

Choosing an Appropriate Test of Significance

Padam Singh, PhD, Gurgaon, India

ABSTRACT

Inferences are drawn about the population parameters based on random samples drawn from the population. While applying statistical test of significance, the first step is to set a null hypothesis as well as the competing alternative hypothesis. To decide whether the null hypothesis is to be accepted or rejected, a 'test statistic', based on the sample values, is computed and compared with a critical value from statistical tables for the test corresponding to the desired level of significance. When 'test statistic' exceeds the critical value the null hypothesis is rejected and the difference is declared as statistically significant. The specific test of significance to be used in a given situation is chosen on the basis of type of variables and the underlying null hypothesis. This paper outlines the types of variables encountered in various situations and the statistical test of significance used in each case. (J Clin Prev Cardiol. 2014;3(1):21-5)

Keywords: Continuous variable, Critical value, Discrete variable, Qualitative variable, Quantitative variable, Test statistic.

Background

While applying statistical tests of significance, the first step is setting up of a *null hypothesis*. In a study on evaluating the efficacy of a new drug, the null hypothesis (H_0) is that the new drug is no better, than the current drug. The *alternative hypothesis*, (H_1) i.e. the competing hypothesis would be that the new drug is better than the current drug.

Once the statistical test has been applied, the inference is either to "reject H_0 in favor of H_1 " or "do not reject H_0 ". If inference is "do not reject H_0 ", it indicates that there is no sufficient evidence against H_0 in favor of H_1 ; whereas if inference is to reject the null hypothesis, then it is indicated that the alternative hypothesis may be true.

Hypotheses are always stated in terms of **population parameter**, such as the mean ' μ ' or proportion 'P'. An alternative hypothesis may be one-sided or two-sided. A one-sided hypothesis claims that the parameter under the alternative hypothesis is either larger or **smaller** than the value given by the null hypothesis. A two-sided hypothesis claims that the parameter is not equal to the value given by the null hypothesis - the **direction** does not matter and could be either side.

The foremost consideration in deciding which statistical test to be used in a given situation is determined by the kind of variables we are dealing with. In this article the various types of variables encountered in research are discussed along with the statistical tests of significance to be used for each situation.

Types of Variables

Variables in general are characteristics or attributes which vary from person to person. Variables are broadly categorized into two types: Quantitative and Qualitative variables.

Quantitative variables are of two types, namely Continuous and Discrete. Similarly the qualitative variables are also of two types, that are Ordinal and Nominal.

As already mentioned, the qualitative variables are of two types-

- Ordinal - The one that can be graded, like severity of cancer graded as stage I, II, III and IV; NYHA Classification-The Stages of Heart Failure (Class I – Mild; Class-II – Mild; Class III – Moderate and Class IV – Severe), etc.
- Nominal - The one that cannot be graded and is meant for categorization, like gender (Male or Female), site of malignancy (lungs, mouth, breast or ovary), blood groups (A, B, AB, O), genotype (AA, Aa, or aa), type of disability (visual, hearing impairments, mental, locomotor), etc.

From: Medanta - The Medicity, Sector 38, Gurgaon, India (P.S.)

Corresponding Author: Dr. Padam Singh, PhD
Medanta - The Medicity, Sector 38, Gurgaon, Haryana-122001,
India

Ph: +91-124-4141414 | Fax: +91-124-4834111

Email: dr.padamsingh2013@gmail.com

Quantitative variables are measurable. The measurement is always a number. A continuous variable is one which can take any value in a given range. Examples of continuous variables are levels of hemoglobin, blood glucose, urinary creatinine excretion, bilirubin, etc. On the other hand, a discrete variable is the one which can take a value based on a count from a set of distinct whole numbers, that is, a discrete variable cannot take be a fraction. Examples of discrete variables are hospitalization days, family size, parity, etc.

Importantly, all nominal and ordinal variables are discrete after coding.

The variables, called “categorical variables,” classify observations into a small number of categories. Nominal variables are often used to divide individuals into classes/groups, so that other variables may be compared among the classes.

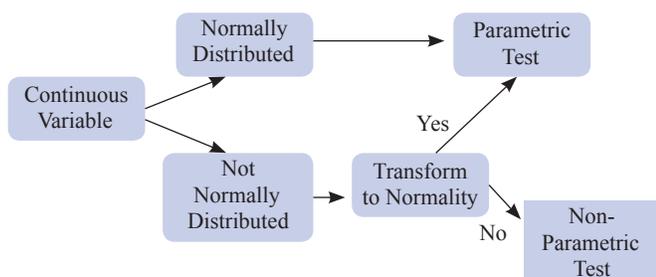
It is possible to convert a continuous variable to a nominal/ordinal variable, which is common in epidemiology. For example, the study subjects whose measurements on HDL levels are taken, could be divided into two groups namely “low HDL” (less than 40 mg/dl) and “normal HDL” (40 or more mg/dl).

Binary variables: Presence of any disease/illness is either yes or no. He/she either survives or dies. Such variables of dichotomous types are called binary.

The test of significance for different types of variables are discussed in the following sections.

Pathway of for Analysis of Continuous Data

Following diagram depicts the pathway for test of significance in case of continuous variables:



For normality mean and median should be approximately equal and the standard deviation is less than 30% of the mean, else the data will not satisfy the normality assumption.

For continuous variables, satisfying the normality assumption, the parameter of interest under test is the population mean. The next section deals with the test of significance for mean in different situations.

Test of Significance for Population Mean

Situation A: One sample test when standard deviation $\{\sigma\}$ is known.

In this situation, the null hypothesis is $H_0: \mu = \mu_0$

The alternative hypothesis could be $\mu \neq \mu_0$ or $\mu < \mu_0$ or $\mu > \mu_0$

For such situations the appropriate significance test is ‘z test’ which defined as

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

Here ‘n’ denotes sample size

The test statistic follows the standard normal distribution, i.e. normal distribution with mean ‘0’ and standard deviation ‘1’.

Example

The normal value of the pulse rate for the population is taken as 72, with standard deviation ‘10’. It is assumed that the pulse rate follows normal distribution. In a sample of 16 persons doing physical exercise the mean pulse rate is observed as 78. Does this provide evidence that the mean pulse rate among the persons doing exercise is higher than the normal?

The null hypothesis H_0 in this situation will be that there is no difference between the mean pulse rate for persons doing physical exercise and the normal pulse rate. The alternative hypothesis would be that the mean pulse rate for persons doing the physical exercise is higher than the normal.

In the example above, $z = (78 - 72)/(10/\sqrt{16}) = 6/2.5 = 2.4$

Since this is a one-sided test, the *P-value* is calculated as -

$$P(z > 2.4) = 1 - P(z \leq 2.4) = 1 - 0.9918 = 0.0082.$$

Since, the *P-value* is less than 0.01, one can confidently reject H_0 given this result. Thus the average pulse rate for persons doing physical exercise is higher than the normal.

Situation B: One sample test when standard deviation is unknown.

In most practical situations in research, the standard deviation for the population of interest is not known. In this case, the standard deviation σ is replaced by the estimated standard deviation s . The distribution of the sample mean \bar{x} is no longer normal. Instead, the sample mean follows the *t distribution* with mean μ and standard deviation s/\sqrt{n} . The *t* distribution is also described by its *degrees of freedom*. For a sample of size n , the *t* distribution will have $n-1$ degrees of freedom. As the sample size n increases, the *t* distribution becomes closer to the normal distribution.

The test statistic ‘*t*’ is defined as

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

Example

Following table summarizes results on body temperature for 9 persons.

Variable - Temperature
 Sample Size (n) = 9
 Sample Mean (\bar{x}) = 98.1
 Sample Standard Deviation (s) = 0.72

Since the normal body temperature is generally assumed to be 98.4 degrees Fahrenheit, one can use the data to test the following one-sided hypothesis:

$$H_0: \mu = 98.4$$

$$H_a: \mu < 98.4$$

The test statistic *t* in this case will be

$$t = (98.1 - 98.4) / \{0.72 / \sqrt{9}\}$$

$$= -0.3 / 0.24$$

$$= -1.5$$

$$= 1.5 \text{ (ignoring the sign)}$$

The table values of *t* with 5% level of significance (1-tailed) for 8 degrees of freedom is 2.306. Since the calculated value of *t* is less than the table value of *t*, there is no sufficient evidence to reject the null hypothesis. Hence the average temperature of the studied sample is not less than the normal temperature.

Situation C: Matched Pairs

In many experiments, one is interested in comparing

the measurements for the same set of individuals in ‘before’ and ‘after’ situation. Such studies have a *matched pairs* design, where the *difference* between the two measurements in each pair is the variable of interest.

Analysis of data from a matched pairs experiment compares the two measurements by subtracting one from the other and basing test hypotheses upon the differences. Usually, the null hypothesis H_0 assumes that the mean of these differences is equal to 0, while the alternative hypothesis H_a would be that the mean of the differences is not equal to zero (the alternative hypothesis may be one- or two-sided, depending on the experiment). In this situation, using the differences between the paired measurements as single observation, the standard t-test for one sample is followed.

The difference for each pair of values is denoted by *d*. Once, this is obtained, the situation reduces to one sample *t* given by

$$t = \bar{d} / \{s_d / \sqrt{n}\}$$

where \bar{d} is the sample mean and S_d the standard deviation of the differences

Example

Consider albumin levels of 8 patients undergoing treatment for liver.

	Patient -1	Patient -2	Patient -3	Patient -4
Before treatment	5.2	3.9	4.1	4.8
After treatment	4.9	3.8	3.9	4.8
Difference	0.3	0.1	0.2	0.0
	Patient -5	Patient -6	Patient -7	Patient -8
Before treatment	4.6	4.9	4.2	3.7
After treatment	4.7	5.1	4.1	3.5
Difference	-0.1	-0.2	0.1	0.2

For above example $\bar{d} = 0.075$; $s_d = 0.17$;

$$t = 0.075 / \{0.17 / \sqrt{8}\} = 1.25$$

The critical value of *t* for 7 degrees of freedom and 5% level of significance (two-tailed) is 2.365. Since the calculated value is less than the critical value, the null hypothesis of no difference cannot be rejected.

Situation D: Two sample test

Comparing averages of one group with that of another group to decide whether they are equal or not, two sample Student's *t*-test is applied which under usual notations is given by:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\{(1/n_1)+(1/n_2)\}}}$$

Here, the degrees of freedom will be $n_1+n_2 -2$

Example

A study on 24 hour creatinine excretion in 15 males and 15 females to examine if a difference exists yielded following results:-

Subject	Male	Female
1	16.8	23.4
2	20.0	22.2
3	17.3	22.1
4	17.2	16.9
5	15.8	14.4
6	20.5	23.4
7	24.9	25.0
8	18.7	25.7
9	17.8	28.3
10	22.2	22.0
11	25.1	21.1
12	18.6	18.2
13	17.1	19.7
14	21.3	20.8
15	23.5	17.5

The null hypothesis for this example is that there is no difference between males and females in 24 hour creatinine excretion. Since it is not known a priori that males would have higher creatinine excretion or females the alternate hypothesis is 2 sided.

In this case the calculated value of *t* is 1.30.

The table value of *t* for 28 degrees of freedom at 5% level of significance (2 tailed) is 2.048.

Since calculated value of *t* is less than table value of *t*, the null hypothesis of equality cannot be rejected. Thus, the samples do not provide sufficient evidence that

the average level of creatinine excretion in men is any different than in women.

Proportion Test

Similar to test for averages the *z* test (based on Gaussian distribution) is also employed in testing for proportions.

Situation A: One Sample Test

In case of one sample test, the test statistic 'z' is given by

$$z = \frac{p - P}{\sqrt{P(1-P)/n}}$$

Here, *p* = Sample proportion
P = Proportion under Null Hypothesis
n = Sample Size

Example

In a hospital, out of 400 births, the male children born were 216 and 184 female. Does the data suggest that the birth of male or female child is equally likely. In this situation the null hypothesis is that the proportion of either sex is 50%, as against the alternative hypothesis that it is not 50%.

The calculated value of *z* in this situation is 1.6 which is less than the critical value 1.96 at 5% level of significance. Hence, the null hypothesis that the sex of the children born is equally likely cannot be rejected.

Situation B - Two Sample Test

Many times proportions (cure rate, prevalence of a condition, etc.) for 2 groups are required to be compared and in such situations two-sample *z* test is used, which is given by

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

Here, *p*₁, *p*₂ = Proportions for Sample 1 and 2
*n*₁, *n*₂ = Sizes for Sample 1 and 2

Example

In samples based on 60 and 40, the proportion of cataract cases among males and females is observed as 62% and 75% respectively. It is of interest to know whether the prevalence of cataract is similar among males and females. Thus, the *z* test is applied. In this case the calculated value of *z* is 1.4 which less than the

critical value 1.96 and hence the null hypothesis cannot be rejected.

Other Commonly used Tests

As already mentioned, the type of data involved determines the specific procedure used to perform the significance test. When the individual observations are categorical and are summarized in frequency tables, the chi-square test is used. There are other tests meant for

situations comparing more than two groups (ANOVA) as well as relating to repeat observations over time (Repeat Measure Analysis). These will be subject matters of the next article.

Reference

1. Indrayan A, L Satyanarayana. Simple Biostatistics. Academia Publishers. 2003.