Chapter 6
Hypothesis Testing for Decision-Making

Learning Objectives

After completing this chapter, you should be able to do the following:

- Understand the purpose of hypothesis testing.
- Learn to construct the hypotheses.
- Know the situations for using one- and two-tailed tests.
- Describe the procedure of hypothesis testing.
- Understand the *p* value.
- Learn the computing procedure manually in different situations by using *t*-tests.
- Identify an appropriate *t*-test in different research situations.
- Know the assumptions under which *t*-test should be used.
- Describe the situations in which one-tailed and two-tailed tests should be used.
- Interpret the difference between one-tailed and two-tailed hypotheses.
- Learn to compute *t*-statistic in different research situations by using SPSS.
- Learn to interpret the outputs of different *t*-tests generated in SPSS.

Introduction

Human beings are progressive in nature. Most of our decisions in life are governed by our past experiences. These decisions may be subjective or objective. Subjective decisions are solely based upon one's own perception of viewing issues. These perceptions keep on changing from person to person. Same thing or situation can be perceived differently by different persons, and therefore, the decision cannot be universalized. On the other hand, if decisions are taken on the basis of scientific law, it is widely accepted and works well in the similar situations.

Decision makers are always engaged in identifying optimum decision in a given situation for solving a problem. Theory of statistical inference which is based on scientific principles provide optimum solution to these decision makers. Statistical inference includes theory of estimation and testing of hypothesis. In this chapter,

J.P. Verma, *Data Analysis in Management with SPSS Software*,
DOI 10.1007/978-81-322-0786-3_6,

different aspects of hypothesis testing regarding population parameter have been discussed. At times one may be interested to know as to whether the population mean is equal to the given value. In testing such hypothesis, a representative sample may be drawn to verify it by using statistical tests. Testing of hypothesis may be done for comparing two population averages on the basis of samples drawn from these populations. For instance, one may like to know whether memory retention power is more in girls or in boys in a particular age category or whether motivation level of employees in two different units of an organization is same or not. By comparing sample means of two groups, we intend to find out whether these samples come from the same population. In other words, we try to test whether their population means are equal or not. In this chapter, procedure of hypothesis testing in comparative studies has been discussed in detail. Besides comparative studies, the researcher may be interested to see the effect of certain treatment on dependent variable. Impact of advertisement campaign on the sale of a product, effect of training on employee's performance, and effect of stress management program on absenteeism are such examples where the posttesting mean may be compared with that of pretesting mean on the dependent variable.

For testing a hypothesis concerning mean, two statistical tests "t" and "z" are normally used. In case of small sample, t-test is used, whereas z-test is used in large-sample case. For all practical purposes, a sample is considered to be small if its size is less than or equal to 30 and large if it is more than 30. Since t-distribution approaches to z-distribution for $n > 30$, t-test can be considered as a specific case of z-test. The t-test is used in a situation where population is normally distributed and population variance is not known, and on the other hand Z test is used when population variance is known. In this chapter, only t-tests in different situations have been discussed.

Usually testing of hypothesis is done for population mean and variance as these are the two indices which are used to describe the nature of data to a great extent. In this chapter, only testing of hypothesis concerning mean in different situations has been discussed in great detail. Plan of choosing a statistical test for testing a hypothesis has been shown graphically in Fig. 6.1.

This chapter describes the procedure of testing a hypothesis concerning single group mean and the difference between two group means for unrelated and related groups.

Hypothesis Construction

Hypotheses are any assertion or statement about certain characteristics of the population. If the characteristics can be quantitatively measured by parameters such as mean or variance, then the hypothesis based on these parameters is said to be parametric. Whereas if the characteristics are qualitatively measured (e.g., assessment of quality, attitude, or perception), then the hypothesis so developed on these characteristics is known as nonparametric hypothesis. These parametric and

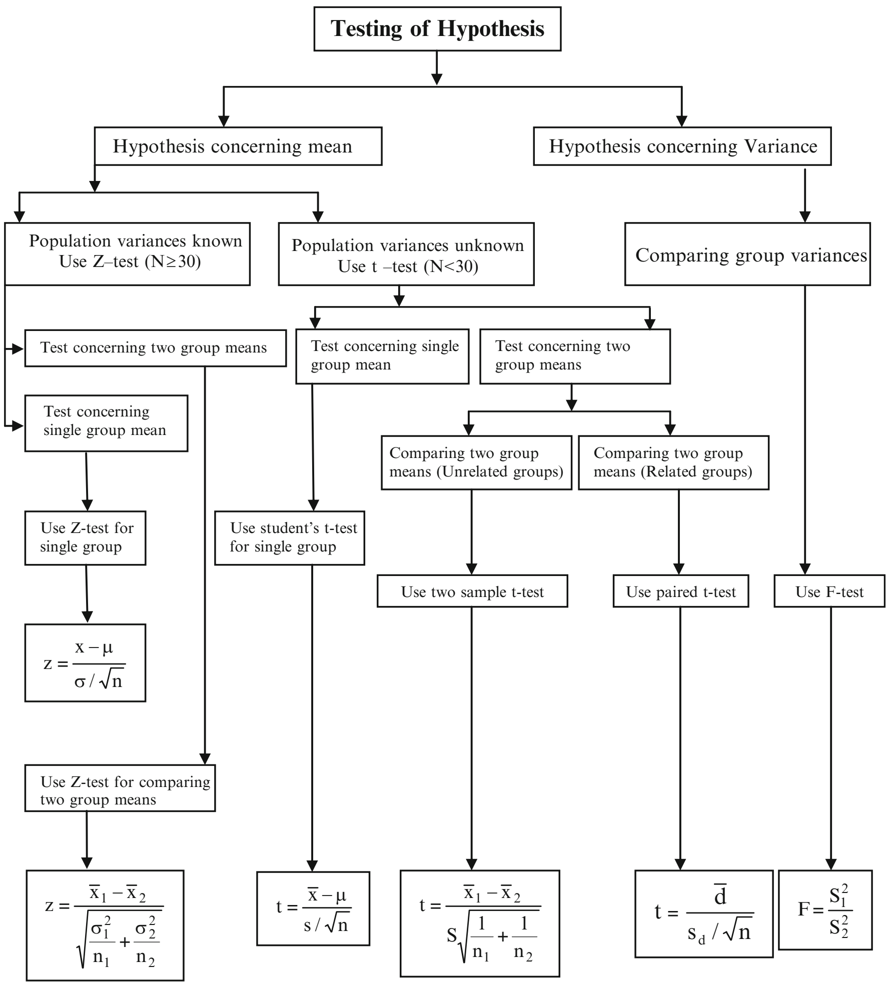


Fig. 6.1 Scheme of selecting test statistic in hypothesis testing

nonparametric hypotheses are known as statistical hypotheses. A hypothesis is said to be statistical hypothesis if the following three conditions prevail:

1. The population may be defined.
2. Sample may be drawn.
3. The sample may be evaluated to test the hypothesis.

Statistical hypotheses are based on the concept of proof by contradiction. For example, consider that a hypothesis concerning population mean (μ) is tested to see if an experiment has caused an increase or decrease in μ. This is done by proof of contradiction by formulating a null hypothesis. Thus, in testing of hypothesis, one needs to formulate a research hypothesis which is required to be tested for some population parameter. Based on research hypothesis, a null hypothesis is formulated. The null and the research (alternative) hypotheses are complementary to each other. In fact the null hypothesis serves as a means of testing the research hypothesis, and therefore, rejection of null hypothesis allows the researcher to accept the research hypothesis.

Null Hypothesis

Null hypothesis is a hypothesis of no difference. It is denoted by H_0. It is formulated to test an alternative hypothesis. Null hypothesis is assumed to be true. By assuming the null hypothesis to be true, the distribution of the test statistic can be well defined. Further, null signifies the unbiased approach of the researcher in testing the research hypothesis. The researcher verifies the null hypothesis by assuming that it is true and rejects it in favor of research hypothesis if any contradiction is observed. In fact the null hypothesis is made for rejection. In case if the null hypothesis cannot be rejected on the basis of the sample data, it is said that the researcher fails to reject the null hypothesis. The sole purpose of the researcher is to try rejecting the null hypothesis in favor of research hypothesis in case the contradiction is observed on the basis of the sample.

Alternative Hypothesis

Alternative hypothesis is also known as research hypothesis. In any research study, the researcher first develops a research hypothesis for testing some parameter of the population, and accordingly null hypothesis is formulated to verify it. The alternative hypothesis is denoted by $H_{1.}$ Alternative hypothesis means that there is a difference between the population parameter and the sample value. In testing of hypothesis, the whole focus is to test whether research hypothesis can be accepted or not, and this is done by contradicting the null hypothesis.

Test Statistic

In hypothesis testing, the decision about rejecting or not rejecting the null hypothesis depends upon the value of test statistic. A test statistic is a random variable X whose value is tested against the critical value to arrive at a decision.

If a random sample of size n is drawn from the normal population with mean, μ and variance, σ^2, then the sampling distribution of mean will also be normal with mean μ and variance σ^2/n. As per the central limit theorem even if the population from which the sample is drawn is not normal, the sample mean will still follow the normal distribution with mean, μ, and variance σ^2/n provided the sample size n is large ($n > 30$).

Thus, in case of large sample ($n > 30$), for testing the hypothesis concerning mean, z-test is used. However, in cases of small sample ($n \leq 30$), the distribution of sample mean follows t-distribution if the population variance is not known. In such situation, t-test is used. In case population standard deviation (σ) is unknown, it is estimated by the sample standard deviation(S). For different sample size, the t-curve is different, and it approaches to normal curve for sample size $n > 30$. All these curves are symmetrical and bell shaped and distributed around $t = 0$. The exact shape of the t-curve depends on the degrees of freedom.

In one-way ANOVA, the comparison between group variance and within-group variance is done by using the F-statistic. The critical value of F can be obtained from the Table A4 or A5 in appendix for a particular level of significance and the degrees of freedom between and within the groups.

Rejection Region

Rejection region is a part of the sample space in which if the value of test statistic falls, null hypothesis is rejected. Rejection region is also known as critical region. The value of the statistic in the distribution that divides sample space into acceptance and rejection region is known as critical value. These can be seen in Fig. 6.2.

The size of the rejection region is determined by the level of significance (α). The level of significance is that probability level below which we reject the null hypothesis. The term statistical significance of a statistic refers only to the rejection of a null hypothesis at some level α. It indicates that the observed difference between the sample mean and the mean of the sampling distribution did not occur by chance alone. So to conclude, if the test statistic falls in the rejection/critical region, H_0 is rejected, else H_0 is failed to be rejected.

Steps in Hypothesis Testing

In experimental research, the inferences are drawn on the basis of testing a hypothesis on population parameters. The following steps are involved in decision-making process:

1. Formulate the null and alternative hypothesis for each of the parameters to be investigated. It is important to mention as to whether the hypothesis required to be tested is one tailed or two tailed.
2. Choose the level of significance at which the hypothesis needs to be tested. Usually in experimental research, the significance levels are chosen as 0.01, 0.05, or 0.10, but any value between 0 and 1 can be used.
3. Identify the test statistic (follow the guidelines shown in Fig. 6.1) that can be used to test the null hypothesis, and compute its value on the basis of the sample data.
4. Obtain the tabulated value of the statistic from the designated table. Care must be taken to obtain its value as its values are different at the same level of significance for one-tailed and two-tailed hypotheses.

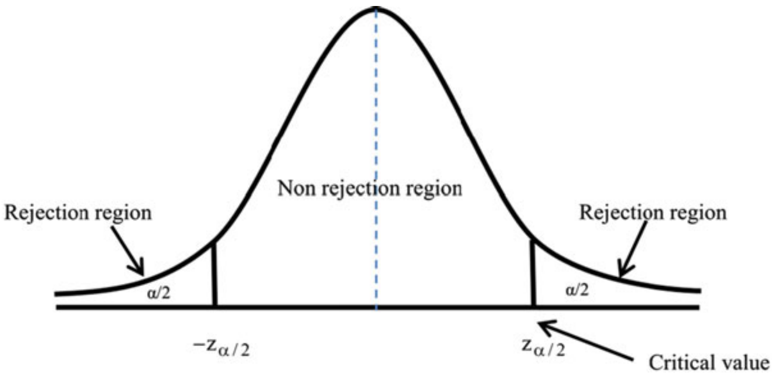


Fig. 6.2 Different regions in two-tailed test

5. If calculated value of statistic is greater than tabulated value, null hypothesis is rejected, and if the calculated value of statistic is less than or equal to its tabulated value, null hypothesis is failed to be rejected. SPSS output provides p value against the computed value of statistic. If the p value is less than .05, the statistic is said to be significant and the null hypothesis may be rejected at significance level of .05; on the other hand, if the p value is more than .05, one would fail to reject the null hypothesis. If the null hypothesis is failed to be rejected, one may state that there is not enough evidence to suggest the truth of the alternative hypothesis.

Type I and Type II Errors

We have seen that the research hypothesis is tested by means of testing the null hypothesis. Thus, the focus of the researcher is to find whether the null hypothesis can be rejected on the basis of the sample data or not. In testing the null hypothesis, the researcher has two options, that is, either to reject the null hypothesis or fail to reject the null hypothesis. Further, the true state of the null hypothesis may be true or false in either of these situations. Thus, the researcher has four courses of actions in testing the null hypothesis. The two actions, that is, rejecting the null hypothesis when it is false and fails to reject the null hypothesis when it is true, are correct decisions. Whereas the remaining two decisions, that is, rejecting the null hypothesis when it is true and fails to reject the null hypothesis when it is false, are the two wrong decisions. These two wrong decisions are known as two different kinds of errors in hypothesis testing. All the four courses of actions have been summarized in Table 6.1.

Thus, in hypothesis testing, a researcher is exposed to two types of errors known as type I and type II errors.

Type I error can be defined as rejecting the null hypothesis, H_0, when it is true. The probability of type I error is known as level of significance and is denoted by α. The choice of α determines the critical values. Looking to the relative importance of the decision, the researcher fixes the value of α. Normally the level of significance is chosen as .05 or .01.

Table 6.1 Decision options in testing of hypothesis

		True state of H_0	
		True	False
Researcher's decision about H_0	Reject H_0	Type I error	Correct decision
	Failed to reject H_0	Correct decision	Type II error

Type II error is said to be committed if we fail to reject the null hypothesis (H_0) when it is false. The probability of type II error is denoted by the Greek letter β and is used to determine the power of the test. The value of β depends on the way the null hypothesis is false. For example, in testing the null hypothesis of equal population means for a fixed sample size, the probability of type II error decreases as the difference between population means increases. The term $1 - \beta$ is said to be the power of test. The power of test is the probability of rejecting the null hypothesis when it is wrong.

Often type I and type II errors are confused with α and β, respectively. In fact α is not the type I error but it is the probability of type I error and similarly β is the probability of type II error and not the type II error. Since α is the probability, hence it can take any value in between 0 and 1, and one should write the statement like "null hypothesis may be rejected at .05 level of significance" instead of "null hypothesis may be rejected at 5% level of significance." Thus, the level of significance (α) should always be expressed in fractions such as .05 and .01, or it may be written as 5 or 1% level. For fixed sample size, the reduction of type I and type II errors simultaneously is not possible because if you try to minimize one error, the other error will increase. Therefore, there are two ways to reducing these two errors.

The *first approach* is to increase the sample size. This is not always possible in research studies because once the data is collected, the same has to be used by the researcher for drawing the inferences. Moreover, by increasing the sample size, a researcher loses the control over experiment, due to which these errors get elevated.

The *second approach* is to identify the error which is more severe, fix it up at a desired level, and then try to minimize the other error to a maximum possible extent. In most of the research studies, type I error is considered to be more severe because wrongly rejecting a correct hypothesis forces us to accept the wrong alternative hypothesis. For example, consider an experiment where it is desired to test the effectiveness of an advertisement campaign on the sales performance. The null hypothesis required to be tested in this case would be, "Advertisement campaign either do not have any impact on sales or may reduce the sales performance." Now if the null hypothesis is wrongly rejected, an organization would go for the said advertisement campaign which in fact is not effective. These decisions will unnecessarily enhance the budget expenditure without any further appreciation in the revenue modal. Severity of type I error can also be seen in the following legal analogy. Convicts are presumed to be innocent until unless they are proved to be guilty. The purpose of the trial is to see whether the null hypothesis of innocence can be rejected based on the evidences. Here the type I error (rejecting a correct null hypothesis) means convicting the innocence, whereas type II error (failing to reject the false null hypothesis) means letting the guilty go free. Here the type I error is more severe than type II error because no innocent should be punished in comparison to guilty may get no punishment. Type I error becomes more serious if the crime is murder and the person gets the punishment of death sentence. Thus, usually in research studies, the type I error is fixed at the desired level of, say, .05 or .01 and then type II error is tried to be minimized as much as possible.

The value of α and β depends upon each other. For a fixed sample size, the only way to reduce the probability of making one type of error is to increase the other.

Consider a situation where it is desired to compare the means of two populations. Let us assume that the rejection regions have critical values $\pm\infty$. Using the statistical test, H_0 will never get rejected as it will exclude every possible difference in sample means. Since the null hypothesis will never be rejected, the probability of rejecting the null hypothesis when it is true will be zero. In other words, the value of $\alpha = 0$. Since the null hypothesis will never be rejected, the probability of type II error (failing to reject the null hypothesis when it is false) will be 1 or to say that $\beta = 1$.

Now consider the rejection regions whose critical values are 0,0. In this case, the rejection region includes every possible difference in sample means. This test will always reject H_0. Since the null hypothesis will always be rejected, the probability of type I error (rejecting H_0 when it is true) will be 1 or the value of $\alpha = 1$. Since the null hypothesis is always rejected, the probability of type II error (failing to reject H_0 when it is false) is 0, or the value of $\beta = 0$.

To conclude, a statistical test having rejection region bounded by the critical values $\pm\infty$ has $\alpha = 0$ and $\beta = 1$, whereas the test with a rejection region bounded by the critical values 0,0 has $\alpha = 1$ and $\beta = 0$. Consider a test having rejection region bounded by the critical values $\pm q$. As q increases from 0 to ∞, α decreases from 1 to 0, while β increases from 0 to 1.

One-Tailed and Two-Tailed Tests

Consider an experiment in which null and alternative hypotheses are H_0 and $H_{1,}$ respectively. We perform a test to determine whether or not the null hypothesis should be rejected in favor of the alternative hypothesis. In this situation, two different kinds of tests can be performed. One may either use a *one-tailed test* to see whether there is an increase or decrease in the parameter or may decide to use a *two-tailed test* to verify for any change in the parameter that can be increased or decrease. The word tail refers to the far left and far right of a distribution curve. These one-tailed and two-tailed tests can be performed at any of the two, 0.01 or 0.05, levels of significance.

One-Tailed Test A statistical test is known as one-tailed test if the null hypothesis (H_0) is rejected only for the values of the test statistic falling into one specified tail of its sampling distribution. In one-tailed test, the direction is specified, that is, we are interested to verify whether population parameter is greater than some value. Or at times we may be interested to know whether the population parameter is less than some value. In other words, the researcher is clear as to what specifically he/she is interested to test. Depending upon the research hypothesis, one-tailed test can be classified as right-tailed or left-tailed tests. If the research hypothesis is to test whether

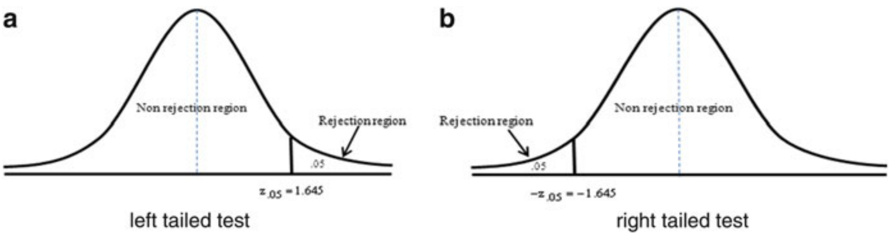
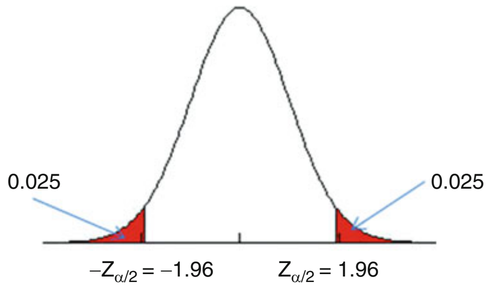


Fig. 6.3 Critical regions at 5% level in (**a**) left-tailed test and (**b**) right-tailed test

Fig. 6.4 Critical regions in two-tailed tests



the population mean is greater than some specified value, then the test is known as right-tailed test and the entire critical region shall lie in the right tail only. And if the test statistic falls into right critical region, the alternative hypothesis will be accepted instead of the null hypothesis. On the other hand, if the research hypothesis is to test whether the population mean is less than some specified value, then the test is known as left-tailed test and the entire critical region shall lie in the left tail only. The critical regions at 5% level in both these situations are shown in Fig. 6.3.

Two-Tailed Test A statistical test is said to be a two-tailed test if the null hypothesis (H_0) is rejected only for values of the test statistic falling into either tail of its sampling distribution. In two-tailed test, no direction is specified. We are only interested to test whether the population parameter is either greater than or less than some specified value. If the test statistic falls into either of the critical regions, the alternative hypothesis will be accepted instead of the null hypothesis. In two-tailed test, the critical region is divided in both the tails. For example, if the null hypothesis is tested at 5% level, the critical region shall be divided in both the tails as shown in Fig. 6.4. Tables A.1 and A.2 in Appendix provide critical values for *z*-test and *t*-test, respectively.

Criteria for Using One-Tailed and Two-Tailed Tests

A one-tailed test is used when we are quite sure about the direction of the difference in advance (e.g., exercise will improve the fitness level). With that assumption, the level of significance (α) is only calculated from one tail of the distribution. However, in standard testing, the probability is calculated from both tails.

For instance if the significance of correlation is tested between age and medical expenses; one might hypothesize that medical expenses may increase or do not increase but will never decrease with age. In such case one-tailed hypothesis should be used. On the other hand, in testing the correlation between people's weights with their income, we may not have reasons to believe that the income will increase with increase in weights or the income will decreases with weights. Here we might be interested just to find out if there was any relationship at all and that is a two-tailed hypothesis.

The issue in deciding between one-tailed and two-tailed tests is not whether or not you expect a difference to exist. Had you known whether or not there was a difference, there is no reason to collect the data. Instead, the question is whether the direction of a difference can only go one way. One should only use a one-tailed test if there is an absolute certainty before data collection that in the overall populations, either there is no difference or there is a difference in a specified direction. Further, if you end up showing a difference in the opposite direction, you should be ready to attribute that difference to random sampling without bothering about the fact that the measured difference might reflect a true difference in the overall populations. If a difference in the "wrong" direction brings even little meaning to your findings, you should use two-tailed test.

The advantage of using one-tailed hypothesis is that you can use a smaller sample to test it. The smaller sample often reduces your cost of the experiment. But on the other hand, it is easier to reject the null hypothesis with a one-tailed test in comparison to two-tailed test. Thus, the level of significance increases in one-tailed test. Because of this reason, it is rarely correct to perform a one-tailed test; usually we want to test whether any difference exists.

Strategy in Testing One-Tailed and Two-Tailed Tests

The strategy in choosing between one-tailed and two-tailed tests is to prefer a two-tailed test unless there is a strong belief that the difference in the population can only be in one direction. If the two-tailed test is statistically significant ($p < \alpha$), interpret the findings in one-tailed manner. Consider an experiment in which it is desired to test the null hypothesis that the average cure time of cold and cough by a newly introduced vitamin C tablet is 4 days against an alternative hypothesis that it is not. If a sample of 64 patients has an average recovery time of 3.5 days with $s = 1.0$ day, the p value in this testing would be 0.0002 and therefore the null hypothesis H_0 will be rejected and we accept the alternative hypothesis H_1. Thus, in this situation, it is concluded that the recovery time is not equal to 4 days for the new prescription of vitamin C.

But we may conclude more than that in saying that the recovery time is less than 4 days with the new prescription of vitamin C. We arrive at this conclusion by combining the two facts: Firstly, we have proved that the recovery time is different than 4 days, which means it must be either less or more than 4 days, and secondly, the sample mean $\bar{X}$ (=3.5 days) in this problem is less than the specified value, that is, 4 days(population mean). After combining these two facts, it may be concluded that the average recovery time (3.5 days) is significantly lower than the 4 days. This conclusion is quite logical because if we again test the null hypothesis $H_0{:}\mu \geq 4$ against the alternative hypothesis $H_1{:}\mu < 4$(one-tailed test), the *p* value would be 0.0001 which is even smaller than 0.0002.

Thus, we may conclude first by answering the original question then going for writing about the directional difference such as "The mean recovery time in cold and cough symptom with the new prescription of vitamin C is different from 4 days"; in fact, it is less than 4 days.

What Is *p* Value?

The *p* value is the probability of wrongly rejecting the null hypothesis. It is analogous to the level of significance. Usually an experimenter decides to test the hypothesis at some desired level of significance. If the absolute value of test statistic increases, the probability of rejecting the correct null hypothesis decreases. Thus, if a null hypothesis is tested at the level of significance .05 and the value of test statistic is large so that its corresponding *p* value is 0.004, in that case if we conclude that the null hypothesis is rejected at 5% level, it would not be logically correct as the error attached to this judgment is only 0.4%. In fact as the absolute value of test statistic increases, the *p* value keeps on decreasing.

One may decide the level of significance in advance say, 0.05, but while explaining the decision, the concept of *p* value should be used to report as to how much error is involved in the decision about rejecting or being unable to reject the null hypothesis. Thus, while testing a hypothesis, a *p* value is calculated against the test statistic which is used to explain the error involved in the decision. In SPSS and other statistical packages, the *p* values are automatically computed against each test statistic. Thus, if an experimenter decides to test the hypothesis at the significance level of 0.05, the test statistic shall be significant so long *p* value is less than 0.05. The general practice is to write the *p* value along with the value of test statistic. For instance, we may write as "Since the calculated $t = 4.0 (p = 0.0002)$ is significant, the null hypothesis may be rejected." The *p* value may be calculated against the value of *t*-statistic by using the *t*-table or by using the free conversion software available on many sites such as http://faculty.vassar.edu/lowry/tabs.html#t.

Degrees of Freedom

Any parameter can be estimated with certain amount of information or data set. The number of independent pieces of data or scores that are used to estimate a parameter is known as degrees of freedom and is usually abbreviated as df. In general, the degrees of freedom of an estimate are calculated as the number of independent scores that are required to estimate the parameter minus the number of parameters estimated as intermediate steps in the estimation of the parameter itself. In general, each item being estimated costs one degree of freedom.

The *degrees of freedom* can be defined as the number of independent scores or pieces of information that are free to vary in computing a statistic.

Since the variance σ^2 is estimated by the statistic S^2 which is computed from a random sample of n independent scores, let us see what the degrees of freedom of S^2 are. Since S is computed from the sample of n scores, its degrees of freedom would have been n, but because one degree of freedom is lost due to the condition that $\sum(X - \bar{X}) = 0$, the degrees of freedom for S^2 are $n - 1$. If we go by the definition, the degrees of freedom of S^2 are equal to the number of independent scores (n) minus the number of parameters estimated as intermediate steps (one, as μ is estimated by $\bar{X}$) and are therefore equal to $n - 1$.

In case of two samples, pooled standard deviation S is computed by using $n_1 + n_2$ observations. In the computation of S, the two parameters μ_1 and μ_2 are estimated by $\bar{X}_1$, and $\bar{X}_2$ hence, the two degrees of freedom are lost and therefore the degrees of freedom for estimating S are $n_1 + n_2 - 2$.

In computing chi-square in a 2×2 contingency table for testing the independence between rows and columns, it is assumed that you already know 3 pieces of information: the row proportions, the column proportions, and the total number of observations. Since the total number of pieces of information in the contingency table is 4, and 3 are already known before computing the chi-square statistic, the degrees of freedom are $4 - 3 = 1$. We know that the degrees of freedom for chi-square are obtained by $(r - 1) \times (c - 1)$; hence, with this formula, also the degrees of freedom in a 2×2 contingency table are 1.

One-Sample *t*-Test

A *t*-test can be defined as a statistical test used for testing of hypothesis in which the test statistic follows a Student's *t*-distribution under the assumption that the null hypothesis is true. This test is used if the population standard deviation is not known and the distribution of the population from which the sample has been drawn is normally distributed. Usually *t*-test is used for small sample size ($n < 30$) in a situation where population standard deviation is not known. Even if the sample is large ($n \geq 30$) but if the population standard deviation is not known in that situation, also *t*-test should be used instead of *z*-test. A one-sample *t*-test is used for testing whether the population mean is equal to a predefined value or not. An example of a one-sample *t*-test may be to see whether population average sleep time is equal to 5 h or not.

In using *t*-test, it is assumed that the distribution of data is approximately normal. The *t*-distribution depends on the sample size. Its parameter is called the degrees of freedom (df) which is equal to $n - 1$, where n is the sample size.

In one-sample test, *t*-statistic is computed by the following formula:

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}} \tag{6.1}$$

Calculated *t* is compared with tabulated *t* at 0.05 level of significance and $n - 1$ degrees of freedom if the hypothesis is to be tested at 5% level. The value of tabulated *t* can be obtained from Table A.2 in Appendix. The *t*-statistic is tested for its significance by finding its corresponding *p* value. If *p* value is less than .05, the *t*-statistic becomes significant, and we reject the null hypothesis against the alternative hypothesis. On the other hand, if the *p* value is more than .05, the null hypothesis is failed to be rejected.

Application of One-Sample Test

In the era of housing boom, everybody is interested to buy a home, and the role of banking institution is very important in this regard. Every bank tries to woe their clients by highlighting their specific features of housing loan like less assessment fee, quick sanctioning of the loans, and waving of penalty for prepayment. One particular bank was more interested to concentrate on loan processing time instead of other attributes and therefore made certain changes in their loan processing procedure without sacrificing the risk features so as to serve their clients with quick processing time. They want to test if their mean loan processing time differs from a competitor's claim of 4 h. The bank randomly selected a sample of few loan applications in their branches and noted the processing time for each cases. On the basis of this sample data, the authorities may be interested to test whether the bank's processing time in all their branches is equal to 4 h or not. One-sample *t*-test can provide the solution to test the hypothesis in this situation.

Example 6.1 A professor wishes to know if his statistics class has a good background of basic math. Ten students were randomly chosen from the class and were given a math proficiency test. Based on the previous experience, it was hypothesized that the average class performance on such math proficiency test cannot be less than 75. The professor wishes to know whether this hypothesis may be accepted or not. Test your hypothesis at 5% level assuming that the distribution of the population is normal. The scores obtained by the students are as follows:

Math proficiency score: 71, 60, 80, 73, 82, 65, 90, 87, 74, and 72

Solution The following steps shall show the procedure of applying the *t*-test for one sample in testing the hypothesis, whether the students of statistics class had their average score on math proficiency test equal to 75 or not.

(a) Here the hypothesis which needs to be tested is

$$\mathrm{H}_0 : \mu \geq 75$$

against the alternative hypothesis

$$\mathrm{H}_1 : \mu < 75$$

(b) *The level of significance:* 0.05
(c) *Statistical test:* As per the test selection scheme shown in Fig. 6.1, the test applicable in this example shall be one-sample *t*-test.

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

To compute the calculated value of *t*, firstly it is required to compute the value of mean and standard deviation of the sample:

X	X^2
71	5,041
60	3,600
80	6,400
73	5,329
82	6,724
65	4,225
90	8,100
87	7,569
74	5,476
72	5,184
$\sum X = 754$	$\sum X^2 = 57,648$

Since n = 10, $\bar{X} = \frac{754}{10} = 75.4$ and

$$S = \sqrt{\frac{1}{n-1}\sum X^2 - \frac{\left(\sum X\right)^2}{n(n-1)}}$$

$$= \sqrt{\frac{1}{9} \times 57648 - \frac{754^2}{10 \times 9}} = \sqrt{6405.33 - 6316.84}$$

$$= 9.41$$

After substituting the value of mean and standard deviation,

$$\text{Calculated} \quad \mathrm{t} = \frac{75.4 - 75}{9.41/\sqrt{10}} = \frac{0.4 \times \sqrt{10}}{9.41}$$

$$= 0.134$$

(d) *Decision criteria:* From Table A.2 in Appendix, the tabulated value of t for one-tailed test at .05 level of significance with 9 degrees of freedom is $t_{.05}(9) = 1.833$.
Since calculated $t(= 0.134) < t_{.05}(9)$, hence the null hypothesis is failed to be rejected at 5% level.
(e) *Inference:* Since the null hypothesis is failed to be rejected, hence the alternative hypothesis that the average math proficiency performance of the students is less than 75 cannot be accepted. Thus, it may be concluded that the average students' performance on math proficiency test is equal or higher than 75.

Two-Sample *t*-Test for Unrelated Groups

The two-sample t-test is used for testing the hypothesis of equality of means of two normally distributed populations. All t-tests are usually called *Student's t-tests*. But strictly speaking, this name should be used only if the variances of the two populations are also assumed to be equal. Two-sample t-test is based on the assumption that the variances of the populations σ_1^2 and σ_2^2 are unknown and population distributions are normal. In case the assumptions of equality of variances are not met, then the test used in such situation is called as Welch's t-test. Readers may read some other text for this test.

We often want to compare the means of two different populations, for example, comparing the effect of two different diets on weights, the effect of two teaching methodologies on the performance, or the IQ of boys and girls. In such situations, two-sample t-test can be used. One of the conditions of using two-sample t-test is that the samples are independent and identically distributed. Consider an experiment in which the job satisfaction needs to be compared among the bank employees working in rural and urban areas. Two randomly selected groups of 30 subjects each may be selected from rural and urban areas. Assuming all other conditions of the employees like salary structure, status and age categories to be similar, null hypothesis of no difference in their job satisfaction scores may be tested by using the two-sample t-test for independent samples. In this case, the two samples are independent because subjects in both the groups are not same.

Assumptions in Using Two-Sample t*-Test*

The following assumptions need to be fulfilled before using the two-sample t-test for independent groups:

- The distributions of both the populations from which the samples have been drawn are normally distributed.
- The variances of the two populations are nearly equal.

- Population variances are unknown.
- The samples are independent to each other.

Since we assume that σ_1^2 and σ_2^2 are equal, we can compute a pooled variance S^2 of both the samples. The purpose of pooling the variances is to obtain a better estimate. The pooled variance is a weighted sum of variances. Thus, if the sample sizes n_1 and n_2 are equal, then S^2 is just an average of the individual variances. The overall degrees of freedom in that case will be the sum of the individual degrees of freedom of the two samples, that is,

$$\mathrm{df} = \mathrm{df} + \mathrm{df}_2 = (\mathrm{n}_1 - 1) + (\mathrm{n}_2 - 1) = \mathrm{n}_1 + \mathrm{n}_2 - 2$$

Computation of *t*-statistic is same irrespective of testing two-tailed or one-tailed hypotheses. The only difference in testing these hypotheses are in its testing criteria and critical values of "*t*." These cases shall be discussed in the following sections.

Application of Two-Sampled t-*Test*

The situation where two-sample *t*-test is applicable can be easily understood by looking to the following case study. A pharmaceutical company decided to conduct an experiment to know as to whether high-protein diet or low-protein diets are more responsible in increasing the weights of male mouse in a controlled environment. Two groups of male mouse of similar age and weights may be selected randomly to serve as the two experimental groups. The number of mouse may be equal or unequal in both the groups. The first group may be fed with low-protein diet, whereas the other may be on the high-protein diet. To compare the average increase in their weights, two-sample *t*-test may be used to answer the research question.

Since one of the conditions of using the two-sample *t*-test is that the variance of the two groups must be equal, therefore *F*-test may be used to compare the variability. Only if the variability of the two groups is equal the two-sample *t*-test should be used. Here the null hypothesis of no difference in the increased weights of the high and low-protein groups is tested against the alternative hypothesis that the difference exists. In case the two-sample *t*-statistic is significant at some specified level of significance, the null hypothesis may be rejected, and it may be concluded that the effect of low-protein and high-protein diets on weights is different. On the other hand, if the *t*-statistic is not significant, we failed to reject the null hypothesis, and it may be concluded that it is not possible to find any significant difference in the rats' weight kept on high- and low-protein diets.

Further, if the null hypothesis is rejected, the mean weight of the high- and low-protein groups is seen, and if the average weight of the high-protein group is higher than that of the low-protein group, it may be concluded that the high-protein diet is more effective than the low-protein diet in increasing the weights of the rats.

Case 1: Two-Tailed Test

Since we have already discussed the general procedure of testing of hypothesis and the situations under which the two-tailed tests should be used, here the working method for two-tailed test shall be discussed. In two-tailed test, null hypothesis is tested against the alternative hypothesis that the groups are different in their means. Acceptance of alternative hypothesis suggests that the difference exists between the two group means. Further, by looking to the mean values of the two groups, one may draw the conclusion as to which group means is greater than the other. There may be many situations where two-tailed test can be used. For example, consider an experiment where it is desired to see the impact of different kinds of music on the hours of sleep. The two groups of the subjects are randomly selected, and the first group is exposed to classical music, whereas the second group is exposed to Jazz music for 1 h before sleep for a week. To test whether average sleep hour remains same or different in two different kinds of music groups, a two-tailed test may be used. Here it is not known that a particular music may increase the sleep hour or not, and hence, two-tailed test would be appropriate. In case of two-tailed test, the testing protocol is as follows:

(a) *Hypotheses need to be tested*

$$\mathrm{H}_0 : \mu_1 = \mu_2$$
$$\mathrm{H}_1 : \mu_1 \neq \mu_2$$

(b) *Test statistic*

$$\text{Calculated} \quad t = \frac{\bar{X}_1 - \bar{X}_2}{S\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{6.2}$$

where $S = \sqrt{\dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$

(c) *Degrees of freedom* $n_1 + n_2 - 2$

(d) *Decision criteria*

In two-tailed test, the critical region is divided in both the tails. If the level of significance is α, then the area in each tail would be $\alpha/2$. If the critical value is $t_{\alpha/2}$ and

if calculated $|t| \leqslant t_{\alpha/2}$, H_0 is failed to be rejected at α level of significance

and if calculated $|t| > t_{\alpha/2}$, H_0 may be rejected at α level of significance

Note: The value of calculated t is taken as absolute because the difference in the two means may be positive or negative.

Case II: Right-Tailed Test (One-Tailed Test)

We have already discussed the situations in which one-tailed test should be used. One-tailed test should only be used if an experimenter, on the basis of past information, is absolutely sure that the difference can go only in one direction. One-tailed test can be either right tailed or left tailed. In right-tailed test, it is desired to test the hypothesis, whether mean of first group is greater than that of the mean of the second group. In other words, the researcher is interested in a particular group only. In such testing, if the null hypothesis is rejected, it can be concluded that the first group mean is significantly higher than that of the second group mean. The situation where right-tailed test can be used is to test whether frustration level is less among those employees whose jobs are linked with incentives in comparison to those whose jobs are not linked with the incentives. Here the first group is the one whose jobs are linked with the incentives, whereas the second group's jobs are not linked with the incentives. In this situation, it is assumed that the employees feel happy in their jobs if it is linked with incentives. The testing protocol in testing the right-tailed hypothesis is as follows:

(a) *Hypotheses need to be tested*

$$\mathrm{H}_0 : \mu_1 \leqslant \mu_2$$
$$\mathrm{H}_1 : \mu_1 > \mu_2$$

(b) *Test statistic*

$$\text{Calculated} \quad t = \frac{\bar{X}_1 - \bar{X}_2}{S\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $S = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$

(c) *Degrees of freedom* $n_1 + n_2 - 2$

(d) *Decision criteria*

In one-tailed test, the entire critical region lies in one tail only. Here the research hypothesis is the right tailed; hence, the entire critical region would lie in the right tail only, and therefore, the sign of the critical value would be positive. If the critical value is represented by t_α and

if calculated $t \leqslant t_\alpha$, H_0 is failed to be rejected at α level of significance

and if calculated $t > t_\alpha$, H_0 may be rejected at α level of significance

Case III: Left-Tailed Test (One-Tailed Test)

At times the researcher is interested in testing whether a particular group mean is less than the second one. In this type of hypothesis testing, it is desired to test whether mean of first group is less than that of mean of the second group. Here if the null hypothesis is rejected, it can be concluded that the first group mean is significantly smaller than that of the second group mean. Consider a situation where an exercise therapist is interested to know whether a 4-week weight reduction program is effective or not if implemented on the housewives. The two groups consisting 20 women each are selected for the study, and the first group is exposed to the weight reduction program, whereas the second group serves as a control and does not take part in any special activities except daily normal work. If the therapist is interested to know whether on an average first treatment group shows the reduction in their weight in comparison to those who did not participate in the program, the left-tailed test may be used. In this situation, as per the experience, it is known that any weight reduction program will always reduce the weight in general in comparison to those who do not participate in it, and therefore, one tailed test would be appropriate in this situation. The testing protocol in applying the left-tailed test is as follows:

(a) *Hypotheses need to be tested*

$$\begin{aligned} H_0 &: \mu_1 \geq \mu_2 \\ H_1 &: \mu_1 < \mu_2 \end{aligned}$$

(b) *Test statistic*

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

where $$S = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

(c) *Degrees of freedom* $n_1 + n_2 - 2$

(d) *Decision criteria*

In one-tailed test, the entire critical region lies in one tail. Since this is a case of left-tailed test, hence the entire critical region lies in the left tail only and therefore the critical value would be negative. If the critical value is represented by $-t_\alpha$ and

if calculated $t \geq -t_\alpha$, H_0 is failed to be rejected at α level of significance

and if calculated $< -t_\alpha$, H_0 may be rejected at α level of significance

Example 6.2 Counseling cell of a college keeps conducting sessions with the problematic students by using different methods. Since the number of visitors keeps increasing every day in the center, they have decided to test whether audiovisual-based counseling and personal counseling are equally effective in reducing the stress level. Eighteen women students were randomly chosen among those who visited the center. Nine of them were given the personal counseling, whereas the other nine were given the sessions with the audiovisual presentation. After the session, the students were tested for their stress level. The data so obtained are shown in Table 6.2.

Table 6.2 Data on stress level for the students in both the counseling groups

Personal counseling:	27	22	28	21	23	22	20	31	26
Audiovisual counseling:	35	28	24	28	31	32	33	34	30

Test your hypothesis at 1% level, whether any one method of counseling is better than other. It is assumed that population variances are equal and both the populations are normally distributed.

Solution To test the required hypothesis, the following steps shall explain the procedure.

(a) *Here the hypothesis which needs to be tested is*

$$\mathrm{H}_0 : \mu_{\text{Personal}} = \mu_{\text{Audi-visual}}$$

against the alternative hypothesis

$$\mathrm{H}_1 : \mu_{\text{Personal}} \neq \mu_{\text{Audi-visual}}$$

(b) *The level of significance:* 0.05

(c) *Statistical test:* In this example, it is required to test a two-tailed hypothesis for comparing the means of two groups. Thus, as per the scheme of the test selection shown in Fig. 6.1, a two-sample *t*-test for independent groups shall be appropriate in this case which is given by

$$t = \frac{\bar{X} - \bar{Y}}{S\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where the pooled standard deviation S is computed as

$$S = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

In order to compute the value of *t*-statistic, the mean and standard deviation of both the groups along with the pooled standard deviation S will have to be computed first (Table 6.3).

Since $n_1 = n_2 = 9$, $\bar{X} = \frac{220}{9} = 24.44$ and $\bar{Y} = \frac{275}{9} = 30.56$

$$\begin{aligned} S_x &= \sqrt{\frac{1}{n_1 - 1}\sum X^2 - \frac{\left(\sum X\right)^2}{n_1(n_1 - 1)}} \\ &= \sqrt{\frac{1}{8} \times 5488 - \frac{(220)^2}{9 \times 8}} = \sqrt{686 - 672.22} \\ &= 3.71 \end{aligned}$$

Table 6.3 Computation for mean and standard deviation

Personal counseling		Audiovisual counseling	
X	X^2	Y	Y^2
27	729	35	1,225
22	484	28	784
28	784	24	576
21	441	28	784
23	529	31	961
22	484	32	1,024
20	400	33	1,089
31	961	34	1,156
26	676	30	900
$\sum X = 220$	$\sum X^2 = 5,488$	$\sum Y = 275$	$\sum Y^2 = 8,499$

Similarly

$$S_Y = \sqrt{\frac{1}{n_2 - 1}\sum Y^2 - \frac{(\sum Y)^2}{n_2(n_2 - 1)}}$$

$$= \sqrt{\frac{1}{8} \times 8499 - \frac{(275)^2}{9 \times 8}} = \sqrt{1062.38 - 1050.35}$$

$$= 3.47$$

Further, pooled standard deviation S is equal to

$$S = \sqrt{\frac{(n_1 - 1)S_X^2 + (n_2 - 1)S_Y^2}{n_1 + n_2 - 2}} = \sqrt{\frac{8 \times 3.71^2 + 8 \times 3.47^2}{9 + 9 - 2}}$$

$$= \sqrt{\frac{110.11 + 96.33}{16}} = 3.59$$

One of the conditions of using the two-sample t-test for independent groups is that the variance of the two populations must be same. This hypothesis can be tested by using the F-test.

Thus, $$F = \frac{S_X^2}{S_Y^2} = \frac{3.71^2}{3.47^2} = 1.14$$

From Table A.4 in Appendix, tabulated $F_{.05}(8,8) = 3.44$

Since calculated value of F is less than the tabulated F, hence it may not be concluded that the variances of the two groups are different, and therefore, two-sample t-test for two independent samples can be applied in this example.

Remark: In computing F-statistic, the larger variance must be kept in the numerator, whereas the smaller one should be in the denominator.

After substituting the values of $\bar{X}$, $\bar{Y}$, and pooled standard deviation S, we get

$$\text{calculated t} = \frac{\bar{X} - \bar{Y}}{S\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{24.44 - 30.56}{3.59\sqrt{\left(\frac{1}{9} + \frac{1}{9}\right)}}$$

$$= -\frac{6.12}{1.69}$$

$$= -3.62$$

$$\Rightarrow \quad \text{calculated} \quad |t| = 3.62$$

(d) *Decision criteria:* From Table A.2 in Appendix, the tabulated value of t for two-tailed test at .05 level of significance with 16(=$n_1 + n_2 - 2$) degrees of freedom is $t_{.05}(16) = 2.12$.
Since calculated $t(= 3.62) > t_{.05}(16)$, the null hypothesis may be rejected at 5% level against the alternative hypothesis.
Further, since the mean stress score of the personal counseling group is lower than that of the audiovisual group, hence it may be concluded that the stress score of the personal counseling group is significantly less than that of the audiovisual group.
(e) *Inference:* Since the null hypothesis is rejected, hence the alternative hypothesis that the average stress scores of the personal counseling group as well as audiovisual counseling groups are not same is accepted. Further, since the mean stress score of the personal counseling group is significantly lower than that of the audiovisual group, it may be concluded that the personal counseling is more effective in comparison to that of the audiovisual counseling in reducing stress among women.

Example 6.3 A researcher wishes to know whether girls' marriage age in metro cities is higher than that of class B cities. Twelve families from metro cities and 11 families from class B cities were randomly chosen and were asked about their daughter's age at which they got married. The data so obtained are shown in Table 6.4. Can it be concluded from the given data that the girls' marriage age was higher in metro cities in comparison to class B cities? Test your hypothesis at 5% level assuming that the population variances are equal and the distribution of both the populations from which the samples have been drawn are normally distributed.

Solution In order to test the hypothesis, the following steps shall be performed:

(a) *The hypothesis which needs to be tested is*

Table 6.4 Marital age of the girls

Metro city:	29,	28,	27,	31,	32,	25,	28,	24,	27,	30,	35,	26
Class B city:	28,	25,	24,	28,	22,	24,	23,	21,	25,	24,	28	

$$\mathrm{H}_0 : \mu_{\text{Metro\_City}} \leqslant \mu_{\text{Class\_B\_City}}$$

against the alternative hypothesis

$$\mathrm{H}_0 : \mu_{\text{Metro\_City}} > \mu_{\text{Class\_B\_City}}$$

(b) *The level of significance:* 0.05

(c) *Statistical test:* In this example, it is required to test one-tailed hypothesis for comparing means of the two groups. Thus, as per the scheme of the test selection shown in Fig. 6.1, a two-sample t-test for independent groups shall be appropriate in this case which is

$$t = \frac{\bar{X} - \bar{Y}}{S\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where the pooled standard deviation S is given by

$$S = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

To compute the value of t statistic, the mean and standard deviation of both the groups along with the pooled standard deviation S need to be computed first (Table 6.5).

Here $n_1 = 12$ and $n_2 = 11$ $\bar{X} = \frac{342}{12} = 28.5$ and $\bar{Y} = \frac{272}{11} = 24.73$

$$\begin{aligned} S_X &= \sqrt{\frac{1}{n_1 - 1}\sum X^2 - \frac{\left(\sum X\right)^2}{n_1(n_1 - 1)}} \\ &= \sqrt{\frac{1}{11} \times 9854 - \frac{(342)^2}{12 \times 11}} = \sqrt{895.82 - 886.09} \\ &= 3.12 \end{aligned}$$

Table 6.5 Computation for mean and standard deviation

Metro city		Class B city	
X	X^2	Y	Y^2
29	841	28	784
28	784	25	625
27	729	24	576
31	961	28	784
32	1,024	22	484
25	625	24	576
28	784	23	529
24	576	21	441
27	729	25	625
30	900	24	576
35	1,225	28	784
26	676		
$\sum X = 342$	$\sum X^2 = 9,854$	$\sum Y = 272$	$\sum Y^2 = 6,784$

Similarly

$$S_Y = \sqrt{\frac{1}{n_2 - 1}\sum Y^2 - \frac{\left(\sum Y\right)^2}{n_2(n_2 - 1)}}$$
$$= \sqrt{\frac{1}{10} \times 6784 - \frac{(272)^2}{11 \times 10}} = \sqrt{678.4 - 672.58}$$
$$= 2.41$$

The pooled standard deviation S is equal to

$$S = \sqrt{\frac{(n_1 - 1)S_X^2 + (n_2 - 1)S_Y^2}{n_1 + n_2 - 2}} = \sqrt{\frac{11 \times 3.12^2 + 10 \times 2.41^2}{12 + 11 - 2}}$$
$$= \sqrt{\frac{107.08 + 58.08}{21}} = 2.80$$

Since t-test can only be applied if the variance of both the populations is same, this hypothesis can be tested by using the F-test.

Thus, $F = \frac{S_X^2}{S_Y^2} = \frac{3.12^2}{2.41^2} = 1.67$

The tabulated value of F can be seen from Table A.4 in Appendix.

Thus, tabulated $F_{.05}(11,10) = 2.85$

Since calculated value of F is less than that of tabulated F, hence hypothesis of equality of variances in two groups may not be rejected, and therefore, the two-sample t-test for independent samples can be applied in this example.

After substituting the values of $\bar{X}$, $\bar{Y}$, and pooled standard deviation, S we get

$$\text{calculated t} = \frac{\bar{X} - \bar{Y}}{S\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{28.5 - 24.73}{2.80\sqrt{\left(\frac{1}{12} + \frac{1}{11}\right)}}$$
$$= \frac{3.77}{2.80 \times 0.42}$$
$$= 3.21$$

(d) *Decision criteria:* From Table A.2 in Appendix, the tabulated value of t for one-tailed test at .05 level of significance with $21(=n_1 + n_2 - 2)$ degrees of freedom is $t_{.05}(21) = 1.721$. Similarly for one-tailed test, tabulated value of t at .01 level of significance is $t_{.01}(21) = 2.518$.
Since calculated $t(= 3.21) > t_{.05}(21)$, the null hypothesis may be rejected at 5% level. Further, calculated value of t is also less than that of tabulated value of t at 1% level as well; hence, t-value is also significant at 1% level.

(e) *Inference:* Since the null hypothesis is rejected, hence the alternative hypothesis that the marriage age of the girls in metro cities is higher than that of class B cities is accepted. It may thus be concluded that girls in metro cities prefers to marry late in age in comparison to that of class B cities.

Paired *t*-Test for Related Groups

Paired t-test is used to test the null hypothesis that the difference between the two responses measured on the same experimental units has a mean value of zero. This statistical test is normally used to test the research hypothesis as to whether the posttreatment response is better than the pretreatment response. Paired t-test is used in all those situations where there is only one experimental group and no control group. The question which is tested here is to know whether the treatment is effective or not. This is done by measuring the responses of the subjects in the experimental group before and after the treatment. There can be several instances in which the paired t-test may be used. Such situations may be, for instance, to see the effectiveness of management development program on the functional efficiency, effectiveness of the weight training program in weight reduction, effectiveness of the psychological training in enhancing memory retention power, etc.

The paired t-test is also known as "repeated measures" t-test. In using the paired t-test, the data must be obtained in pair on the same set of subjects before and after the experiment.

While applying the paired t-test for two related groups, the pairwise differences, d_i, is computed for all n paired data. The mean, $\bar{d}$ and standard deviation, S_d, of the differences d_i are calculated. Thus, paired t-statistic is computed as follows:

$$t = \frac{\bar{d}}{S_d/\sqrt{n}} \tag{6.3}$$

where "t" follows the Student's t-distribution with $n - 1$ degrees of freedom.

An assumption in using the paired t-test is that the difference d_i follows the normal distribution. An experiment where paired difference is computed is often more powerful, since it can eliminate differences in the samples that increase the total variance σ^2. When the comparison is made between groups (of similar experimental units), it is called blocking. The paired difference experiment is an example of a randomized block experiment.

Note: The blocking has to be done before the experiment is performed.

Assumptions in Using Paired t*-Test*

While using the paired t-test, the following assumptions need to be satisfied:

1. The distribution of the population is normal.
2. The distribution of scores obtained by pairwise difference is normal, and the differences are a random sample.
3. Cases must be independent of each other.

Remark: If the normality assumption is not fulfilled, you may use the nonparametric Wilcoxon sign rank test for paired difference designs.

Testing Protocol in Using Paired t*-Test*

Testing protocol of using paired t-test is similar to that of two-sample t-test for independent groups discussed above. In applying paired t-test, the only difference is that the test statistic is

$$t = \frac{\bar{d}}{S_d/\sqrt{n}}$$

instead of the one used in two-sample t-test. Further, in paired t-test, the degrees of freedom are $n - 1$. While using paired t-test, one should normally construct the two-tailed test first, and if the difference is significant, then by looking to the values of the samples mean of the pre- and posttesting responses, one may interpret as to which group mean is higher than the other. In general using one-tailed test should be avoided until there is strong evidence that the difference can go only in one direction. In one-tailed test, the probability of rejecting the correct null hypothesis becomes more in comparison to two-tailed test for the same level of significance.

Table 6.6 Calorie intake of the women participants before and after the nutrition educative program

Before:	2,900	2,850	2,950	2,800	2,700	2,850	2,400	2,200	2,650	2,500	2,450	2,650
After:	2,800	2,750	2,800	2,800	2,750	2,800	2,450	2,250	2,550	2,450	2,400	2,500

The trade-off using one- and two-tailed tests has been discussed in details while discussing the criteria for using one-tailed and two-tailed tests earlier in this chapter.

Application of Paired *t*-Test

The application of paired *t*-test can be understood by considering the following situation. An herbal company has come out with a drug useful for lowering the cholesterol level if taken for a week. In order to claim its effectiveness, it has been decided to administer the drug on the patients with high cholesterol level. In this situation, the paired *t*-test can be used to test the hypothesis to know as to whether there is any difference in the cholesterol level between post- and pretest data after administering this new drug. In this situation, the null hypothesis of no difference may be tested to test the two-tailed hypothesis first. If the *t*-statistic is significant, we reject the null hypothesis in favor of the alternative, and it is concluded that the difference exists between the cholesterol levels of the patients before and after the administration of the drug. On the other hand, if the *t*-statistic is not significant, we may fail to reject the null hypothesis and we may end up in concluding that the claim of the drug being effective may not be proved with this sample. After having rejected the null hypothesis, by looking to the average cholesterol level of the patients before and after the administration of the drugs, one may conclude as to whether the drug is effective or not.

Example 6.4 Twelve women participated in a nutritional educative program. Their calorie intake, before and after the program, was measured which are shown in Table 6.6.

Can you draw the conclusion that the nutritional educative program was successful in reducing the participant's calorie requirements? Test your hypothesis at 5% level assuming that the differences of the scores are normally distributed.

Solution In this example, data is paired. In other words, post- and pretest data belongs to the same person, and therefore, the groups may be called as related or paired. To test a hypothesis as to whether the nutritional educative program is effective or not in reducing the calorie intake, the following steps shall be performed:

(a) *Here the hypothesis which needs to be tested is*

$H_0 : \mu_D = 0$(Difference of means of the two groups is zero.)
against the alternative hypothesis
$H_1 : \mu_D \neq 0$(Difference of means of the two groups is not equal to zero.)

(b) *The level of significance:* 0.05

(c) *Statistical test:* In this example, it is required to test the effectiveness of the nutritional educative program in reducing the calorie consumption in diet. Here the null hypothesis that there is no difference in the means of the two groups is to be tested against the alternative hypothesis that there is a difference. Once the null hypothesis is rejected, then on the basis of mean values of the pre- and posttesting data of calorie consumption, the conclusion would be drawn as to whether the program was effective or not.

Thus, first a two-tailed test will be used, and if the null hypothesis is rejected against the alternative hypothesis, then the directional interpretation would be made by looking to the mean values. It is because of the fact that if the t-statistic is significant in two-tailed test, then it will also be significant at one-tailed test. This can be understood like this: for the two-tailed test, the critical value $t_{\alpha/2}$ at α level of significance will always be greater than that of the critical value t_α in one-tailed test. And therefore, if the calculated value of t is greater than $t_{\alpha/2}$, it will also be greater than t_α.

Since this example is a case of paired samples, hence as per the scheme of the test selection shown in Fig. 6.1, the paired t-test for related groups shall be appropriate in this case which is given by

$$t = \frac{\bar{d}}{S_d/\sqrt{n}}$$

where $\bar{d}$ is the mean of the difference between X and Y, and S_d is the standard deviation of these differences as given by

$$S_d = \sqrt{\frac{1}{n-1}\sum d^2 - \frac{(\sum d)^2}{n(n-1)}}$$

The $\bar{d}$ and S_d shall be computed first to find the value of t-statistic (Table 6.7).

Remark: Difference d can be computed by subtracting postdata from predata or vice versa because two-tailed test is being used here

Here number of paired data, $n = 12$,

$$\bar{d} = \frac{600}{12} = 50$$

and

Table 6.7 Computation for $\bar{d}$and S_d for paired t-ratio

Before X	After Y	$d = X - Y$	d^2
2,900	2,800	100	10,000
2,850	2,750	100	10,000
2,950	2,800	150	22,500
2,800	2,800	0	0
2,700	2,750	−50	2,500
2,850	2,800	50	2,500
2,400	2,450	−50	2,500
2,200	2,250	−50	2,500
2,650	2,550	100	10,000
2,500	2,450	50	2,500
2,450	2,400	50	2,500
2,650	2,500	150	22,500
		$\sum d = 600$	$\sum d^2 = 90,000$

$$
\begin{aligned}
S_d &= \sqrt{\frac{1}{n-1}\sum d^2 - \frac{\left(\sum d\right)^2}{n(n-1)}} = \sqrt{\frac{1}{11} \times 90000 - \frac{(600)^2}{12 \times 11}} \\
&= \sqrt{8181.82 - 2727.27} \\
&= 73.85
\end{aligned}
$$

After substituting the values of $\bar{d}$ and S_d, we get

$$
\begin{aligned}
\text{calculated t} &= \frac{\bar{d}}{S_d/\sqrt{n}} = \frac{50}{73.85/\sqrt{12}} \\
&= \frac{50 \times \sqrt{12}}{73.85} \\
&= 2.345
\end{aligned}
$$

(d) *Decision criteria:* From Table A.2 in Appendix, the tabulated value of t for two-tailed test at .05 level of significance with 11(=$n - 1$) degrees of freedom is $t_{.05/2}(11) = 2.201$.
Since calculated $t(= 2.345) > t_{.05/2}(11)$, the null hypothesis may be rejected at 5% level against the alternative hypothesis. It may therefore be concluded that the mean calorie intakes before and after the nutritional educative program are not same.
Since the mean calorie intake of the after-testing group is lower than that of the before-testing group, it may be concluded that the mean calorie score of the after-testing group is significantly less than that of the before-testing group.

(e) *Inference:* It is therefore concluded that the nutritional educative program is effective in reducing the calorie intake among the participants.

Solved Example of Testing Single Group Mean with SPSS

Example 6.5 The age of the 15 randomly chosen employees of an organization is shown in Table 6.8. Can it be concluded that the average age of the employees in the organization is 28 years? Test your hypothesis at 5% level and interpret your findings.

Solution The hypothesis that needs to be tested here is

$$H_0 : \mu = 28$$

against the alternative hypothesis

$$H_1 : \mu \neq 28$$

After using the SPSS commands as mentioned below for testing the population mean to be equal to 28, the output will generate the value of t-statistic along with its p value. If p value is less than .05, then the t-statistic will be significant and the null hypothesis shall be rejected at 5% level in favor of alternative hypothesis; otherwise, we would fail to reject the null hypothesis.

Computation of t*-Statistic and Related Outputs*

(a) ***Preparing Data File***

Before using the SPSS commands for computing the value of t-statistic and other related statistics for single group, a data file needs to be prepared. The following steps will help you to prepare the data file:

Table 6.8 Data on age

27
31
34
29
28
34
33
36
26
28
29
36
35
31
36

(i) *Starting the SPSS:* Use the following command sequence to start SPSS:
Start → All Programs → IBM SPSS Statistics → IBM SPSS Statistics 20
After checking the option **Type in Data** on the screen you will be taken to the **Variable View** option for defining the variables in the study.

(ii) *Defining variables:* There is only one variable in this example which needs to be defined in SPSS along with its properties. Since the variable is measured on interval scale, hence will be defined as 'Scale' variable. The procedure of defining the variable in the SPSS is as follows:

1. Click **Variable View** to define the variable and its properties.
2. Write short name of the variable as *Age* under the column heading **Name.**
3. Under the column heading **Label**, define full name of the variable as *Employees' Age*.
4. Under the column heading **Measure**, select the option 'Scale' for the variable.
5. Use default entries in all other columns.

After defining the variable in variable view, the screen shall look like Fig. 6.5.

(iii) *Entering data:* After defining the variable in the **Variable View**, click **Data View** on the left bottom of the screen to enter the data. Enter the data for the variable column wise. After entering the data, the screen will look like Fig. 6.6. Save the data file in the desired location before further processing.

(b) **SPSS *Commands for Computing t-Statistic***
After entering the data in the data view, follow these steps for computing *t*-statistic:

(i) *Initiating the SPSS commands for one-sample t-test:* In data view, go to the following commands in sequence:

Analyze ⇨ Compare Means ⇨ One-Sample *t* Test

The screen shall look like Fig. 6.7 as shown below.

(ii) *Selecting variables for t-statistic:* After clicking the **One-Sample *t* Test** option you will be taken to the next screen for selecting variable for computing *t*-statistic. Select the variable *Age* from left panel and bring it to the right panel by clicking the arrow sign. In case of computing *t*-value for more than one variable simultaneously, all the variables can be selected together. The screen shall look like Fig. 6.8.

(iii) *Selecting the options for computation:* After selecting the variable, option needs to be defined for the one-sample *t*-test. Do the following:

– In the screen shown in Fig. 6.8, enter the 'test value' as 28. This is the assumed population mean age that we need to verify in the hypothesis.
– Click the tag **Options**, you will get the screen shown in Fig. 6.9. Enter the confidence interval as 95% and click ***Continue*** and then you will be taken back to the screen shown in Fig. 6.8.

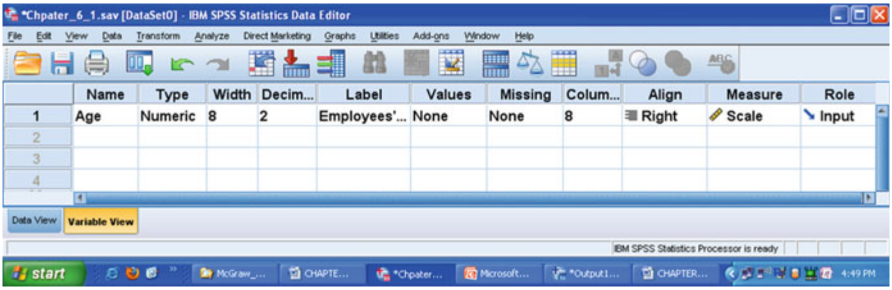
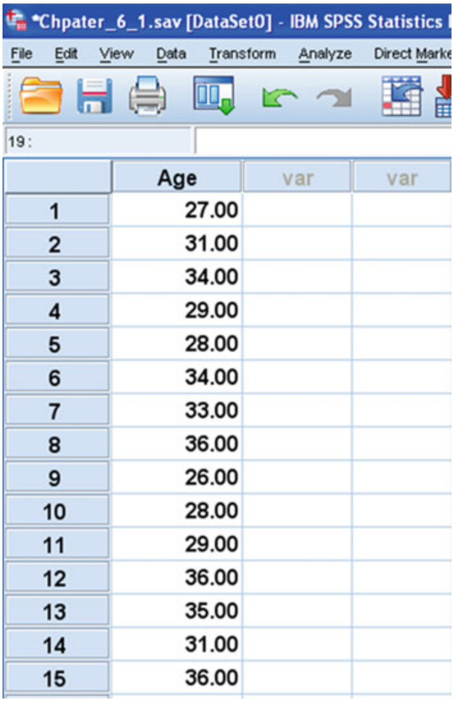


Fig. 6.5 Defining variable and its characteristics for the data shown in Table 6.8

Fig. 6.6 Screen showing entered data for the age variable in the data view



The confidence interval is chosen to get the confidence limits of mean based on sample data. Since in this example the null hypothesis needs to be tested at 5% level, choose the confidence interval as 95%.

– Click ***OK***.

(c) ***Getting the Output***

After clicking the **OK** tag in Fig. 6.8, you will get the output window. In the output window, the relevant outputs can be selected by using the right click of

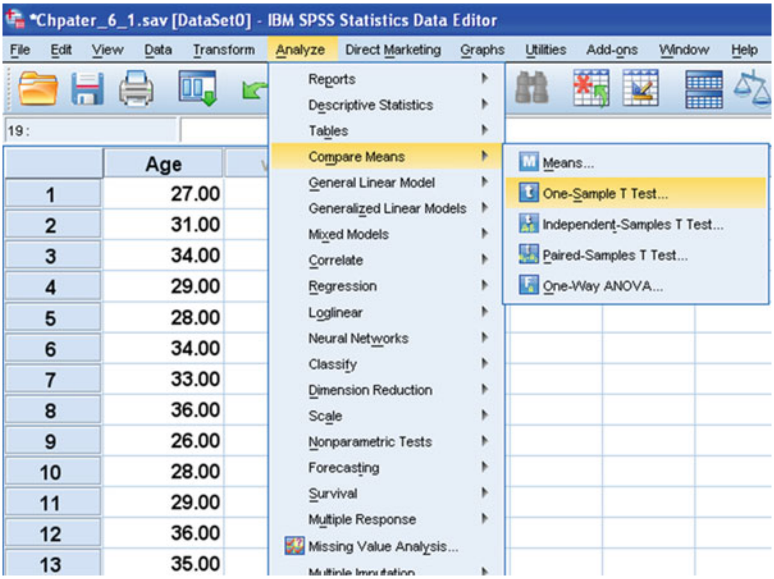


Fig. 6.7 Screen showing SPSS commands for one sample *t*-test

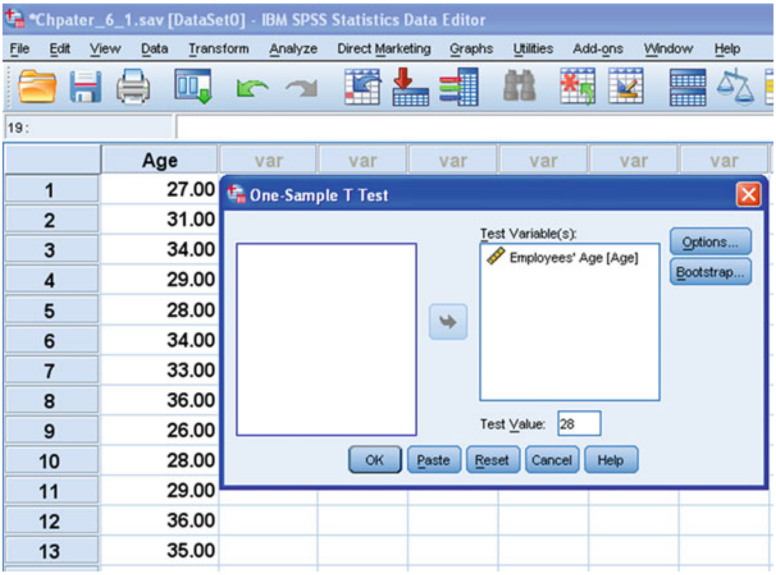


Fig. 6.8 Screen showing selection of variable for one-sample *t*-test

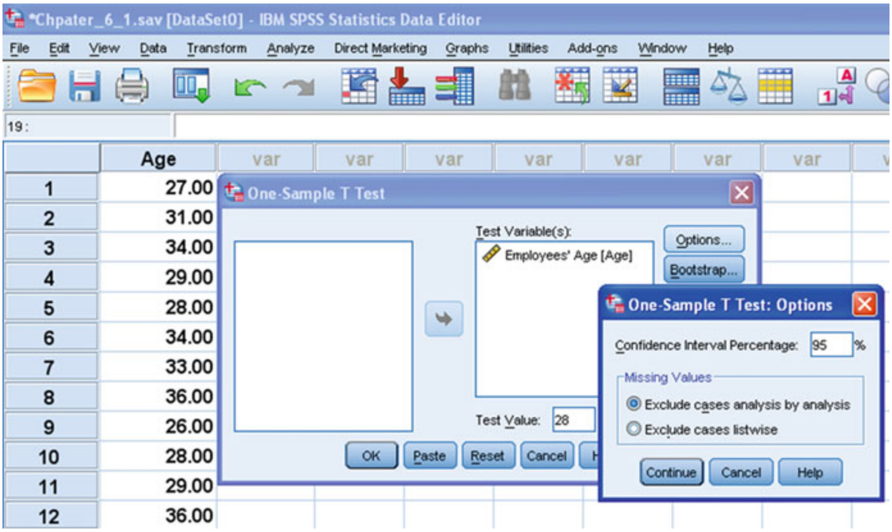


Fig. 6.9 Screen showing options for computing one-sample *t*-test and selecting significance level

Table 6.9 One-sample statistics

	N	Mean	Std. deviation	Std. error mean
Employees' age	15	31.5333	3.54293	.91478

Table 6.10 One-sample *t* test

	Test value = 28				95% confidence interval of the difference	
	t	df	Sig. (two-tailed)	Mean difference	Lower	Upper
Employees' age	3.862	14	.002	3.53333	1.5713	5.4953

Table 6.11 *t*-table for the data on employees' age

Mean	SD	Mean diff.	t-value	p value
31.53	3.54	3.53	3.862	.002

the mouse, and the content may be copied in the word file. The output panel shall have the following results:

1. Sample statistics showing mean, standard deviation, and standard error
2. Table showing the value of t and its significance level

In this example, all the outputs so generated by the SPSS will look like Tables 6.9 and 6.10. The model way of writing the results of one-sample t-test has been shown in Table 6.11.

Interpretation of the Outputs

The mean, standard deviation, and standard error of mean for the data on age are given in Table 6.9. These values may be used for further analysis.

From Table 6.11, it can be seen that the t-value is equal to 3.862 along with its p value .002. Since p value is less than 0.05, it may be concluded that the t-value is significant and the null hypothesis may be rejected at 5% level. Further, since the average age of the employees in this problem is 31.5 which is higher than the assumed age of 28 years, hence it may be inferred that average age of the employees in the organization is higher than 28 years.

Solved Example of Two-Sample *t*-Test for Unrelated Groups with SPSS

Example 6.6 An experiment was conducted to assess delivery performance of the two pizza companies. Customers were asked to reveal the delivery time of the pizza they have ordered from these two companies. Following are the delivery time in minutes of the two pizza companies as reported by their customers (Table 6.12). Can it be concluded that the delivery time of the two companies is different? Test your hypothesis at 5% level.

Solution Here the hypothesis which needs to be tested is

$$H_0 : \mu_A = \mu_B$$

against the alternative hypothesis

$$H_1 : \mu_A \neq \mu_B$$

After computing the value of t-statistic for two independent samples by the SPSS, it will be tested for its significance. The SPSS output also gives the significance value (p value) corresponding to the t-value. The t-value would be significant if its corresponding p value is less than .05, and in that case, the null hypothesis shall be rejected at 5% level; otherwise, null hypothesis is failed to be rejected.

One of the conditions in using the two sample t-test is that the variance of the two groups must be equal or nearly equal. The SPSS uses Levene's F-test to test

Table 6.12 Data of delivery time (in minutes) in two pizza companies

S.N.	Company A	Company B
1	20.5	20.5
2	24.5	17
3	15.5	18.5
4	21.5	17.5
5	20.5	20.5
6	18.5	16
7	21.5	17
8	20.5	18
9	19.5	18
10	21	18.5
11	21.5	
12	22	

this assumption. If the p value for F-test is more than .05, null hypothesis may be accepted, and this will ensure the validity of t-test.

Another important feature in this example is the style of feeding the data for SPSS analysis. The readers should note the procedure of defining the variables and feeding the data carefully in this example. Here there are two variables Pizza *Company* and *Delivery Time*. Pizza *Company* is a nominal variable, whereas *Delivery Time* is a scale variable.

Computation of Two-Sample t-Test for Unrelated Groups

(a) ***Preparing Data File***

Before using the SPSS commands for computing the t-value and other related statistics for two independent groups, a data file needs to be prepared. The following steps will help you to prepare the data file:

(i) *Starting the SPSS:* Use the following command sequence to start SPSS:
Start → All Programs → IBM SPSS Statistics → IBM SPSS Statistics 20
After checking the option **Type in Data** on the screen you will be taken to the **Variable View** option for defining the variables in the study.

(ii) *Defining variables:* There are two variables in this example which need to be defined in SPSS along with their properties. Variable *Pizza Company* is defined as 'Nominal,' whereas *Delivery Time* is defined as 'Scale' as they are measured on nominal as well as interval scale, respectively. The procedure of defining the variables in SPSS is as follows:

1. Click **Variable View** to define the variables and their properties.
2. Write short name of the variables as *Company* and *Del_Time* under the column heading **Name.**

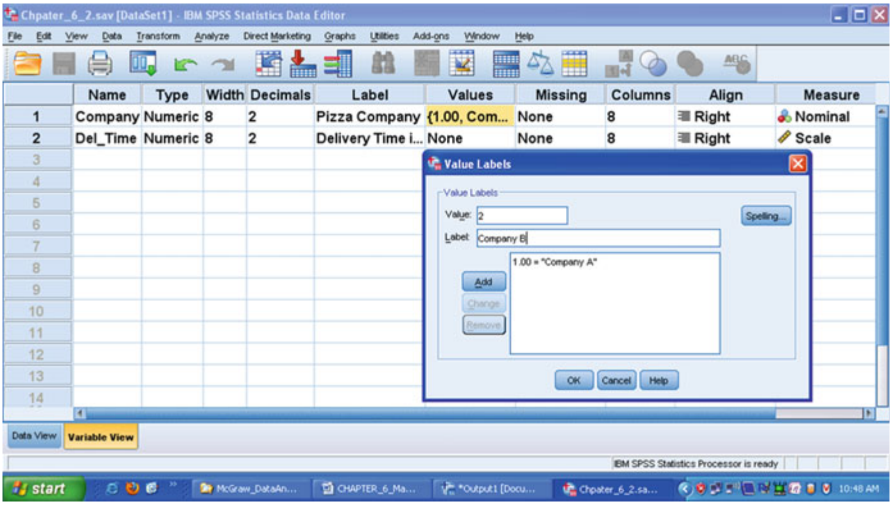


Fig. 6.10 Defining variables along with their characteristics

3. Under the column heading **Label**, full names of these variables may be defined as *Pizza Company* and *Delivery Time,* respectively. Readers may choose some other names of these variables as well.
4. For the variable *Company,* double-click the cell under the column heading **Values** and add the following values to different levels:

Value	Label
1	Company A
2	Company B

The screen for defining the values can be seen in Fig. 6.10.

5. Under the column heading **Measure,** select the option 'Nominal' for the *Company* variable and 'Scale' for the *Del_Time* variable.
6. Use default entries in rest of the columns.

After defining the variables in variable view, the screen shall look like Fig. 6.10

(iii) *Entering data*

After defining both the variables in **Variable View,** click **Data View** on the left corner in the bottom of the screen shown in Fig. 6.10 to open the data entry format column wise. For the *Company* variable, type first twelve scores as 1 and the next ten scores as 2 in the column. This is because the value '1' denotes Company A and there are 12 delivery time scores reported by the customers. Similarly, the value '2' denotes Company B and there are 10 delivery time scores as reported by the customers. After entering the data, the screen will look like Fig. 6.11.

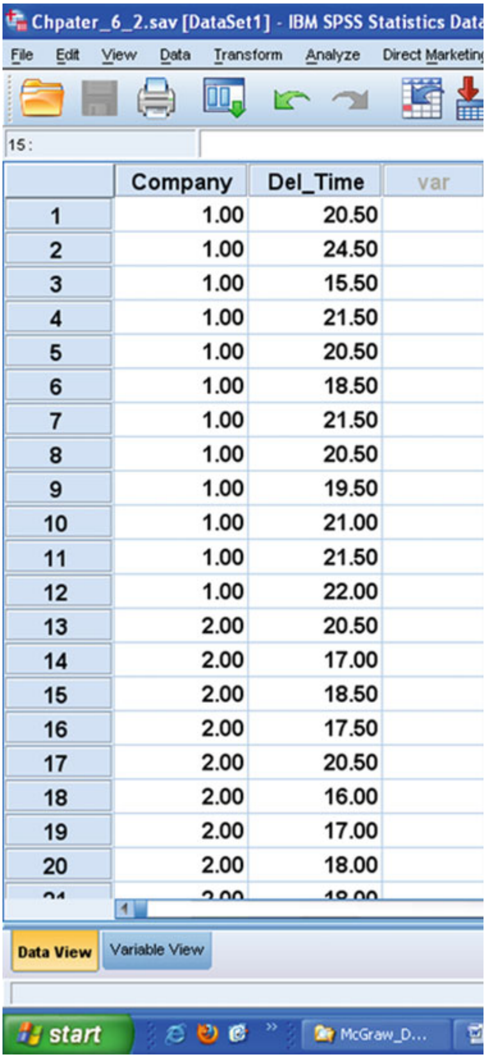


Fig. 6.11 Screen showing entered data for company and delivery time in the data view

(b) **SPSS *Commands for Two-Sample t-Test***
After preparing the data file in data view, take the following steps for two-sample *t*-test:

(i) *Initiating the SPSS commands for two-sample t-test:* In data view, click the following commands in sequence:
Analyze ⇨ Compare means ⇨ Independent-Samples *t* test
The screen shall look like Fig. 6.12.

(ii) *Selecting variables for analysis:* After clicking the **Independent-Samples *t* test** option, you will be taken to the next screen for selecting variables for the two-sample *t*-test. Select the variable *Delivery Time* from left panel and

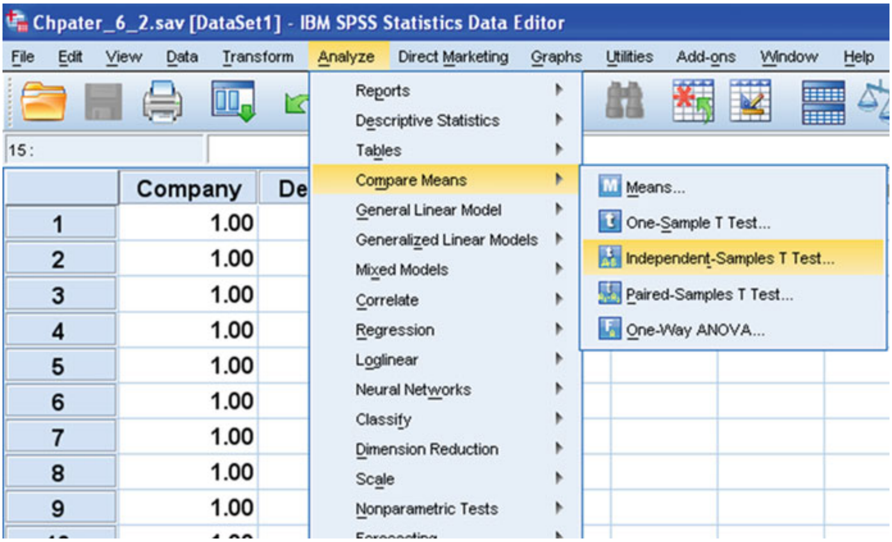


Fig. 6.12 Screen showing SPSS commands for two-sample *t*-test

bring it in the "Test Variable" section of the right panel. Similarly, select the variable *Pizza Company* from the left panel and bring it to the "Grouping Variable" section of the right panel.

Select variable from the left panel and bring it to the right panel by using the arrow key. After selecting both the variables, the values '1' and '2' need to be defined for the grouping variable *Pizza Company* by pressing the tag 'Define Groups.' The screen shall look like Fig. 6.13.

Note: Many variables can be defined in the variable view in the same data file for computing several *t*-values for different independent groups.

(iii) *Selecting options for computation:* After selecting the variables, option needs to be defined for the two-sample *t*-test. Do the following:

- In the screen shown in Fig. 6.13, click the tag **Options** and you will get the screen shown in Fig. 6.14.
- Enter the confidence interval as 95% and click ***Continue*** to get back to the screen shown in Fig. 6.13. By default, the confidence interval is 95%; however, if desired, it may be changed to some other level.
  The confidence level is the one at which the hypothesis needs to be tested. In this problem, the null hypothesis is required to be tested at .05 level of significance, and therefore, the confidence level here shall be 95%. One can choose the confidence level as 90 or 99% if the level of significance for testing the hypothesis is .10 or .01, respectively.
- Click **OK** on the screen shown in Fig. 6.13.

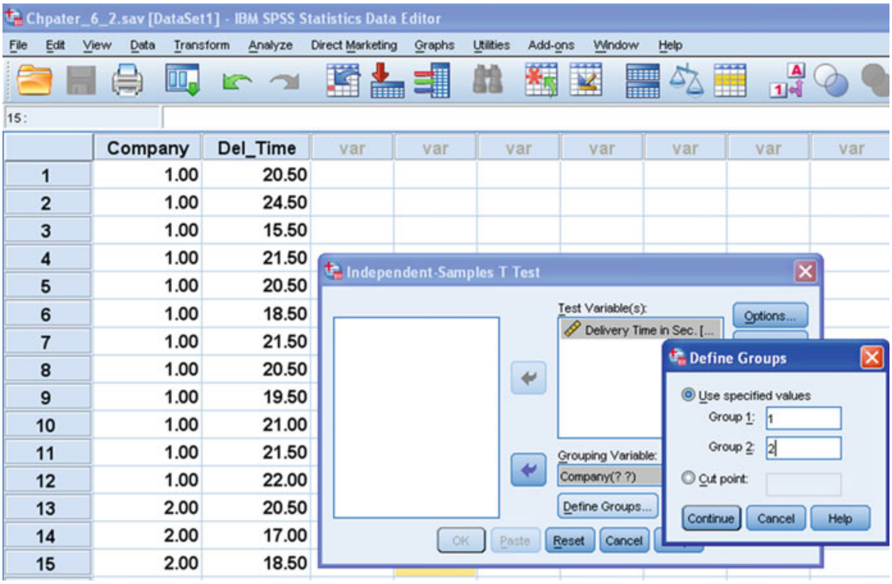


Fig. 6.13 Screen showing selection of variable for two-sample *t*-test for unrelated groups

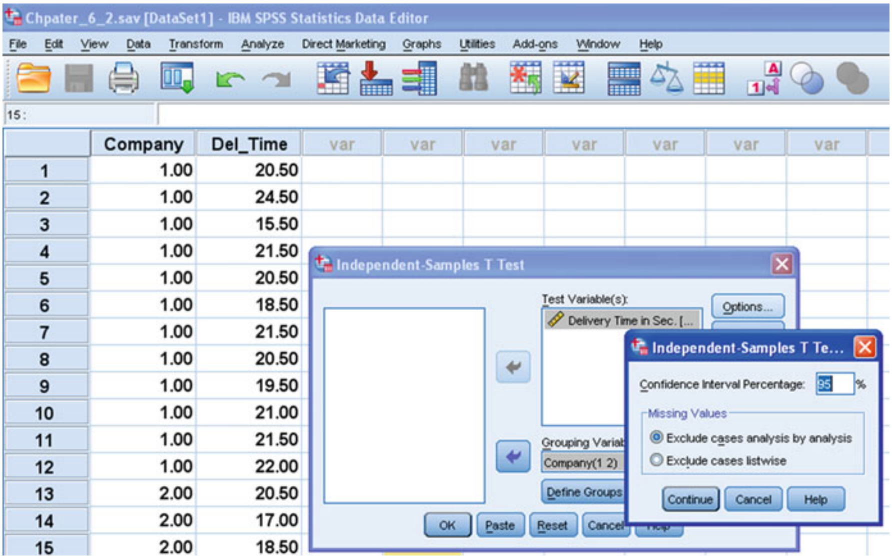


Fig. 6.14 Screen showing the option for choosing the significance level

(c) ***Getting the Output***

Clicking the OK key in Fig. 6.14 will lead you to the output window. In the output window of the SPSS, the relevant outputs may be selected by using the

Table 6.13 Descriptive statistics of the groups

	Pizza company	*N*	Mean	Std. deviation	Std. error mean
Delivery time in sec	A	12	20.58	2.16	.62412
	B	10	18.15	1.45	.45977

Table 6.14 *F*- and *t*-table for testing the equality of variances and equality of means of two unrelated groups

	Levene's test for equality of variance		*t*-test for equality of means					95% confidence interval of the difference	
	F	Sig.	*t*	df	Sig. (two-tailed)	Mean diff.	SE diff.	Lower	Upper
Delivery time in sec.									
Equal variances assumed	.356	.557	3.028	20	.007	2.43	.804	0.76	4.11
Equal variances not assumed			3.139	19.3	.005	2.43	.775	0.81	4.05

right click of the mouse, and it may be copied in the word file. The output panel shall have the following results:

1. Descriptive statistics for the data in different groups
2. '*F*-' and '*t*-'values for testing the equality of variances and equality of means, respectively

(i) In this example, all the outputs so generated by the SPSS will look like Tables 6.13 and 6.14. The model way of writing the results of two-sample *t*-test for unrelated samples has been shown in Table 6.15.

Interpretation of the Outputs

The following interpretations can be made on the basis of the results shown in the above outputs:

1. Table 6.13 shows the mean, standard deviation, and standard error of the mean for the data on delivery time of both the pizza companies. The mean delivery time of the company B is less than that of the delivery time of company A. However, whether this difference is significant or not shall be revealed by looking to the *t*-value and its associated *p* value. However, if the *t*-value is not significant, no one should draw the conclusion about the delivery time of the pizza companies by looking to the sample means.

Table 6.15 *t*-table for the data on delivery time along with *F*-value

Groups	Means	S.D.	Mean. diff	SE of mean diff	*t*-value	*p* value	*F*-value	*p* value
Company A	20.58	2.16	2.43	.804	3.028	.007	.356	.557
Company B	18.15	1.45						

2. One of the conditions for using the two-sample *t*-ratio for unrelated groups is that the variance of the two groups must be equal. To test the equality of variances, Levene's test was used. In Table 6.14, *F*-value is .356 which is insignificant as the *p* value is .557 which is more than .05. Thus, the null hypothesis of equality of variances may be accepted, and it is concluded that the variances of the two groups are equal.
3. It can be seen from Table 6.15 that the value of *t*-statistic is 3.028. This *t*-value is significant as its *p* value is 0.007 which is less than .05. Thus, the null hypothesis of equality of population means of two groups is rejected, and it may be concluded that the average delivery time of the pizza in both the companies is different. Further, average delivery time of the company B is less than that of the company A, and therefore, it may be concluded that the delivery of pizza by the company B to their customers is faster than that of the company A.

Remark: The readers can note that initially the two-tailed hypothesis was tested in this example, but the final conclusion has been made similar to the one-tailed test. This is because of the fact that if the *t*-statistic is significant in two-tailed test then it will also be significant at one-tailed test. To make it clearer, let us consider that for two-tailed test, the critical value is $t_{\alpha/2}$ at level of significance. This value will always be greater than that of the critical value of t_α in one-tailed test, and therefore, if the calculated value of t is greater than $t_{\alpha/2}$, it will also be greater than t_α.

Solved Example of Paired *t*-Test with SPSS

Example 6.7 An experiment was conducted to know the impact of new advertisement campaign on sale of television of a particular brand. The number of television units sold on 12 consecutive working days before and after launching the advertisement campaign in a city was recorded. The data obtained are shown in Table 6.16.

Solution Here the hypothesis which needs to be tested is

$$H_0 : \mu_D = 0 \text{ (Difference of average sales after and before the advertisement is zero.)}$$

against the alternative hypothesis

Table 6.16 Number of TV units sold in a city before and after the advertisement campaign

Days	Before advertisement	After advertisement
1	25	28
2	36	42
3	22	38
4	26	40
5	18	35
6	8	12
7	23	29
8	31	52
9	25	26
10	22	26
11	20	25
12	5	7

$H_1 : \mu_D \neq 0$ (Difference of average sales after and before the advertisement is not equal to zero.)

After getting the value of *t*-statistic for paired sample in the output of SPSS, it needs to be tested for its significance. The output so generated by the SPSS also gives the significance level (*p* value) along with *t*-value. The null hypothesis may be rejected if the *p* value is less than .05; otherwise, it is accepted. If the null hypothesis is rejected, an appropriate conclusion may be drawn regarding the effectiveness of the advertisement campaign by looking to the mean values of the sales before and after the advertisement.

In this problem, there are two variables *TV Sold before Advertisement* and *TV Sold after Advertisement*. For both these variables, data shall be entered in two different columns unlike the way it was entered in two-sample *t*-test for unrelated groups.

Computation of Paired t*-Test for Related Groups*

(a) ***Preparing Data File***

The data file needs to be prepared first for using the SPSS commands for the computation of *t*-value and other related statistics. Follow the below-mentioned steps in preparing the data file.

(i) *Starting the SPSS:* Follow the below-mentioned command sequence to start SPSS on your computer:

Start → All Programs → IBM SPSS Statistics → IBM SPSS Statistics 20

(ii) *Defining variables:* In this example, two variables *TV Sold before Advertisement* and *TV Sold after Advertisement* need to be defined along with their properties. Both these variables are scalar as they are measured on

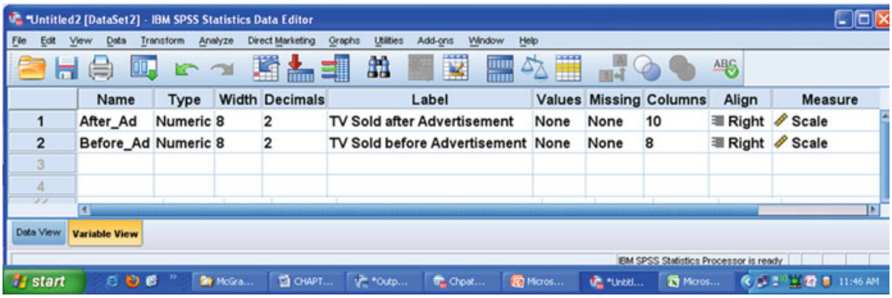


Fig. 6.15 Variables along with their characteristics for the data shown in Table 6.16

ratio scale. These variables can be defined along with their properties in SPSS by using the following steps:

1. After clicking the Type in Data above, click the **Variable View** to define the variables and their properties.
2. Write short name of the variables as *After_Ad* and *Before_Ad* under the column heading **Name.**
3. Under the column heading **Label,** full name of these variables may be defined as *TV Sold before Advertisement* and *TV Sold after Advertisement,* respectively. Readers may choose some other names of these variables if so desired.
4. Under the column heading **Measure,** select the option 'Scale' for both the variables.
5. Use default entries in rest of the columns.

After defining the variables in variable view, the screen shall look like Fig. 6.15.

(iii) *Entering the data*

Once both these variables are defined in the **Variable View**, click **Data View** on the left corner in the bottom of the screen as shown in Fig. 6.15 to open the format for entering the data column wise. For both these variables, data is entered column wise. After entering the data, the screen will look like Fig. 6.16.

(b) **SPSS *Commands for Paired t-Test***

After entering all the data in the data view, take following steps for paired *t*-test.

(i) *Initiating SPSS commands for paired t-test:* In data view, click the following commands in sequence:

Analyze → Compare means → Paired-Samples *t* Test

The screen shall look like Fig. 6.17.

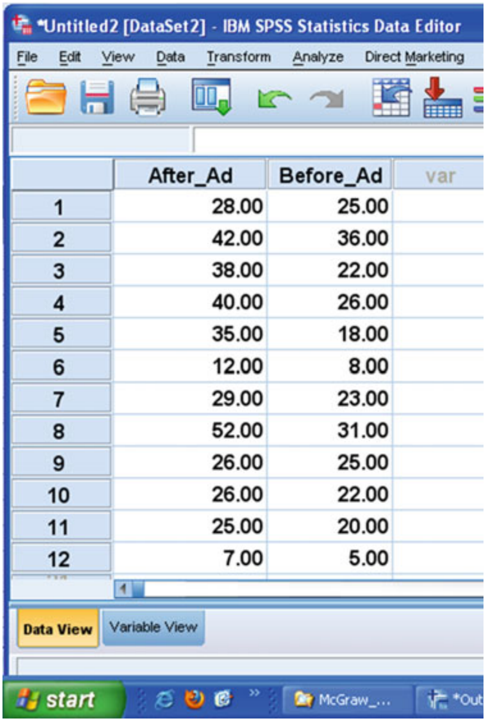


Fig. 6.16 Screen showing entered data on TV sales before and after the advertisement campaign

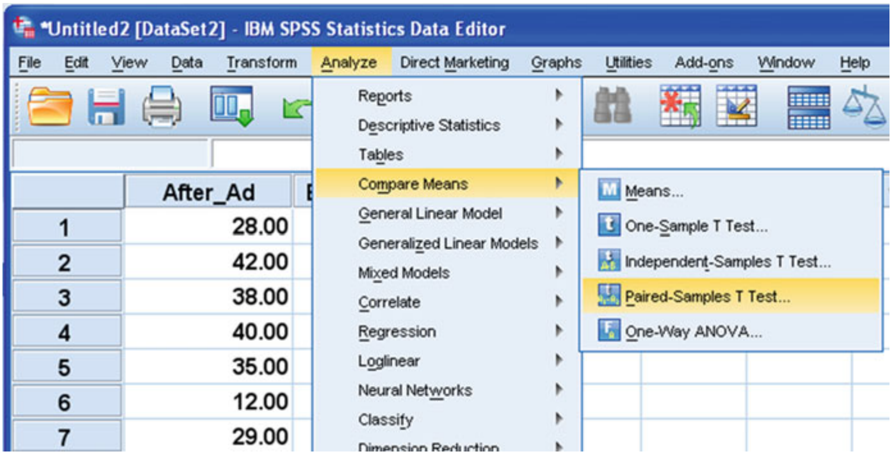


Fig. 6.17 Screen showing SPSS commands for paired *t*-test

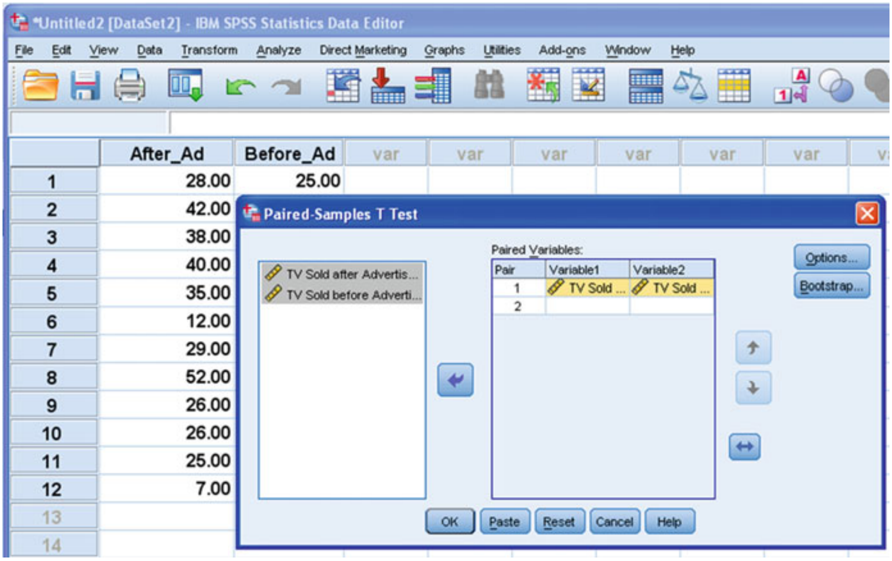


Fig. 6.18 Screen showing selection of variables for paired *t*-test

(ii) *Selecting variables for analysis:* After clicking the **Paired-Samples *t* Test,** the next screen will follow for variable selection. Select the variable *TV Sold before Advertisement* and *TV Sold after Advertisement* from left panel and bring them to the right panel as variable 1 and variable 2 of pair 1. After selecting both the variables, the screen shall look like Fig. 6.18.
Note: Many pairs of variables can be defined in the variable view in the same data file for computing several paired *t*-tests. These pairs of variables can be selected together in the screen as shown in Fig. 6.18.

(iii) *Selecting options for computation:* After selecting the variables, option needs to be defined for computing paired *t*-test. Do the following:

– In the screen as shown in Fig. 6.18, click the tag **Options** and you will get the screen where by default confidence level is selected 95%. No need of doing anything except to click ***Continue***.
One can define the confidence level as 90 or 99% if the level of significance for testing the hypothesis is .10 or .01, respectively.
– Click **OK** on the screen shown in Fig. 6.18.

(c) ***Getting the Output***
Clicking the **OK** tag in Fig. 6.18 will lead you to the output window. In the output window, the relevant results can be selected by using right click of the mouse and may be copied in the word file. The output panel shall have the following results:

Table 6.17 Paired sample statistics

		Mean	*N*	SD	SE(mean)
Pair 1					
	TV sold before advertisement	21.75	12	8.61	2.49
	TV sold after advertisement	30.00	12	12.55	3.62

Table 6.18 Paired *t*-test for the data on number of TV sold

	Paired differences			95% confidence interval of the difference				
	Mean	SD	SE(M)	Lower	Upper	*t*	df	Sig.(two-tailed)
Pair 1								
Before advertisement	8.25	6.797	1.96	3.93	12.57	4.204	11	.001
After advertisement								

1. Paired samples statistics
2. Paired *t*-test table

In this example, all the outputs so generated by the SPSS will look like Tables 6.17 and 6.18.

Interpretation of the Outputs

The following interpretations can be made on the basis of the results shown in the above output:

1. The values of the mean, standard deviation, and standard error of the mean for the data on TV sales before and after the advertisement are shown in Table 6.17. These values can be used to draw conclusion as to whether the advertisement campaign was effective or not.
2. It can be seen from Table 6.18 that the value of *t*-statistic is 4.204. This *t*-value is significant as the *p* value is 0.001 which is less than .05. Thus, the null hypothesis of equality of average TV sales before and after advertisement is rejected, and therefore, it may be concluded that the average sale of the TV units before and after the advertisement is not same.

   Further, by looking to the values of the mean sales of the TV units before and after advertisement in Table 6.17, you may note that the average sales have increased after the advertisement campaign. Since the null hypothesis has been rejected, it may thus be concluded that the increase in the TV units has been significantly increased due to advertisement campaign.

You may notice that we started with testing two-tailed test but ended up in testing one-tailed test. This is because of the fact that if the *t*-value is significant at 5% level in two-tailed test, then this will also be significant in one-tailed test.

Summary of SPSS Commands for *t*-Tests

(a) **For One-Sample *t*-Test**

(i) Start the SPSS by using the following commands:

Start → All Programs → IBM SPSS Statistics → IBM SPSS Statistics 20

(ii) Click **Variable View** tag and define the variable *Age*.
(iii) Once the variables are defined, type the data for each variable by clicking **Data View.**
(iv) In the data view, follow the below-mentioned command sequence for computing one-sample *t*-test:

Analyze ⟹ Compare Means ⟹ One-Sample *t* Test

(v) Select the variable *Age* from left panel to the right panel by using the arrow command.
(vi) Enter the test value as 28. This is the population mean age which is required to be tested.
(vii) By clicking the tag **Options,** ensure that confidence interval is selected as 95% and click ***Continue***. Confidence level can be entered as 90 or 99% if the level of significance for testing the hypothesis is .10 or .01, respectively.
(viii) Press **OK** for output.

(b) **For Two-Sample *t*-Test for Unrelated Groups**

(i) Start the SPSS the way it is done in case of one-sample *t*-test.
(ii) In the variable view, define *Company* and *Del_Time* as a 'Nominal' and 'Scale' variables, respectively.
(iii) In the variable view under column heading **Values,** define the values '1' for Company A and '2' for Company B for the variable *Company*.
(iv) In the data view, feed the data of *Company A* as 1 for first twelve entries (as there are twelve scores in the Company A) and 2 as next ten entries (as ten scores are in Company B) column wise under the column *Company*. Under the column *Del_Time*, enter the first group of delivery time data and then in the same column, enter the second group of delivery time data.
(v) In the data view, follow the below-mentioned command sequence for computing the value of *t*:

Analyze → Compare means → Independent-Samples *t* test

(vi) Select the *Company* and *Del_Time* variables from left panel and bring them in the "Test Variable" and "Grouping Variable" sections of the right panel, respectively.
(vii) Define the values 1 and 2 as two groups for the grouping variable *Company*.
(viii) By clicking the tag **Options,** ensure that confidence interval is selected as 95% and click ***Continue***.
(ix) Press **OK** for output.

(c) **For Paired *t*-Test**

(i) Start the SPSS the way it is done in case of one-sample *t*-test.
(ii) In variable view, define the variables *After_Ad* and *Before_Ad* as scale variables.
(iii) In the data view, follow the below-mentioned command sequence for computing the value of *t* after entering the data for both the variables:
Analyze → Compare means → Paired-Samples *t* Test
(iv) Select the variables *After_Ad* and *Before_Ad* from left panel and bring them to the right panel as variable 1 and variable 2 of pair 1.
(v) By clicking the tag **Options,** ensure that confidence interval is selected as 95% and click ***Continue***.
(vi) Press **OK** for output.

Exercise

Short-Answer Questions

Note: Write the answer to each of the following questions in not more than 200 words.

Q.1. What do you mean by pooled standard deviation? How will you compute it?
Q.2. Discuss the criteria of choosing a statistical test in testing hypothesis concerning mean and variances.
Q.3. What are the various considerations in constructing null and alternative hypotheses?
Q.4. What are the various steps in testing a hypothesis?
Q.5. Discuss the advantages and disadvantages of one- and two-tailed tests.
Q.6. Explain the situations in which one- and two-tailed tests should be used.
Q.7. Discuss the concept of one- and two-tailed hypotheses in terms of rejection region.
Q.8. What do you mean by type I and type II errors? Discuss the situations when type II error is to be controlled.
Q.9. What do you mean by *p* value? How it is used in testing the significance of test statistic?

Q.10. What do you mean by degrees of freedom? Discuss it in computing different statistics.
Q.11. Discuss a situation where one-sample t-test can be used. Explain the formula and procedure of testing the hypothesis.
Q.12. What are the various assumptions in using two-sample t-tests for unrelated groups? What is the solution if the assumptions are not met?
Q.13. Write the steps in testing a hypothesis in comparing the means of two unrelated groups.
Q.14. Discuss the procedure of testing a hypothesis by using paired t-test.
Q.15. Under what situations should paired t-test be used? Can it be used if sample sizes differ?

Multiple-Choice Questions

Note: Question no. 1–10 has four alternative answers for each question. Tick marks the one that you consider the closest to the correct answer.

1. If the value of t-statistic increases, then its associated p value

   (a) Increases
   (b) Decreases
   (c) Remains constant
   (d) Depends upon the level of significance chosen in the study

2. At a particular level of significance, if a null hypothesis is rejected in two-tailed test, then

   (a) It will be accepted in one-tailed test.
   (b) May be accepted or rejected in one-tailed test depending upon the level of significance.
   (c) It may be rejected in one-tailed test.
   (d) It will also be rejected in one-tailed test.

3. Choose the most appropriate statement.

   (a) t-test cannot be used for large sample.
   (b) z-test cannot be used for large sample.
   (c) t-test can be used for large sample.
   (d) Both t-test and z-test can be used for small sample.

4. Sample is said to be small if it is

   (a) 39
   (b) 31
   (c) 29
   (d) 32

5. In two-tailed hypothesis, the critical region is

   (a) Divided in both the tails in 1:4 proportion
   (b) Lying in right tail only

(c) Lying in left tail only
(d) Divided in both the tails

6. If a researcher wishes to test whether rural youth is less intelligent than the urban youth by means of the following hypotheses,

$$H_0 : \mu_{\text{Rural}} \geq \mu_{\text{Urban}}$$
$$H_1 : \mu_{\text{Rural}} < \mu_{\text{Urban}}$$

the critical region lies

(a) In the left tail only.
(b) In the right tail only.
(c) In both the tails.
(d) None of the above is correct.

7. In using two-sample t-test, which assumption is used?

(a) Variances of both the populations are equal.
(b) Variances of both the populations are not necessarily equal.
(c) No assumption is made on the population variance.
(d) Variance of one population is larger than other.

8. If Cal $t < t_\alpha$, choose the most appropriate statement.

(a) H_0 is failed to be rejected.
(b) H_1 may be rejected.
(c) H_0 may be rejected.
(d) H_1 is failed to be accepted.

9. If it is desired to compare the anxiety of male and female, which is the most appropriate set of hypotheses?

(a)
$$H_0 : \mu_{\text{Male}} = \mu_{\text{Female}}$$
$$H_1 : \mu_{\text{Male}} \neq \mu_{\text{Female}}$$

(b)
$$H_0 : \mu_{\text{Male}} = \mu_{\text{Female}}$$
$$H_1 : \mu_{\text{Male}} > \mu_{\text{Female}}$$

(c)
$$H_0 : \mu_{\text{Male}} = \mu_{\text{Female}}$$
$$H_1 : \mu_{\text{Male}} < \mu_{\text{Female}}$$

(d)
$$H_0 : \mu_{\text{Male}} \neq \mu_{\text{Female}}$$
$$H_1 : \mu_{\text{Male}} = \mu_{\text{Female}}$$

10. In testing the following set of hypotheses,

$$H_0 : \mu_1 \geq \mu_2$$
$$H_1 : \mu_1 < \mu_2$$

choose the most appropriate statement.

(a) If Cal $t \leq t_\alpha$, H_0 may not be rejected.
(b) If Cal $t < -t_\alpha$, H_0 may be rejected.
(c) If Cal $t > -t_\alpha$, H_0 may be rejected.
(d) None of the above is correct.

11. If there are N pairs of score and paired t-test is used for comparing the means of both the groups, what will be the degrees of freedom for t-statistic?

(a) N
(b) $2N - 2$
(c) $N + 1$
(d) $N - 1$

12. If attitude toward science is to be compared among 22 male and 18 women students of class XII by using t-ratio, what would be its degrees of freedom?

(a) 40
(b) 39
(c) 2
(d) 38

13. To see the effectiveness of observation method on learning skill, which of the SPSS command shall be used?

(a) One-sample t-test
(b) Independent-sample t-test
(c) Paired-sample t-test
(d) None of the above

14. Power of statistical test is given by

(a) β
(b) $1 + \beta$
(c) $\beta - 1$
(d) $1 - \beta$

Assignments

1. A random sample of 25 management students was tested for their IQ in a university. Their scores were as follows:

   92, 101, 94, 93, 97, 98, 120, 104, 98, 96, 85, 121, 87, 96, 111, 102, 99, 95, 89, 102, 131, 107, 109, 99, 97

   Can it be concluded that the management students in the university have a mean IQ score equal to 101? Test your hypothesis at 5% level.

2. The following data set represents the weight of the average daily household waste (kg/day/house) generated from 20 houses in a locality:

4.1	3.7	4.3	2.5	2.5	6.8	4.0	4.5	4.6	7.1
3.5	3.1	6.6	5.5	6.5	4.1	4.2	4.8	5.1	4.8

Can it be concluded that the average daily household waste of that community is 5.0 kg/day/house? Test your hypothesis at 1% level.

3. A feeding experiment was conducted with two random samples of pigs on the relative value of limestone and bone meal for bone development. The data so obtained on ash content are shown in the following table:

Ash contents (%) in the bones of pigs

S.N.	Lime stone	Bone stone
1	48.9	52.5
2	52.3	53.9
3	51.4	53.2
4	50.6	49.9
5	52	51.6
6	45.8	48.5
7	50.5	52.6
8	52.1	44.6
9	53	52.8
10	46.5	48.8

Test the significance of the difference between the mean ash content of the two groups at 5% level.

4. A company wanted to know as to which of the two pizza types, that is, fresh veggie and peppy paneer, was most popular among the people. An experiment was conducted in which 12 men were given two types of pizza, that is, fresh veggie pizza and pepper paneer pizza, to eat on two different days. Each pizza was carefully weighed at exactly 16 oz. After 20 min, the leftover pizzas were weighed, and the amount of each type of pizza remaining per person was calculated assuming that the subjects would eat more if they preferred the pizza type. The data so obtained is shown in the following table.

Weights of the leftover pizzas in both varieties

S.N.	Fresh veggie (in oz.)	Pepper paneer (in oz.)
1	12.5	15
2	5.87	7.1
3	14	14
4	12.3	13.7
5	3.5	14.2
6	2.6	5.6
7	14.4	15.4
8	10.2	11.3
9	4.5	15.6
10	6.5	10.5
11	4.3	8.5
12	8.4	9.3

Apply the paired *t*-test and interpret your findings. Do people seem to prefer fresh veggie pizza over pepper veggie pizza? Test your hypothesis at 5% level.

Answers to Multiple-Choice Questions

Q.1 b	Q.2 d
Q.3 c	Q.4 c
Q.5 d	Q.6 a
Q.7 a	Q.8 a
Q.9 a	Q.10 b
Q.11 d	Q.12 d
Q.13 c	Q.14 d

Assignments

1. Calculated value of $t = -0.037$; average IQ score of the students is 101.
2. Calculated value of $t = -1.286$; average daily household waste of the community is 5 kg/day/house.
3. Calculated value of $t = -0.441$; mean ash contents of both the groups are same.
4. Calculated value of $t = 3.193$ which is significant. People prefer fresh veggie pizza.