis, H_0 is :

$$Z = \frac{p_1 - p_2}{\sqrt{\hat{P}\hat{Q}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} \sim N(0, 1), \text{ where}$$

$$\hat{P} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{200 + 325}{400 + 600} = 0{\cdot}525 \quad \Rightarrow \quad \hat{Q} = 1 - \hat{P} = 1 - 0{\cdot}525 = 0{\cdot}475$$

$$\therefore\ Z = \frac{0{\cdot}500 - 0{\cdot}541}{\sqrt{0{\cdot}525 \times 0{\cdot}475 \times \left(\dfrac{1}{400} + \dfrac{1}{600}\right)}} = \frac{-0{\cdot}041}{\sqrt{0{\cdot}001039}} = \frac{-0{\cdot}041}{0{\cdot}0323} = -1{\cdot}269$$

Conclusion. Since $|Z| = 1{\cdot}269$ which is less than $1{\cdot}96$, it is not significant at 5% level of significance. Hence H_0 may be accepted at 5% level of significance and we may conclude that men and women do not differ significantly as regards proposal of flyover is concerned.

Critical value (z_α)	Level of significance (α) 1%	5%	10%
Two-tailed test	$\lvert Z_\alpha \rvert = 2{\cdot}58$	$\lvert Z_\alpha \rvert = 1{\cdot}96$	$\lvert Z_\alpha \rvert = 1{\cdot}645$
Right-tailed test	$Z_\alpha = 2{\cdot}33$	$Z_\alpha = 1{\cdot}645$	$Z_\alpha = 1{\cdot}28$
Left-tailed test	$Z_\alpha = -2{\cdot}33$	$Z_\alpha = -1{\cdot}645$	$Z_\alpha = -1{\cdot}28$

Remark

15·6·1. Inferences About a Population Variance. Suppose we want to test if a random sample x_i, $(i = 1, 2, \dots, n)$ has been drawn from a normal population with a specified variance $\sigma^2 = \sigma_0^2$ (say).

Under the null hypothesis that the population variance is $\sigma^2 = \sigma_0^2$, the statistic

$$\chi^2 = \sum_{i=1}^{n} \left[\frac{(x_i - \bar{x})^2}{\sigma_0^2}\right] = \frac{1}{\sigma_0^2}\left[\sum_{i=1}^{n} x_i^2 - \frac{(\Sigma x_i)^2}{n}\right] = \frac{ns^2}{\sigma_0^2} \qquad \dots(15\cdot14)$$

follows chi-square distribution with $(n-1)$ *d.f.*

By comparing the calculated value with the tabulated value of χ^2 for $(n-1)$ *d.f.* at certain level of significance (usually 5%), we may retain or reject the null hypothesis.

Remarks 1. The above test (15·14) can be applied only if the population from which the sample is drawn is normal.

2. If the sample size n is large (>30), then we can use Fisher's approximation

$$\sqrt{2\chi^2} \sim N(\sqrt{2n-1},\ 1), \quad i.e., Z = \sqrt{2\chi^2} - \sqrt{2n-1} \sim N(0, 1) \qquad \dots(15\cdot14a)$$

and apply Normal Test.

Example 15·9. *It is believed that the precision (as measured by the variance) of an instrument is no more than 0·16. Write down the null and alternative hypothesis for testing this belief. Carry out the test at 1% level given 11 measurements of the same subject on the instrument :*

2·5, 2·3, 2·4, 2·3, 2·5, 2·7, 2·5, 2·6, 2·6, 2·7, 2·5.

Solution.

COMPUTATION OF SAMPLE VARIANCE

X	$X - \bar{X}$	$(X - \bar{X})^2$
2·5	−0·01	0·0001
2·3	−0·21	0·0441
2·4	−0·11	0·0121
2·3	−0·21	0·0441
2·5	−0·01	0·0001
2·7	+0·19	0·0361
2·5	−0·01	0·0001
2·6	+0·09	0·0081
2·6	+0·09	0·0081
2·7	+0·19	0·0361
2·5	−0·01	0·0001
$\bar{X} = \frac{27\cdot6}{11} = 2\cdot51$		$\Sigma(X - \bar{X})^2 = 0\cdot1891$

Null Hypothesis, $H_0 : \sigma^2 = 0\cdot16$

Alternative Hypothesis, $H_1 : \sigma^2 > 0\cdot16$

Under the null hypothesis, $H_0 : \sigma^2 = 0\cdot16$, the test statistic is :

$$\chi^2 = \frac{ns^2}{\sigma^2} = \frac{\Sigma(X-\bar{X})^2}{\sigma^2} = \frac{0.1891}{0.16} = 1.182,$$

which follows χ^2-distribution with *d.f.* $n-1 = (11-1) = 10$.

Since the calculated value of χ^2 is less than the tabulated value 23·2 of χ^2 for 10 *d.f.* at 1% level of significance, it is not significant. Hence H_0 may be accepted and we conclude that the data are consistent with the hypothesis that the precision of the instrument is 0·16.

Example 15·10. *Test the hypothesis that $\sigma = 10$, given that $s = 15$ for a random sample of size 50 from a normal population.*

Solution. *Null Hypothesis,* $H_0 : \sigma = 10$.

We are given $n = 50$, $s = 15$. Now $\chi^2 = \frac{ns^2}{\sigma^2} = \frac{50 \times 225}{100} = 112.5$

Since n is large, using(15·14*a*), the test statistic is : $Z = \sqrt{2\chi^2} - \sqrt{2n-1} \sim N(0, 1)$

$\therefore \quad Z = \sqrt{225} - \sqrt{99} = 15 - 9.95 = 5.05$

Since $|Z| > 3$, it is significant at all levels of significance and hence H_0 is rejected and we conclude that $\sigma \neq 10$.

15·6·2. Goodness of Fit Test. A very powerful test for testing the significance of the discrepancy between theory and experiment was given by Prof. Karl Pearson in 1900 and is known as "*Chi-square test of goodness of fit*". It enables us to find if the deviation of the experiment from theory is just by chance or is it really due to the inadequacy of the theory to fit the observed data.

If f_i $(i = 1, 2, ..., n)$ is a set of observed (experimental) frequencies and e_i $(i = 1, 2, ..., n)$ is the corresponding set of expected (theoretical or hypothetical) frequencies, then Karl Pearson's chi-square, given by :

$$\chi^2 = \sum_{i=1}^{n} \left[\frac{(f_i - e_i)^2}{e_i}\right], \quad \left(\sum_{i=1}^{n} f_i = \sum_{i=1}^{n} e_i\right) \qquad ...(15·15)$$

follows chi-square distribution with $(n-1)$ *d.f.*

Remark. This is an approximate test for large values of n. Conditions for the validity of the χ^2-test of goodness of fit have already been given in § 15·4 Remark 2 on page 15·12.

The goodness of fit test uses the chi-square distribution to determine if a hypothesized probability distribution for a population provides a good fit. Acceptance or rejection of the hypothesized population distribution is based upon differences between observed frequencies (f_i's) in a sample and the expected frequencies (e_i's) obtained under null hypothesis H_0.

Decision rule :Accept H_0 if $\chi^2 \leq \chi^2_\alpha(n-1)$ and reject H_0 if $\chi^2 > \chi^2_\alpha(n-1)$, where χ^2 is the calculated value of chi-square obtained on using (15·15) and $\chi^2_\alpha(n-1)$ is the tabulated value of chi-square for $(n-1)$ *d.f.* and level of significance α.

Example 15·11. *The demand for a particular spare part in a factory was found to vary from day-to-day. In a sample study the following information was obtained :*

Days	:	*Mon.*	*Tues.*	*Wed.*	*Thurs.*	*Fri.*	*Sat.*
No. of parts demanded	:	*1124*	*1125*	*1110*	*1120*	*1126*	*1115*

Test the hypothesis that the number of parts demanded does not depend on the day of the week. (Given : the values of chi-square significance at 5, 6, 7, d.f. are respectively 11·07, 12·59, 14·07 at the 5% level of significance.)

Solution. Here we set up the *null hypothesis*, H_0 that the number of parts demanded does not depend on the day of week.

Under the null hypothesis, the expected frequencies of the spare part demanded on each of the six days would be :

$$\frac{1}{6}(1124 + 1125 + 1110 + 1120 + 1126 + 1115) = \frac{6720}{6} = 1120$$

TABLE 15·2 : CALCULATIONS FOR χ^2

Days	Frequency		$(f_i - e_i)^2$	$\frac{(f_i - e_i)^2}{e_i}$
	Observed (f_i)	Expected (e_i)		
Mon.	1124	1120	16	0·014
Tues.	1125	1120	25	0·022
Wed.	1110	1120	100	0·089
Thurs.	1120	1120	0	0
Fri.	1126	1120	36	0·032
Sat.	1115	1120	25	0·022
Total	6720	6720		0·179

$$\chi^2 = \sum \frac{(f_i - e_i)^2}{e_i} = 0{\cdot}179$$

The number of degrees of freedom = 6 − 1 = 5 (since we are given 6 frequencies subjected to only one linear constraint : $\sum f_i = \sum e_i = 6720$)

The tabulated $\chi^2_{0{\cdot}05}$ for 5 *d.f.* = 11·07.

Since calculated value of χ^2 is less than the tabulated value, it is not significant and the null hypothesis may accepted at 5% level of significance. Hence we conclude that the number of parts demanded are same over the 6-day period.

Example 15·12. *The following figures show the distribution of digits in numbers chosen at random from a telephone directory :*

Digits :	0	1	2	3	4	5	6	7	8	9	*Total*
Frequency :	1026	1107	997	966	1075	933	1107	972	964	853	10,000

Test whether the digits may be taken to occur equally frequently in the directory.

Solution. Here we set up the *null hypothesis* that the digits occur equally frequently in the directory.

Under the null hypothesis, the expected frequency for each of the digits 0, 1, 2, ..., 9 is 10,000/10 = 1000. The value of χ^2 is computed as follows :

TABLE 15·3 : CALCULATIONS FOR χ^2

Digits	Frequency		$(f_i - e_i)^2$	$\frac{(f_i - e_i)^2}{e_i}$
	Observed (f_i)	Expected (e_i)		
0	1026	1000	676	0·676
1	1107	1000	11449	11·449
2	997	1000	9	0·009
3	966	1000	1156	1·156
4	1075	1000	5625	5·625
5	933	1000	4489	4·489
6	1107	1000	11149	11·449
7	972	1000	784	0·784
8	964	1000	1296	1·296
9	853	1000	21609	21·609
Total	10,000	10,000		58·542

$$\chi^2 = \sum \frac{(f_i - e_i)^2}{e_i}$$

$$= 58{\cdot}542$$

The number of degrees of freedom

= Number of observations – Number of independent constraints

$= 10 - 1 = 9$

Tabulated $\chi^2_{0.05}$ for 9 *d.f.* $= 16{\cdot}919$

Since the calculated value of χ^2 is much greater than the tabulated value, it is highly significant and we reject the null hypothesis. Thus we conclude that the digits are not uniformly distributed in the directory.

Example 15·13. *A sample analysis of examination results of 200 MBA's was made. It was found that 46 students had failed, 68 secured a third division, 62 secured a second division and the rest were placed in first division. Are these figures commensurate with the general examination result which is in the ratio of 4 : 3 : 2 : 1 for various categories respectively ?*

Solution. Set up the *null hypothesis* that the observed figures do not differ significantly from the hypothetical frequencies which are in the ratio of 4 : 3 : 2 : 1. In other words *the given data are commensurate with the general examination result* which is in the ratio of 4 : 3 : 2 : 1 for the various categories.

Under the null hypothesis, the expected frequencies can be computed as shown in the adjoining table :

Category	Frequency	
	Observed (f_i)	Expected (e_i)
Failed	46	$\frac{4}{10} \times 200 = 80$
III Division	68	$\frac{3}{10} \times 200 = 60$
II Division	62	$\frac{2}{10} \times 200 = 40$
I Division	24	$\frac{1}{10} \times 200 = 20$
Total	200	200

TABLE 15·4 : CALCULATIONS FOR χ^2

Category	Frequency: Observed (f_i)	Frequency: Expected (e_i)	$(f_i - e_i)^2$	$\frac{(f_i - e_i)^2}{e_i}$
Failed	46	80	1156	14·450
III Division	68	60	64	1·067
II Division	62	40	484	12·100
I Division	24	20	16	0·800
Total	200	200		28·417

$$\chi^2 = \sum \frac{(f_i - e_i)^2}{e_i} = 28{\cdot}417$$

$d.f. = 4 - 1 = 3$, tabulated $\chi^2_{0\cdot05}$ for 3 $d.f. = 7{\cdot}815$

Since the calculated value of χ^2 is greater than the tabulated value, it is significant and the null hypothesis is rejected at 5% level of significance. Hence we may conclude that data are not commensurate with the general examination result.

Example 15·14. *A survey of 800 families with four children each revealed the following distribution :*

No. of boys	:	0	1	2	3	4
No. of girls	:	4	3	2	1	0
No. of families	:	32	178	290	236	64

Is this result consistent with the hypothesis that male and female births are equally probable ?

Solution. Let us set up the *null hypothesis* that *the data are consistent with the hypothesis of equal probability for male and female births*. Then under the null hypothesis :

$$p = \text{Probability of male birth} = \tfrac{1}{2} = q.$$

$$p(r) = \text{Probability of 'r' male births in a family of } 4 = {}^4C_r \left(\tfrac{1}{2}\right)^2 \left(\tfrac{1}{2}\right)^{4-r} = {}^4C_r \left(\tfrac{1}{2}\right)^4$$

The frequency of r male births is given by :

$$f(r) = N.\ p(r) = 800 \times {}^4C_r \left(\tfrac{1}{2}\right)^4 = 50 \times {}^4C_r\ ;\ r = 0, 1, 2, 3, 4. \qquad \ldots (*)$$

Substituting $r = 0, 1, 2, 3, 4$ successively in (*), we get the expected frequencies as follows :

$f(0) = 50 \times 1 = 50,$ $\quad f(1) = 50 \times {}^4C_1 = 200,$ $\quad f(2) = 50 \times {}^4C_2 = 300,$

$f(3) = 50 \times {}^4C_3 = 200,$ $\quad f(4) = 50 \times {}^4C_4 = 50.$

given data.

15·6·3. Test of Independence of Attributes—Contingency Tables. Let us consider two attributes A and B, A divided into r classes $A_1, A_2, \ldots, A_r$ and B divided into s classes $B_1, B_2, \ldots, B_s$. Such a classification in which attributes are divided into more than two classes is known as *manifold classification*. The various cell frequencies can be expressed in the following table known as $r \times s$ *manifold contingency table* where (A_i) is the number of persons possessing the attribute A_i, $(i = 1, 2, \ldots, r)$, (B_j) is the number of persons possessing the attribute B_j $(j = 1, 2, \ldots, s)$ and (A_iB_j) is the number of persons possessing both the attributes A_i and B_j, $(i = 1, 2, \ldots, r; j = 1, 2, \ldots, s)$.

Also $\sum\limits_{i=1}^{r} (A_i) = \sum\limits_{j=1}^{s} (B_j) = N$, where N is the total frequency.

TABLE 15·7 : $r \times s$ CONTINGENCY TABLE

B \ A	A_1	A_2	$\cdots$	A_i	$\cdots$	A_r	*Total*
B_1	(A_1B_1)	(A_2B_1)	$\cdots$	(A_iB_1)	$\cdots$	(A_rB_1)	(B_1)
B_2	(A_1B_2)	(A_2B_2)	$\cdots$	(A_iB_2)	$\cdots$	(A_rB_2)	(B_2)
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
B_j	(A_1B_j)	(A_2B_j)	$\cdots$	(A_iB_j)	$\cdots$	(A_rB_j)	(B_j)
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
B_s	(A_1B_s)	(A_2B_s)	$\cdots$	(A_iB_s)	$\cdots$	(A_rB_s)	(B_s)
Total	(A_1)	(A_2)	$\cdots$	(A_i)	$\cdots$	(A_r)	N

The problem is to test if the two attributes A and B under consideration are independent or not.

Under the *null hypothesis that the attributes are independent,* the theoretical cell frequencies are calculated as follows :

$P[A_i]$ = Probability that a person possesses the attribute $A_i = \frac{(A_i)}{N}; i = 1, 2, ..., r$

$P[B_j]$ = Probability that a person possesses the attribute $B_j = \frac{(B_j)}{N}; j = 1, 2, ..., s$

$P[A_iB_j]$ = Probability that a person possesses the attributes A_i and $B_j = P(A_i)P(B_j)$

(By compound probability theorem, since the attributes A_i and B_j are independent, under the null hypothesis.)

$$\therefore \quad P[A_iB_j] = \frac{(A_i)}{N} \cdot \frac{(B_j)}{N}; i = 1, 2, ..., r; j = 1, 2, ..., s \text{ and}$$

$(A_iB_j)_0$ = Expected number of persons possessing both the attributes A_i and B_j

$= N.P[A_i B_j] = \frac{(A_i)(B_j)}{N}$

$$\Rightarrow \quad (A_i B_j)_0 = \frac{(A_i)(B_j)}{N}, (i = 1, 2, ..., r; j = 1, 2, ..., s) \quad \text{... (15·16)}$$

By using this formula, we can find out expected frequencies for each of the cell-frequencies $(A_i B_j)$ $(i = 1, 2, ..., r; j = 1, 2, ..., s)$, under the null hypothesis of independence of attributes.

The exact test for the independence of attributes is very complicated but a fair degree of approximation is given, for large samples, (large N), by the χ^2-test of goodness of fit, *viz.*,

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} \left[\frac{\{(A_iB_j) - (A_iB_j)_0\}^2}{(A_iB_j)_0} \right] = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \quad \text{...(15·16a)}$$

where f_{ij} = observed frequency for contingency table category in column i and row j,

e_{ij} = expected frequency for contingency table category in column i and row j,

which is distributed as a χ^2-variate with $(r-1)(s-1)$ *d.f.* [*c.f.* Note below on degrees of freedom].

Remarks 1. $\phi^2 = \chi^2/N$ is known as *mean-square contingency.*

Since the limits for χ^2 and ϕ^2 vary in different cases, they cannot be used for establishing the closeness of the relationship between qualitative characters under study. Prof. Karl Pearson suggested another measure, known as "*coefficient of mean square contingency*" which is denoted by C and is given by :

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} = \sqrt{\frac{\phi^2}{1 + \phi^2}} \quad \text{...(15·17)}$$

Obviously C is always less than unity. The maximum value of C depends on r and s, the number of classes into which A and B are divided. In a $r \times r$ contingency table, the maximum value of $C = \sqrt{(r - 1/r)}$. Since the maximum value of C differs for different classification, *viz.*, $r \times r$ $(r = 2, 3, 4, ...)$, strictly speaking, the values of C obtained from different types of classifications are not comparable.

2. Note on Degrees of Freedom (d.f.). The number of independent variates which make up the statistic (*e.g.*, χ^2) is known as the *degrees of freedom* (*d.f.*) and is usually denoted by υ (the letter 'Nu' of the Greek alphabet).

The number of degrees of freedom, in general, is the total number of observations less the number of independent constraints imposed on the observations. For example, if k is the number of independent constraints in a set of data of n observations then $v = (n - k)$.

Thus in a set of n observations usually, the degrees of freedom for χ^2 are $(n - 1)$, one *d.f.* being lost because of the linear constraint $\sum_i f_i = \sum_i e_i = N$, on the frequencies (*c.f.* Theorem 15·3). If 'r' independent linear constraints are imposed on the cell frequencies, then the *d.f.* are reduced by 'r'.

In addition, if any of the population parameter(s) is (are) calculated from the given data and used for computing the expected frequencies then in applying χ^2-test of goodness of fit, we have to subtract one *d.f.* for each parameter calculated. Thus if 's' is the number of population parameters estimated from the sample observations (n in number), then the required number of degrees of freedom for χ^2-test is $(n - s - 1)$.

If any one or more of the theoretical frequencies is less than 5, then in applying χ^2-test we have also to subtract the degrees of freedom lost in pooling these frequencies with the preceding or succeeding frequency (or frequencies).

In a $r \times s$ contingency table, in calculating the expected frequencies, the row totals, the column totals and the grand totals remain fixed. The fixation of 'r' column totals and 's' row totals imposes $(r + s)$ constraints on the cell frequencies. But since $\sum_{i=1}^{r} (A_i) = \sum_{j=1}^{s} (B_j) = N$, the total number of independent constraints is only $(r + s - 1)$. Further, since the total number of cell-frequencies is $r \times s$, the required number of *d.f.* is : $v = rs - (r + s - 1) = (r - 1)(s - 1)$ ...(5·17*a*)

Example 15·16. *Two sample polls of votes for two candidates A and B for a public office are taken, one from among the residents of rural areas. The results are given in the adjoining table. Examine whether the nature of the area is related to voting preference in this election.*

Area	*Votes for*		*Total*
	A	B	
Rural	620	380	1000
Urban	550	450	1000
Total	1170	830	2000

Solution. Under the *null hypothesis* that the nature of the area is independent of the voting preference in the election, we get the expected frequencies as follows :

$$E(620) = \frac{1170 \times 1000}{2000} = 585, \qquad E(380) = \frac{830 \times 1000}{2000} = 415,$$

$$E(550) = \frac{1170 \times 1000}{2000} = 585, \quad \text{and} \quad E(450) = \frac{830 \times 1000}{2000} = 415$$

Aliter. In a 2×2 contingency table, since *d.f.* $= (2 - 1)(2 - 1) = 1$, only one of the cell frequencies can be filled up independently and the remaining will follow immediately, since the observed and theoretical marginal totals are fixed. Thus having obtained any one of the theoretical frequencies (say) $E(620) = 585$, the remaining theoretical frequencies can be easily obtained as follows :

$$E(380) = 1000 - 585 = 415, \; E(550) = 1170 - 585 = 585, \text{ and } \; E(450) = 1000 - 585 = 415.$$

$$\therefore \quad \chi^2 = \sum_i \left[\frac{(f_i - e_i)^2}{e_i}\right] = \frac{(620 - 585)^2}{585} + \frac{(380 - 415)^2}{415} + \frac{(550 - 585)^2}{585} + \frac{(450 - 415)^2}{415}$$

$$= (35)^2\left(\frac{1}{585}+\frac{1}{415}+\frac{1}{585}+\frac{1}{415}\right) = (1225)[2 \times 0{\cdot}002409 + 2 \times 0{\cdot}001709] = 10{\cdot}0891$$

Tabulated $\chi^2_{0\cdot05}$ for $(2-1)(2-1) = 1$ *d.f.* is 3·841. Since calculated χ^2 is much greater than the tabulated value, it is highly significant and null hypothesis is rejected at 5% level of significance. Thus we conclude that nature of area is related to voting preference in the election.

Example 15·17. (2 × 2 CONTINGENCY TABLE). *For the* 2×2 *table,*

a	b
c	d

, prove that chi-square test of independence gives

$$\chi^2 = \frac{N(ad-bc)^2}{(a+c)(b+d)(a+b)(c+d)},\ N = a+b+c+d \qquad \ldots(15{\cdot}18)$$

Solution. Under the hypothesis of independence of attributes,

$$E(a) = \frac{(a+b)(a+c)}{N}$$

$$E(b) = \frac{(a+b)(b+d)}{N}$$

$$E(c) = \frac{(a+c)(c+d)}{N}$$

and

$$E(d) = \frac{(b+d)(c+d)}{N}$$

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	N

$$\therefore \quad \chi^2 = \frac{[a-E(a)]^2}{E(a)} + \frac{[b-E(b)]^2}{E(b)} + \frac{[c-E(c)]^2}{E(c)} + \frac{[d-E(d)]^2}{E(d)} \qquad \ldots(*)$$

$$a - E(a) = a - \frac{(a+b)(a+c)}{N} = \frac{a(a+b+c+d)-(a^2+ac+ab+bc)}{N} = \frac{ad-bc}{N}$$

Similarly, we will get : $\quad b - E(b) = -\dfrac{ad-bc}{N} = c - E(c)\ ; \quad d - E(d) = \dfrac{ad-bc}{N}$

Substituting in (*), we get

$$\chi^2 = \frac{(ad-bc)^2}{N^2}\left[\frac{1}{E(a)}+\frac{1}{E(b)}+\frac{1}{E(c)}+\frac{1}{E(d)}\right]$$

$$= \frac{(ad-bc)^2}{N}\left[\left\{\frac{1}{(a+b)(a+c)}+\frac{1}{(a+b)(b+d)}\right\}+\left\{\frac{1}{(a+c)(c+d)}+\frac{1}{(b+d)(c+d)}\right\}\right]$$

$$= \frac{(ad-bc)^2}{N}\left[\frac{b+d+a+c}{(a+b)(a+c)(b+d)}+\frac{b+d+a+c}{(a+c)(c+d)(b+d)}\right]$$

$$= (ad-bc)^2\left[\frac{c+d+a+b}{(a+b)(a+c)(b+d)(c+d)}\right] = \frac{N(ad-bc)^2}{(a+b)(a+c)(b+d)(c+d)}.$$

Remark. We can calculate the value of χ^2 for 2×2 contingency table by using (15·18) directly. The reader is advised to obtain the value of χ^2 in Example 15·16 by using (15·18).

Example 15·18. *Out of 8,000 graduates in a town 800 are females, out of 1,600 graduate employees 120 are females. Use* χ^2 *to determine if any distinction is made in appointment on the basis of sex. Value of* χ^2 *at 5% level for one degree of freedom is 3·84.*

TABLE 15.

SINGNIFICANT VALUES $\chi^2(\alpha)$ OF CHI-SQUARE DISTRIBUTION (RIGHT TAIL AREAS) FOR GIVEN PROBABILITY α,

where $$P = P_r[\chi^2 > \chi_\upsilon^2(\alpha)] = \alpha$$

AND υ IS DEGREES OF FREEDOM (*d.f.*)

* χ^2-DISTRIBUTION VALUES OF $\chi_\upsilon^2(\alpha)$

Degrees of freedom (υ)	*Probability* (α)							
	0·995	0·99	0·995	0·95	0·05	0·025	0·01	0·005
1	0·000	0·000	0·001	0·004	3·841	5·024	6·635	7·879
2	0·010	0·020	0·051	0·103	5·991	7·378	9·210	10·597
3	0·072	0·115	0·216	0·352	7·815	9·348	11·345	12·838
4	0·207	0·297	0·484	0·711	9·488	11·143	13·277	14·860
5	0·412	0·554	0·831	1·145	11·070	12·832	15·086	16·750
6	0·676	0·872	1·237	1·634	12·592	14·449	16·812	18·548
7	0·989	1·239	1·690	2·167	14·067	16·013	18·475	20·278
8	1·344	1·646	2·180	2·733	15·507	17·535	20·090	21·955
9	1·735	2·088	2·700	3·325	16·919	19·023	21·666	23·589
10	2·156	2·558	3·247	3·940	18·307	20·483	23·209	25·188
11	2·603	3·053	3·816	4·575	19·675	21·920	24·725	26·757
12	3·074	3·571	4·404	5·226	21·026	23·337	26·217	28·300
13	3·565	4·107	5·009	5·892	22·362	24·736	24·888	29·819
14	4·075	4·660	5·629	6·571	23·685	26·119	29·141	31·319
15	4·601	5·229	6·262	7·261	24·996	27·488	30·578	32·801
16	5·142	5·812	6·908	7·962	26·296	28·845	32·000	34·267
17	5·697	6·408	7·564	8·672	27·587	30·191	33·409	35·718
18	6·265	7·015	8·231	9·390	28·869	31·526	34·805	37·156
19	6·844	7·633	8·907	10·117	30·144	32·852	36·191	38·582
20	7·434	8·260	9·591	10·851	31·410	34·170	37·566	39·997
21	8·034	8·897	10·283	11·591	32·671	35·479	38·932	41·401
22	8·643	9·542	10·982	12·338	33·924	36·781	40·289	42·796
23	9·260	10·196	11·688	13·091	35·172	38·076	41·638	44·181
24	9·886	10·856	12·401	13·848	36·415	39·364	42·980	45·558
25	10·520	11·524	13·120	14·611	37·652	40·646	44·314	46·928
26	11·160	12·198	13·844	15·379	38·885	41·923	45·642	48·290
27	11·808	12·879	14·573	16·151	40·113	43·194	46·963	49·645
28	12·461	13·565	15·308	16·928	41·337	44·461	48·278	50·993
29	13·121	14·256	16·047	17·708	42·557	45·722	49·588	52·336
30	13·787	14·953	16·791	18·493	43·773	46·979	50·892	53·672
40	20·706	22·164	24·433	26·509	55·759	59·342	63·691	66·766
50	27·991	29·707	32·357	34·764	67·505	71·420	76·154	79·490
60	35·535	37·485	40·482	43·188	79·082	83·298	88·379	91·952
70	43·275	45·442	48·758	51·739	90·531	95·023	100·425	104·215
80	51·172	53·540	57·153	60·391	101·879	106·629	112·329	116·321
90	59·196	61·754	65·647	69·126	113·145	118·136	124·116	128·299
100	67·328	70·065	74·222	77·929	124·342	129·561	135·807	140·169

For larger values of v, quantity $\sqrt{2\chi^2} - \sqrt{2v-1}$ may be used as a standard normal variable.

---

* *Abridged from Table 8 of Biometrika Tables for Statisticians, Vol. I.*

16·3. APPLICATIONS OF t-DISTRIBUTION

The t-distribution has a wide number of applications in Statistics, some of which are enumerated below.

(i) To test if the sample mean ($\bar{x}$) differs significantly from the hypothetical value μ of the population mean :

(ii) To test the significance of the difference between two sample means.

(iii) To test the significance of an observed sample correlation coefficient and sample regression coefficient.

(iv) To test the significance of observed partial correlation coefficient.

In the following sections we will discuss these applications in detail, one by one.

16·3·1. *t*-Test for Single Mean. Suppose we want to test :

(i) if a random sample x_i ($i = 1, 2, \ldots, n$) of size n has been drawn from a normal population with a specified mean, say μ_0, or

(ii) if the sample mean differs significantly from the hypothetical value μ_0 of the population mean.

Under the null hypothesis, H_0 :

(i) The sample has been drawn from the population with mean μ_0 or

(ii) there is no significant difference between the sample mean $\bar{x}$ and the population mean μ_0,

the statistic
$$t = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}, \qquad \ldots(16\cdot6)$$

where
$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2, \qquad \ldots[16\cdot6a)$$

follows Student's t-distribution with $(n - 1)$ *d.f.*

We now compare the calculated value of t with the tabulated value at certain level of significance. If calculated $|t|$ > tabulated t, null hypothesis is rejected and if calculated $|t|$ < tabulated t, H_0 may be accepted at the level of significance adopted.

Remarks **1.** ***On computation of S^2 for numerical problems.*** If $\bar{x}$ comes out in integers, the formula (16·6a) can be conveniently used for computing S^2. However, if $\bar{x}$ comes in fractions then the formula (16·6a) for computing S^2 is very cumbersome and is not recommended. In that case, step deviation method, given below, is quite useful.

If we take $d_i = x_i - A$, where A is any arbitrary number, then

$$S^2 = \frac{1}{n-1}\left[\Sigma(x_i - \bar{x})^2\right] = \frac{1}{n-1}\left[\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}\right] \qquad \ldots(16\cdot6b)$$

$$= \frac{1}{n-1}\left[\Sigma d_i^2 - \frac{(\Sigma d_i)^2}{n}\right], \text{ since variance is independent of change of origin.} \qquad \ldots(16\cdot6c)$$

Also, in this case
$$\bar{x} = A + \frac{\Sigma d_i}{n}. \qquad \ldots(16\cdot6d)$$

2. We know, the sample variance : $s^2 = \frac{1}{n}\sum_i (x_i - \bar{x})^2 \quad \Rightarrow \quad ns^2 = (n-1)\,S^2$

$$\therefore \qquad \frac{S^2}{n} = \frac{s^2}{n-1} \qquad \ldots(16\cdot6e)$$

Hence for numerical problems, the test statistic (16·6) on using [16·6(e)] becomes

$$t = \frac{\bar{x} - \mu_0}{\sqrt{S^2/n}} = \frac{\bar{x} - \mu_0}{\sqrt{s^2/(n-1)}} \sim t_{n-1} \qquad \ldots(16\cdot6f)$$

3. Assumption for Student's t-test. The following assumptions are made in the Student's t-test :

(i) The parent population from which the sample is drawn is normal.
(ii) The sample observations are independent, *i.e.*, the sample is random.
(iii) The population standard deviation σ is unknown.

Example 16·5. *A machinist is making engine parts with axle diameters of 0·700 inch. A random sample of 10 parts shows a mean diameter of 0·742 inch with a standard deviation of 0·040 inch. Compute the statistic you would use to test whether the work is meeting the specifications. Also state how you would proceed further.*

Solution. Here we are given :

$$\mu = 0{\cdot}700 \text{ inche}, \quad \bar{x} = 0{\cdot}742 \text{ inche}, \quad s = 0{\cdot}040 \text{ inche} \quad \text{and} \quad n = 10$$

Null Hypothesis, $H_0 : \mu = 0{\cdot}700$, *i.e., the product is conforming to specifications.*

Alternative Hypothesis, $H_1 : \mu \neq 0{\cdot}700$

Test Statistic. Under H_0, the test statistic is : $t = \dfrac{\bar{x} - \mu}{\sqrt{S^2/n}} = \dfrac{\bar{x} - \mu}{\sqrt{s^2/(n-1)}} \sim t_{(n-1)}$

$$\therefore \qquad t = \frac{\sqrt{9}\,(0{\cdot}742 - 0{\cdot}700)}{0{\cdot}040} = 3{\cdot}15$$

How to proceed further. Here the test statistic 't' follows Student's t-distribution with $10 - 1 = 9$ *d.f.* We will now compare this calculated value with the tabulated value of t for 9 *d.f.* and at certain level of significance, say 5%. Let this tabulated value be denoted by t_0.

(i) If calculated 't', *viz.*, $3{\cdot}15 > t_0$, we say that the value of t is significant. This implies that $\bar{x}$ differs significantly from μ and H_0 is rejected at this level of significance and we conclude that the product is not meeting the specifications.

(ii) If calculated $t < t_0$, we say that the value of t is not significant, *i.e.*, there is no significant difference between $\bar{x}$ and μ. In other words, the deviation $(\bar{x} - \mu)$ is just due to fluctuations of sampling and null hypothesis H_0 may be retained at 5% level of significance, *i.e.*, we may take the product conforming to specifications.

Example 16·6. *The mean weekly sales of soap bars in departmental stores was 146·3 bars per store. After an advertising campaign the mean weekly sales in 22 stores for a typical week increased to 153·7 and showed a standard deviation of 17·2. Was the advertising campaign successful ?*

Solution. We are given : $n = 22$, $\bar{x} = 153{\cdot}7$, $s = 17{\cdot}2$.

Null Hypothesis. The advertising campaign is not successful, *i.e.*, $H_0 : \mu = 146{\cdot}3$

Alternative Hypothesis, $H_1 : \mu > 146{\cdot}3$ (Right-tail).

Test Statistic. Under H_0, the test statistic is : $t = \dfrac{\bar{x} - \mu}{\sqrt{s^2/(n-1)}} \sim t_{22-1} = t_{21}$

$$\therefore \qquad t = \frac{153{\cdot}7 - 146{\cdot}3}{\sqrt{(17{\cdot}2)^2/21}} = \frac{7{\cdot}4 \times \sqrt{21}}{17{\cdot}2} = 9{\cdot}03$$

Conclusion. Tabulated value of t for 21 *d.f.* at 5% level of significance for *single-tailed test* is 1·72. Since calculated value is much greater than the tabulated value, it is

Under the null hypothesis, H_0 that (*a*) *samples have been drawn from the populations with same means, i.e.,* $\mu_X = \mu_Y$, *or* (*b*) *the sample means* $\bar{x}$ *and* $\bar{y}$ *do not differ significantly,* the statistic :

$$t = \frac{\bar{x} - \bar{y}}{S\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad [\because \mu_X = \mu_Y, \text{ under } H_0] \qquad \ldots(16.8)$$

where symbols are defined in (16·7*a*), follows Student's *t*-distribution with $(n_1 + n_2 - 2)$ *d.f.*

3. *On the assumption of t-test for difference of means.* Here we make the following three fundamental assumptions :

(*i*) Parent populations, from which the samples have been drawn are normally distributed.

(*ii*) The population variances are equal and unknown, *i.e.*, $\sigma_X^2 = \sigma_Y^2 = \sigma^2$ (say), where σ^2 is unknown.

(*iii*) The two samples are random and independent of each other.

Thus before applying *t*-test for testing the equality of means it is theoretically desirable to test the equality of population variances by applying *F*-test. (*c.f* § 16·6·1) If the variances do not come out to be equal then *t*-test becomes invalid and in that case Behren's '*d*'-test based on fiducial intervals is used. For practical problems, however, the assumptions (*i*) and (*ii*) are taken for granted.

16·3·3. Paired t-test for Difference of Means. Let us now consider the case when (*i*) the sample sizes are equal, *i.e.*, $n_1 = n_2 = n$ (say), and (*ii*) the two samples are not independent but the sample observations are paired together, *i.e.*, the pair of observations (x_i, y_i), $(i = 1, 2, \ldots, n)$ corresponds to the same (*i*th) sample unit. The problem is to test if the sample means differ significantly or not.

For example, suppose we want to test the efficacy of a particular drug, say, for inducing sleep. Let x_i and y_i $(i = 1, 2, \ldots, n)$ be the readings, in hours of sleep, on the *i*th individual, before and after the drug is given respectively. Here instead of applying the difference of the means test discussed in § 16·3·2, we apply the paired *t*-test given below.

Here we consider the increments, $d_i = x_i - y_i$, $(i = 1, 2, \ldots, n)$.

Under the null hypothesis, H_0 that *increments are due to fluctuations of sampling, i.e., the drug is not responsible for these increments,* the statistic :

$$t = \frac{\bar{d}}{S/\sqrt{n}} \qquad \ldots(16.9)$$

where

$$\bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i \text{ and } S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (d_i - \bar{d})^2 \qquad \ldots(16.9a)$$

follows Student's *t*-distribution with $(n - 1)$ *d.f.*

Example 16·10. *Below are given the gain in weights (in kgs.) of pigs fed on two diets A and B.*

Gain in weight

Diet A : 25, 32, 30, 34, 24, 14, 32, 24, 30, 31, 35, 25

Diet B : 44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21, 35, 29, 22

Test, if the two diets differ significantly as regards their effect on increase in weight.

Solution. *Null hypothesis,* $H_0 : \mu_X = \mu_Y$, *i.e., there is no significant difference between the mean increase in weight due to diets A and B.*

Alternative hypothesis, H_1: $\mu_X \neq \mu_Y$ (two-tailed).

Diet A			Diet B		
x	$x-\bar{x}$	$(x-\bar{x})^2$	y	$y-\bar{y}$	$(y-\bar{y})^2$
25	−3	9	44	14	196
32	4	16	34	4	16
30	2	4	22	−8	64
34	6	36	10	−20	400
24	−4	16	47	17	289
14	−14	196	31	1	1
32	4	16	40	10	100
24	−4	16	30	0	0
30	2	4	32	2	4
31	3	9	35	5	25
35	7	49	18	−12	144
25	−3	9	21	−9	81
			35	5	25
			29	−1	1
			22	−8	64
$\Sigma x = 336$	$\Sigma(x-\bar{x}) = 0$	$\Sigma(x-\bar{x})^2 = 380$	$\Sigma y = 450$	$\Sigma(y-\bar{y}) = 0$	$\Sigma(y-\bar{y})^2 = 1{,}410$

$$\bar{x} = \frac{336}{12} = 28,\ \bar{y} = \frac{450}{15} = 30,\ S^2 = \frac{1}{n_1+n_2-2}\left[\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2\right] = 71{\cdot}6$$

and $n_1 = 12$, $n_2 = 15$

Under null hypothesis (H_0) :

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}$$

$$\therefore \qquad t = \frac{28-30}{\sqrt{71{\cdot}6\left(\frac{1}{12}+\frac{1}{15}\right)}} = \frac{-2}{\sqrt{10{\cdot}74}} = -0{\cdot}609$$

Tabulated $t_{0.05}$ for $(12 + 15 - 2) = 25$ *d.f.* is 2·06.

Conclusion. Since calculated $|t|$ is less than tabulated t, H_0 may be accepted at 5% level of significance and we may conclude that the two diets do not differ significantly as regards their effect on increase in weight.

Remark. Here $\bar{x}$ and $\bar{y}$ come out to be integral values and hence the direct method of computing $\Sigma(x-\bar{x})^2$ and $\Sigma(y-\bar{y})^2$ is used. In case $\bar{x}$ and (or) $\bar{y}$ comes out to be fractional, then the step deviation method is recommended for computation of $\Sigma(x-\bar{x})^2$ and $\Sigma(y-\bar{y})^2$.

Example 16·11. *Samples of two types of electric light bulbs were tested for length of life and following data were obtained :*

	Type I	*Type II*
Sample No.	$n_1 = 8$	$n_2 = 7$
Sample Means	$\bar{x}_1 = 1{,}234$ *hrs.*	$\bar{x}_2 = 1{,}036$ *hrs.*
Sample S.D.'s	$s_1 = 36$ *hrs.*	$s_2 = 40$ *hrs.*

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

Solution. *Null Hypothesis*, $H_0 : \mu_X = \mu_Y$, *i.e.*, the two types I and II of electric bulbs are indentical.

Alternative Hypothesis, $H_1 : \mu_X > \mu_Y$, *i.e.*, type I is superior to type II.

Test Statistic. Under H_0, the test statistic is :

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t_{n_1 + n_2 - 2} = t_{13},$$

where

$$S^2 = \frac{1}{n_1 + n_2 - 2}\left[\Sigma(x_1 - \bar{x}_1)^2 + \Sigma(x_2 - \bar{x}_2)^2\right]$$

$$= \frac{1}{n_1 + n_2 - 2}\left(n_1 s_1^2 + n_2 s_2^2\right) = \frac{1}{13}\left[8 \times (36)^2 + 7 \times (40)^2\right] = 1{,}659{\cdot}08$$

$$\therefore \quad t = \frac{1234 - 1036}{\sqrt{1659{\cdot}08\left(\frac{1}{8} + \frac{1}{7}\right)}} = \frac{198}{\sqrt{1659{\cdot}08 \times 0{\cdot}2679}} = 9{\cdot}39$$

Tabulated value of t for 13 *d.f.* at 5% level of significance for right (single)-tailed test is 1·77. [This is the value of $t_{0{\cdot}10}$ for 13 *d.f.* from two-tail tables given at the end of the chapter.]

Conclusion. Since calculated 't' is much greater than tabulated 't', it is highly significant and H_0 is rejected. Hence the two types of electric bulbs differ significantly. Further, since $\bar{x}_1$ is much greater than $\bar{x}_2$, we conclude that type I is definitely superior to type II.

Example 16·16. *Two laboratories carry out independent estimates of a particular chemicals in a medicine produced by a certain firm. A sample is taken from each batch, halved and the separate halves sent to the two laboratories. The following data is obtained :*

n	*No. of samples*	10
$\bar{d}$	*Mean value of the difference of estimates*	0·6
$\Sigma(d_i - \bar{d})^2$	*Sum of the squares of the differences (from their means)*	20

Is the difference significant ? (Value of t at 5% level for 9 d.f. is 2·262.)

Solution. Let d stand for the difference between the estimates of the chemical between the two halves of each batch, and $\bar{d}$ the mean value of the difference of estimates. In usual notations, we are given :

$$n = 10,\ \bar{d} = 0{\cdot}6,\ \Sigma(d - \bar{d})^2 = 20$$

Null hypothesis, $H_0 : \mu_1 = \mu_2$, *i.e.*, the difference is insignificant.

Alternative hypothesis, $H_1 : \mu_1 \neq \mu_2$.

Test Statistic. Under H_0, the test statistic is : $t = \dfrac{\bar{d}}{\sqrt{S^2/n}} \sim t_{10-1}$

where $S^2 = \dfrac{1}{n-1}\Sigma(d - \bar{d})^2 = \dfrac{20}{9} = 2{\cdot}22 \quad \therefore \quad t = \dfrac{0{\cdot}6}{\sqrt{2{\cdot}22/10}} = \dfrac{0{\cdot}6}{0{\cdot}471} = 1{\cdot}274.$

The tabulated value of t at 5% level for 9 *d.f.*, is 2·262 (given).

Conclusion. Since calculated value of t is less than tabulated value of t, it is not significant. Hence, we may accept the null hypothesis and conclude that the difference is not significant.

16·6·1. F-test for Equality of Two Population Variances. Suppose we want to test (*i*) whether two independent samples x_i, $(i = 1, 2, \ldots, n_1)$ and y_j, $(j = 1, 2, \ldots, n_2)$ have been drawn from the normal populations with the same variance σ^2 (say), or (*ii*) whether the two independent estimates of the population variance are homogeneous or not.

Under the null hypothesis (H_0) that (*i*) $\sigma_X^2 = \sigma_Y^2 = \sigma^2$, *i.e., the population variances are equal, or (ii) Two independent estimates of the population variance are homogeneous,* the statistic F is given by :

$$F = \frac{S_X^2}{S_Y^2} \qquad \ldots(16\cdot17)$$

where $\quad S_X^2 = \dfrac{1}{n_1 - 1} \sum\limits_{i=1}^{n_1} (x_i - \bar{x})^2 \quad$ and $\quad S_Y^2 = \dfrac{1}{n_2 - 1} \sum\limits_{j=1}^{n_2} (y_j - \bar{y})^2 \qquad \ldots(16\cdot17a)$

are unbiased estimates of the common population variance σ^2 obtained from two independent samples and it follows Snedecor's F-distribution with (ν_1, ν_2) *d.f.* where $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$.

Proof. $F = \dfrac{S_X^2}{S_Y^2} = \left[\dfrac{n_1}{n_1 - 1} s_X^2\right] \Big/ \left[\dfrac{n_2}{n_2 - 1} s_Y^2\right]$

$$= \left[\frac{n_1 s_X^2}{\sigma_X^2} \cdot \frac{1}{(n_1 - 1)}\right] \Big/ \left[\frac{n_2 s_Y^2}{\sigma_Y^2} \cdot \frac{1}{(n_2 - 1)}\right] \qquad (\because \sigma_X^2 = \sigma_Y^2 = \sigma^2, \text{ under } H_0)$$

Since $\dfrac{n_1 s_X^2}{\sigma_X^2}$ and $\dfrac{n_2 s_Y^2}{\sigma_Y^2}$ are independent chi-square variates with $(n_1 - 1)$ and $(n_2 - 1)$ *d.f.* respectively, F follows Snedecor's F-distribution with $(n_1 - 1, n_2 - 1)$ *d.f.* (*c.f.* § 16·5).

Example 16·26. *In one sample of 8 observations, the sum of the squares of deviations of the sample values from the sample mean was 84·4 and in the other sample of 10 observations it was 102·6. Test whether this difference is significant at 5 per cent level, given that the 5 per cent point of F for $n_1 = 7$ and $n_2 = 9$ degrees of freedom is 3·29.*

Solution. Here $n_1 = 8, n_2 = 10$ and $\Sigma(x - \bar{x})^2 = 84{\cdot}4, \ \Sigma(y - \bar{y})^2 = 102{\cdot}6$

$$\therefore \quad S_X^2 = \frac{1}{n_1 - 1}\Sigma(x - \bar{x})^2 = \frac{84{\cdot}4}{7} = 12{\cdot}057$$

$$S_Y^2 = \frac{1}{n_2 - 1}\Sigma(y - \bar{y})^2 = \frac{102{\cdot}6}{9} = 11{\cdot}4$$

Under $H_0 : \sigma_X^2 = \sigma_Y^2 = \sigma^2$, *i.e.*, the estimates of σ^2 given by the samples are homogeneous, the test statistic is :

$$F = \frac{S_X^2}{S_Y^2} = \frac{12{\cdot}057}{11{\cdot}4} = 1{\cdot}057$$

Tabulated $F_{0{\cdot}05}$ for (7, 9) *d.f.* is 3·29.

Since calculated $F < F_{0{\cdot}05}$, H_0 may be accepted at 5% level of significance.