The Single-Sample t-Test

Hypothesis testing using single-sample tests: We have already studied the single-sample z-test, and we will now take up the single-sample t-test. Both of these tests are used when we have a single random sample from some population. There is no control group in these studies other than the general population from which the sample is chosen, so random sampling is critical, and we must know something about that population.

Single-sample z-test: The single-sample z-test for a sample mean is used when both μ and σ are known. The z-test can also be used when the sample is so large that σ can be very accurately estimated by the sample standard deviation, *s*. For that reason, the z-test is sometimes referred to as the large sample test. The following assumptions must be met:

1) The subjects must be randomly sampled from the population.
2) The sample must consist of independent observations.
3) The treatment cannot change the variability (σ) of the scores.
4) The distribution of sample means (sampling distribution) must be normal.

Single-sample t-test: The single-sample t-test for a sample mean is used when μ is known but the value of σ is not known and must be estimated by *s*. When the sample is small, this estimate will not be especially accurate, and for that reason the sampling distribution will not be normal. Thus, the t-test is sometimes referred to as the small sample test. Its assumptions are similar to those above:

1) The subjects must be randomly sampled from the population.
2) The sample must consist of independent observations.
3) The distribution of raw scores must be normal. (Note: This is especially important when the sample size is very small. With larger samples the t-test is somewhat *robust* to violations of this assumption, but the distribution of raw scores should still be "more or less" unimodal and symmetrical.)

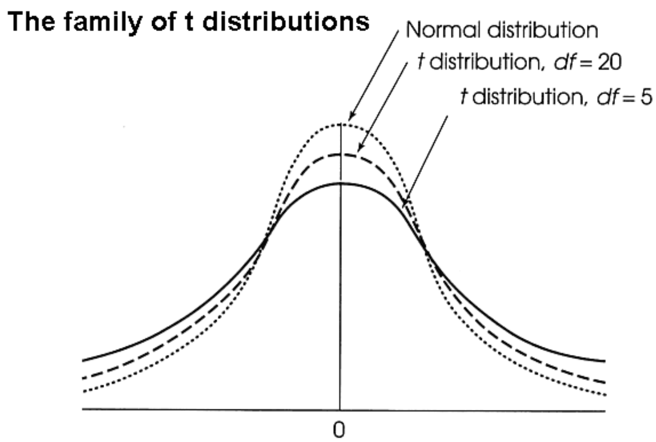
Formulas: The similarities between these two tests are also illustrated by their formulas, which are very similar.

$$z = \frac{M - \mu_0}{\sigma_M} \qquad \text{vs.} \qquad t = \frac{M - \mu_0}{s_M},\ df = n - 1$$

There are two important differences here. The first is in the denominator, where the z-test uses the standard error of the mean (s.e.m.), but the t-test uses the *estimated standard error of the mean* (still often abbreviated s.e.m.), which is calculated as follows:

$$s_M = \frac{s}{\sqrt{n}} = \sqrt{\frac{s^2}{n}}$$

The second difference is the t-test requires the calculation of *degrees of freedom* (*df*). This is required because the t-distribution is not just one distribution as the normal distribution is, but is actually a family of distributions. Each different value of *n* gives a slightly different distribution. Notice as *n* becomes large, the t-distributions gets closer and closer in shape to the normal distribution.



In all cases, the distribution is symmetrical around zero and more or less bell-shaped.

Hypothesis testing using the single-sample t-test: One of the original tests for ESP involved the use of special cards called Zener cards, which is a deck of cards marked with the following symbols:

In a deck of 100 cards, each symbol occurs 20 times, and a subject guessing at random (i.e., without ESP) would be expected to guess correctly 20 times. A researcher will take a random sample of subjects and test them on the Zener cards in an attempt to find evidence of the existence of ESP.

Step 1) State the null and alternative hypotheses.

H_0:

H_1:

Step 2) Establish a decision criterion (set alpha).

Step 3) Go out and collect the data, calculate summary statistics, and calculate the value of the test statistic.

The following values represent number of correct guesses in 100 trials for subjects tested in this study:

20 28 25 23 20 21 24 19 16 18 27 23 26 18 19 16 25

Step 4) Make a statistical decision concerning the null hypothesis.

Step 5) Write a conclusion describing the result of the statistical analysis.

Question) Does this result indicate the presence of ESP in these subjects?

A researcher hypothesizes that people who listen to classical music will score differently from the general population on a test of spatial ability. On a standardized test of spatial ability, $\mu = 58$.

Step 1) State the null and alternative hypotheses.

H_0:

H_1:

Step 2) Establish a decision criterion (set alpha).

Step 3) Go out and collect the data, calculate summary statistics, and calculate the value of the test statistic.

A random sample of 14 people who listen to classical music is given this test. Their scores appear below:

55 65 63 69 58 50 62 66 53 68 57 61 72 64

Step 4) Make a statistical decision concerning the null hypothesis.

Step 5) Write a conclusion describing the result of the statistical analysis.

Effect size: A common misunderstand is that big *t* values, or big mean differences, imply large effects. This is not the case. With enough subjects, even tiny little effects can yield large values of *t*. And big mean differences can turn out nonsignificant if there is too much error variability. Thus, it is necessary to have separate measures of effect size. Two are commonly used and are discussed in the book: Cohen's d, and r^2.

$$\text{Cohen's } d = \frac{M - \mu_0}{s}, \quad r^2 = \frac{t^2}{t^2 + df}$$