TESTS OF SIGNIFICANCE

Seema Jaggi
I.A.S.R.I., Library Avenue, New Delhi – 110012
seema@iasri.res.in

In applied investigations, one is often interested in comparing some characteristic (such as the mean, the variance or a measure of association between two characters) of a group with a specified value, or in comparing two or more groups with regard to the characteristic. For instance, one may wish to compare two varieties of wheat with regard to the mean yield per hectare or to know if the genetic fraction of the total variation in a strain is more than a given value or to compare different lines of a crop in respect of variation between plants within lines. In making such comparisons one cannot rely on the mere numerical magnitudes of the index of comparison such as the mean, variance or measure of association. This is because each group is represented only by a sample of observations and if another sample were drawn the numerical value would change. This variation between samples from the same population can at best be reduced in a well-designed controlled experiment but can never be eliminated. One is forced to draw inference in the presence of the sampling fluctuations which affect the observed differences between groups, clouding the real differences. Statistical science provides an objective procedure for distinguishing whether the observed difference connotes any real difference among groups. Such a procedure is called a **test of significance**. The test of significance is a method of making due allowance for the sampling fluctuation affecting the results of experiments or observations. The fact that the results of biological experiments are affected by a considerable amount of uncontrolled variation makes such tests necessary. These tests enable us to decide on the basis of the sample results, if

i) the deviation between the observed sample statistic and the hypothetical parameter value, or
ii) the deviation between two sample statistics,

is significant or might be attributed to chance or the fluctuation of sampling.

For applying the tests of significance, we first set up a hypothesis - a definite statement about the population parameters. In all such situations we set up an exact hypothesis such as, the treatments or variate in question do not differ in respect of the mean value, or the variability, or the association between the specified characters, as the case may be, and follow an objective procedure of analysis of data which leads to a conclusion of either of two kinds:

i) reject the hypothesis, or
ii) not reject the hypothesis

1. Test of Significance for Large Samples

For large n (sample size), almost all the distributions can be approximated very closely by a normal probability curve, we therefore use the **normal test** of significance for large samples. If t is any statistic (function of sample values), then for large sample

$$Z = \frac{t - E(t)}{\sqrt{V(t)}} = N(0.1)$$

Thus if the discrepancy between the observed and the expected (hypothetical) value of a statistic is greater than Z_α times the standard error (S.E.), hypothesis is rejected at α level of significance. Similarly if

$$|t - E(t)| \le Z_\alpha \times S.E(t),$$

the deviation is not regarded significant at 5% level of significance. In other words the deviation t - E(t), could have arisen due to fluctuations of sampling and the data do not provide any evidence against the null hypothesis which may, therefore be accepted at α level of significance.

If $|Z| \le 1.96$, then the hypothesis H_0 is accepted at 5% level of significance. Thus the steps to be used in the normal test are as follows:

i) Compute the test statistic Z under H_0.
ii) If $|Z| > 3$, H_0 is always rejected
iii) If $|Z| < 3$, we test its significance at certain level of significance

The table below gives some critical values of Z:

Level of Significance	Critical value (Z_α) of Z	
	Two-tailed test	**Single tailed test**
10%	1.645	1.28
5%	1.96	1.645
1%	2.58	2.33

1.1 Test for Single Mean

A very important assumption underlying the tests of significance for variables is that the sample mean is asymptotically normally distributed even if the parent population from which the sample is drawn is not normal.

If x_i (i =1,…,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., $\bar{x} \sim N(\mu, \sigma^2/n)$

H_0 : Population mean $\mu =$ a given value μ_0 ; $H_1 : \mu \ne \mu_0$

Test Statistic:

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

If σ^2 is unknown, then it is estimated by sample variance i.e., $\sigma^2 = s^2$ (for large n).

Example 1.1.1: A sample of 900 members has a mean of 3.4 cms and standard deviation (s.d.) 2.61 cms. Is the sample drawn from a large population of mean 3.25 cms?

Solution:

H_0 : The sample has been drawn from the population with mean $\mu = 3.25$ cm
H_1: $\mu \neq 3.25$ (two tailed test)
Here $\bar{x} = 3.4$ cm, n = 900,
$\mu = 3.25$ cm, $\sigma = 2.61$ cm

Under H_0, $Z = \frac{3.40 - 3.25}{2.61/\sqrt{900}} = 1.73$.

Since $|Z| < 1.96$, we conclude that the data does not provide any evidence against the null hypothesis H_0 which may therefore be accepted at 5% level of significance.

1.2 Test for Difference of Means

Let $\bar{x}_1 (\bar{x}_2)$ be the mean of a sample of size n_1 (n_2) from a population with mean μ_1 (μ_2) and variance $\sigma_1^2 (\sigma_2^2)$. Therefore

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

The difference $(\bar{x}_1 - \bar{x}_2)$ is also a normal variate
Test Statistic:

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

$$Z = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Under the null hypothesis $H_0 : \mu_1 = \mu_2$

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

Therefore $Z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$, If $\sigma_1^2 = \sigma_2^2 = \sigma^2$

If σ is not known, then its estimate is used

$$\hat{\sigma}^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}$$

1.3 Test for Single Proportion

Suppose in a sample of size n, x be the number of persons possessing the given attribute. Then observed proportion of successes = $\frac{x}{n} = p$

$$E(p) = E\left(\frac{x}{n}\right) = \frac{1}{n} E(x) = P \text{ (population proportion)}$$

and $V(p) = \frac{PQ}{n}$, $Q = 1 - P$

The normal test for the proportion of successes becomes

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

Example 1.3.1: In a sample of 1000 people, 540 are rice eaters and the rest are wheat eaters. Can we assume that both rice and wheat are equally popular at 1% level of significance.

Solution: It is given that n = 1000, x = No. of rice eaters = 540,

p = sample proportion of rice eaters = $\frac{540}{1000} = 0.54$,

P = Population proportion of rice eaters = $\frac{1}{2} = 0.5$.

H_0 : Both rice and wheat are equally popular; $H_1 : P \neq 0.5$

$$Z = \frac{p - P}{\sqrt{PQ/n}} = \frac{0.54 - 0.5}{\sqrt{0.5 \times 0.5/1000}} = 2.532$$

Since computed Z < 2.58 at 1% level of significance, therefore H_o is not rejected and we conclude that rice and wheat are equally popular.

1.4 Test for Difference of Proportions

Suppose we want to compare two populations with respect to the prevalence of a certain attribute A. Let x_1 (x_2) be the number of persons possessing the given attribute A in random sample of size n_1 (n_2) from 1^{st} (2^{nd}) population. Then sample proportions will be

$$p_1 = \frac{x_1}{n_1}, \ p_2 = \frac{x_2}{n_2}$$

Let P_1 and P_2 be the population proportions.

$$V(p_1) = \frac{P_1Q_1}{n_1}, \quad V(p_2) = \frac{P_2Q_2}{n_2} ,$$

Test Statistic:

$$Z = \frac{p_1 - p_2 - (P_1 - P_2)}{\sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}} \sim N(0,1)$$

Under $H_0 : P_1 = P_2 = P$ i.e. no significant difference between population proportions

$$Z = \frac{p_1 - p_2}{\sqrt{PQ(\frac{1}{n_1} + \frac{1}{n_2})}}$$

2. Test of Significance for Small Samples

In this section, the statistical tests based on t, χ^2 and F are given.

2.1 Tests Based on t-Distribution

2.1.1 Test for an Assumed Population Mean

Suppose a random sample $x_1, .., x_n$ of size n ($n \geq 2$) has been drawn from a normal population whose variance σ^2 is unknown. On the basis of this random sample the aim is to test

$H_0: \quad \mu = \mu_0$

$H_0: \quad \mu \neq \mu_0$ (two-tailed)

$\mu > \mu_0$ (right-tailed)

$\mu < \mu_0$ (left-tailed)

Test statistic:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

The table giving the value of t required for significance at various levels of probability and for different degrees of freedom are called the t – tables which are given in Statistical Tables by Fisher and Yates. The computed value is compared with the tabulated value at 5 or 1 percent levels of significance and at (n-1) degrees of freedom and accordingly the null hypothesis is accepted or rejected.

2.1.2 Test for the Difference of Two Population Means

Let $\bar{x}_1 (\bar{x}_2)$ be the sample mean of a sample of size n_1 (n_2) from a population with mean μ_1 (μ_2) and variance of the two population be same σ^2, which is unknown.

$H_0: \quad \mu_1 - \mu_2 = \delta_0$

Since $\bar{x}_1 \sim N(\mu_1, \sigma^2 / n_1)$; $\bar{x}_2 \sim N(\mu_2, \sigma^2 / n_2)$. Therefore, $\bar{x}_1 - \bar{x}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2})$.

Test statistic: Under H_0

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \delta_0}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1 + n_2 - 2}$$

Since σ^2 is unknown, therefore, it is estimated from the sample

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

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \delta_0}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1 + n_2 - 2}$$

If $\delta_0 = 0$, this test reduces to test the equality of two population means.

Example 2.1.2.1: A group of 5 plots treated with nitrogen at 20 kg/ha. yielded 42, 39, 48, 60 and 41 kg whereas second group of 7 plots treated with nitrogen at 40 kg/ha. yielded 38, 42, 56, 64, 68, 69 and 62 kg. Can it be concluded that nitrogen at level 40 kg/ha. increases the yield significantly?

Solution: H_0: $\mu_1 = \mu_2$, H_1: $\mu_1 < \mu_2$

$$\bar{x}_1 = 46, \quad \bar{x}_2 = 57, \quad s^2 = 121.6$$

$$t = \frac{46 - 57}{\sqrt{121.6(\frac{1}{5} + \frac{1}{7})}} = -1.7 \sim t_{10}$$

Since |t| < 1.81 (value of t at 5% and 10 d.f.), the yield from two doses of nitrogen do not differ significantly.

2.1.3 Paired t-test for Difference of Means

When $n_1 = n_2 = n$ and the two samples are not independent but the sample observations are paired together, then this test is applied. Let (x_i, y_i), i = 1,..,n be a random sample from a bivariate normal population with parameters $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Let $d_i = x_i - y_i$

H_0 : $\mu_1 - \mu_2 = \mu_0$

Test statistic:

$$t = \frac{\bar{d} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$$

where $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2$.

2.1.4 Test for Significance of Observed Correlation Coefficient

Given a random sample (x_i, y_i) , i = 1,…, n from a bivariate normal population. We want to test the null hypothesis that the population correlation coefficient is zero i.e.

$H_0 : \rho = 0$; $H_1 : \rho \neq 0$

Test Statistic: $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t_{n-2}$,

where r is the sample correlation coefficient. H_0 is rejected at level α if $|t| > t_{n-2}(\alpha/2)$. This test can also be used for testing the significance of rank correlation coefficient.

2.2 Test of Significance Based on Chi-Square Distribution

2.2.1 Test for the Variance of a Normal Population

Let $x_1, x_2, \ldots, x_n$ $(n \geq 2)$ be a random sample from $N(\mu, \sigma^2)$. H_0 : $\sigma^2 = \sigma_0^2$.

Test statistic: $\chi^2 = \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma_0} \right)^2 \sim \chi_n^2$, when μ is known

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma_0} \right)^2 \sim \chi_n^2, \text{ when } \mu \text{ is not known}$$

Tables are available for χ^2 at different levels of significance and with different degrees of freedom.

2.2.2 Test for Goodness of Fit

A test of wide applicability to numerous problems of significance in frequency data is the χ^2 test of goodness of fit. It is primarily used for testing the discrepancy between the expected and the observed frequency, for instance, in comparing an observed frequency distribution with a theoretical one like the normal.

H_0 : the fitted distribution is a good fit to the given data; H_1 : not a good fit.

Test statistic: If O_i and E_i , i =1,…,n are respectively the observed and expected frequency of i^{th} class, then the statistic

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{n-r-1}$$

where r is the number of parameters estimated from the sample, n is the number of classes after pooling. H_0 is rejected at level α if calculated χ^2 > tabulated χ^2_{n-r-1} (α)

Example 2.2.1: In an F_2 population of chillies, 831 plants with purple and 269 with non-purple chillies were observed. Is this ratio consistent with a single factor ratio of 3:1?

Solution: On the hypothesis of a ratio of 3:1, the frequencies expected in the purple and non-purple classes are 825 and 275 respectively.

	Frequency		
	Observed (O_i)	Expected (E_i)	O_i - E_i
Purpose	831	825	6
Non-purple	269	275	-6

$$\chi^2 = \sum_{i=1}^{2} \frac{(O_i - E_i)^2}{E_i} = 0.17$$

Here χ^2 is based on one degree of freedom. It is seen from the table that the value of 0.17 for χ^2 with 1 d.f corresponds to a level of probability which lies between 0.5 and 0.7. It is concluded that the result is non-significant.

2.2.3 Test of Independence

Another common use of the χ^2 test is in testing independence of classifications in what are known as contingency tables. When a group of individuals can be classified in two ways the result of the classification in two ways the results of the classification can be set out as follows:

Contingency table

Class	A_1	A_2	A_3
B_1	n_{11}	n_{21}	n_{31}
B_2	n_{12}	n_{22}	n_{32}
B_3	n_{13}	n_{23}	n_{33}

Such a table giving the simultaneous classification of a body of data in two different ways is called contingency table. If there are r rows and c columns the table is said to be an r x c table.

H_0: the attributes are independent

H_1: they are not independent

Test statistic:

$$\chi^2 = \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2_{(r-1)(c-1)}$$

H_0 is rejected at level α if $\chi^2 > \chi^2_{(r-1)(c-1)}$

2.3 Test of Significance Based on F-Distribution

2.3.1 Test for the Comparison of Two Population Variances

Let x_i , i = 1,…,n_1 and x_j, j=1,…,n_2 be the two random samples of sizes n_1 and n_2 drawn from two independent normal populations $N(\mu_1,\sigma_1^2)$ and $N(\mu_1,\sigma_2^2)$ respectively. s_1^2 and s_2^2 are the sample variances of the two samples.

$$s_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(x_i - \bar{x}_1)^2 \text{ and } s_2^2 = \frac{1}{n_2 - 1}\sum_{j=1}^{n_2}(x_j - \bar{x}_2)^2$$

$$\bar{x}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i, \quad \bar{x}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} x_j$$

H_0 : the ratio of two population variances is specified

$$\frac{\sigma_1^2}{\sigma_2^2} = \sigma_0^2$$

Test statistic: Assuming $s_1^2 > s_2^2$

$$F = \frac{\sigma_2^2}{\sigma_1^2}\frac{s_1^2}{s_2^2} \sim F_{n_1-1, n_2-1}$$

Under H_0 : $F = \frac{s_1^2}{s_2^2}\frac{1}{\sigma_0^2} \sim F_{n_1-1, n_2-1}$.

Tables are available giving the values of F required for significance at different levels of probability and for different degrees of freedom. The computed value of F is compared with the tabulated value and the inference is drawn accordingly.