Testing of hypotheses: T test

Mr. Ajit Pawar

Assistant Professor,
Department of Statistics
Vivekanand College, Kolhapur

Introduction to Testing of Hypotheses:

Inferential statistics is all about making decisions or judgments about the value of a particular observation or measurement. One of the most commonly used methods for making such decisions is to perform a hypothesis test. A hypothesis is a proposed explanation for a phenomenon. In the context of statistical hypothesis tests the term hypothesis is a statement about something that is supposed to be true.

Introduction to Testing of Hypotheses:

A hypothesis test involves two hypotheses: the null hypothesis and the alternative hypothesis.
The null hypothesis (H_0) is a statement to be tested.
The alternative hypothesis (H_A) is a statement that is considered to be an alternative to the null hypothesis.

Introduction to Testing of Hypotheses:

The hypothesis test is aimed to test if the null hypothesis should be rejected in favor of the alternative hypothesis. The basic logic of a hypothesis test is to compare two statistical data sets. One data set is obtained by sampling and the other data set originates from an idealized model. If the sample data is consistent with the idealized model, the null hypothesis is not rejected. If the sample data is inconsistent with the idealized model and thus supports an alternative hypothesis, the null hypothesis is rejected in favor of the alternative hypothesis.

Introduction to Testing of Hypotheses:

The criterion for deciding whether to reject the null hypothesis involves a so called test statistic. The test statistic is a number calculated from the data set, which is obtained by measurements and observations, or more general by sampling.

Introduction to T test:

The t-test

1. Parametric: works on normally distributed scale data
2. Compares TWO means
3. There are different versions for different designs

Applications of T test:

1. To test if the sample mean $\bar{x}$ differs significantly from the hypothetical value of μ.
2. To test the significance of the difference between two sample means.
3. To test the significance of observed sample correlation coefficient and sample regression correlation coefficient.
4. To test the significance of observed partial correlation coefficient.

t-test for single mean:

H_o: There is no significant difference between the sample mean and the population mean μ_0 of the population mean
H_1: There is significant difference between the sample mean and the population mean μ_0 of the population mean

t test: Test Statistic

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

Decision Rule: Calculated $|t| > t_{tab}$, then reject H_0 otherwise accept.

Independent Samples T Test

A t test compares the means of two groups.For example, compare whether systolic blood pressure differs between a control and treated group, between men and women, or any other two groups.

1. Don't confuse t tests with correlation and regression. • The t test compares one variable (perhaps blood pressure) between two groups.
2. Use correlation and regression to see how two variables (perhaps blood pressure and heart rate) vary together.
3. Also don't confuse t tests with ANOVA. The t tests (and related nonparametric tests) compare exactly two groups.
4. ANOVA (and related nonparametric tests) compare three or more groups.
5. Finally, don't confuse a t test with analyses of a contingency table (Fishers or chi-square test).

Independent Samples T Test

1. Use a t test to compare a continuous variable (e.g., blood pressure, weight or enzyme activity). Use a contingency table to compare a categorical variable (e.g., pass vs. fail, viable vs. not viable).
2. This simple t-test calculator, provides full details of the t-test calculation, including sample mean, sum of squares and standard deviation.
3. A t-test is used when you're looking at a numerical variable - for example, height - and then comparing the averages of two separate populations or groups (e.g., males and females).

Assumptions:Independent Samples T Test

1. Two independent samples; One independent, categorical variable that has two levels/groups. One continuous dependent variable.
2. Unrelated groups
3. Data should be normally distributed
4. The two samples should have the same variance
5. Assumption of homogeneity of variance

Test statistic: Independent Samples T Test

$$t = \frac{(x_1 - x_2)}{\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}}$$

Where

X_1 is mean of first group

X_2 is mean of second group

S_1 is Standard deviation of first group

S_2 is Standard deviation of second group

n_1 is sample size of first group

n_1 is sample size of second group

PAIRED SAMPLES T-TEST

Paired-samples t-test compares the means between two related groups on the same continuous, dependent variable.
For example, you could use a dependent t-test to understand whether there was a difference in smokers' daily cigarette consumption before and after a 6 week hypnotherapy programme (i.e., your dependent variable would be "daily cigarette consumption", and your two related groups would be the cigarette consumption values "before" and "after" the hypnotherapy programme).

Assumptions: PAIRED SAMPLES T-TEST

1. Your dependent variable should be measured on a continuous scale (i.e., it is measured at the interval or ratio level). Examples of variables include revision time (measured in hours), intelligence (measured using IQ score), exam performance (measured from 0 to 100), weight (measured in kg), and so forth.
2. Your independent variable should consist of two categorical, "related groups". "Related groups" indicates that the same subjects are present in both groups. It is possible to have the same subjects in each group is because each subject has been measured on two occasions on the same dependent variable.

Assumptions: PAIRED SAMPLES T-TEST

3 There should be no significant outliers in the differences between the two related groups. Outliers are simply single data points within your data that do not follow the usual pattern (e.g., in a study of 100 students' IQ scores, where the mean score was 108 with only a small variation between students, one student had a score of 156, which is very unusual, and may even put her in the top 1globally).

4 The distribution of the differences in the dependent variable between the two related groups should be approximately normally distributed.

Test statistic: Paired T Test

$$t = \frac{\Sigma d}{\sqrt{\frac{n\left(\Sigma d^2\right) - \left(\Sigma d\right)^2}{n-1}}}$$

Where; sum of d= Difference

n= No. of cases

d^2= Square of differences

Paired T Test

A paired t-test can determine if the mean of the pre test scores is significantly different than the mean of their post-test scores by testing if the mean difference in scores for these subjects was different from zero. Although the paired t-test is considered a "two-sample" t-test, it is actually the same as running a one-sample t-test on the differences.