

Weighted Data and Significance Testing Tools

When dealing with other than unweighted data based on independent simple random samples WinCross uses sophisticated methodology when testing significance of mean or percentages (proportions). Details can be found in the [Statistical Reference Guide](#) available from the WinCross **Help** menu . Statistical testing becomes much more complicated when dealing with weighted data.

The **Significance Testing** tool shown below, will be accessible from the **Tools** menu in WinCross V16 and only applies to significance testing for *unweighted* statistics based on independent simple random samples. The tool can also be downloaded by going to www.AnalyticalGroup.com/support_free_tools.htm. Our significance testing tool, as well as others like it available on the web, should only be used for statistical tests based on *unweighted* data from independent samples.

Some users might want to use this tool for statistical testing with weighted data. There are two approaches to the treatment of weighted data in statistical tests, the one employed by WinCross and the one employed by some other statistical programs, notably SPSS. The WinCross approach is the more accurate approach; documentation of this claim can be found on The Analytical Group website www.analyticalgroup.com/support_wc_faqs.htm under the **HELPFUL DOCUMENTS** heading.

The following describes ways to use this tool with weighted data.

Averages:

The screenshot shows the 'WinCross Quick Tools' application window. The title bar reads 'WinCross Quick Tools'. The main header area contains the logo 'WinCross Quick Tools', social media icons for Facebook, Twitter, and LinkedIn, and copyright information: 'Copyright © 2015 by The Analytical Group, Inc. All Rights Reserved Worldwide'. The 'TAG' logo and 'The Analytical Group Inc.' logo are also present.

The interface has a tabbed menu with four tabs: 'Statistics and Frequency', 'Significance Testing' (which is selected), 'Sample Size Calculator', and 'Calculator'. Under the 'Significance Testing' tab, there are two sub-tabs: 'Percentages' and 'Averages' (which is selected). Under the 'Averages' sub-tab, there are three sub-sub-tabs: 'Two Independent Samples' (which is selected), 'Average vs. Theoretical', and 'Two Matched Samples'.

The 'Two Independent Samples' sub-tab contains the following input fields and controls:

- Mean #1 (x1):
- Standard deviation #1 (s1):
- Sample size #1 (n1):
- Mean #2 (x2):
- Standard deviation #2 (s2):
- Sample size #2 (n2):
- z score:
- Significance level: (with 'Two tail' and 'One tail' radio buttons above it)
- Confidence level: (with 'Two tail' and 'One tail' radio buttons above it)

Text on the right side of the interface explains the test: 'Here we test whether an observed mean x1 based on a random sample of size n1 is significantly different from an observed mean x2 based on an independent sample of size n2. To perform this test one needs additionally the standard deviations s1 and s2 from samples 1 and 2. The z statistic for testing this hypothesis is given by' followed by the formula:

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

Below the formula, it states: 'This tool reports the significance level and confidence level for both the two-tailed and one-tailed tests of significance of p1 from p2.'

WinCross approach: For each of the samples, fill in the mean box with the weighted mean, fill in the standard deviation box with the unweighted standard deviation, and fill in the sample size box with the effective sample size, calculated as

$$f = \frac{\sum_{i=1}^n w_i}{\sum_{i=1}^n w_i^2}$$

SPSS approach: For each of the samples, fill in the mean box with the weighted mean, fill in the standard deviation box with the weighted standard deviation, and fill in the sample size box with the sum of the weights.

Percentages:

The screenshot shows the WinCross Quick Tools application window. The title bar reads "WinCross Quick Tools". The main header area contains the logo "WinCross Quick Tools" and "The Analytical Group Inc." logo. Below the header, there are navigation tabs: "Statistics and Frequency", "Significance Testing", "Sample Size Calculator", and "Calculator". The "Significance Testing" tab is active, and within it, the "Percentages" sub-tab is selected. Under "Percentages", there are four sub-tabs: "Two Independent Samples", "Two Categories in One Sample", "Proportion vs. Theoretical", and "Two Matched Samples". The "Two Independent Samples" sub-tab is active. The interface contains several input fields: "Percent #1 (p1) %", "Percent #2 (p2) %", "Sample size #1 (n1)", and "Sample size #2 (n2)". To the right of these fields is a text box explaining the test: "Here we test whether an observed proportion p1 based on a random sample of size n1 is significantly different from an observed proportion p2 based on an independent sample of size n2. The z statistic for testing this hypothesis is given by" followed by the formula:
$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$
. Below the formula, it states: "This tool reports the significance level and confidence level for both the two-tailed and one-tailed tests of significance of p1 from p2." At the bottom, there are fields for "z score", "Significance level" (with "Two tail" and "One tail" radio buttons), and "Confidence level".

WinCross approach: This **Quick Tool** cannot replicate the WinCross computation. That computation requires use of both the weighted percent (the quantity required for the percent box), the effective sample size, calculated as

$$f = \frac{\left(\sum_{i=1}^n w_i\right)^2}{\sum_{i=1}^n w_i^2}$$

for the sample size box, and an additional quantity, namely the UNWEIGHTED PERCENTAGE (which is used in the variance computation for weighted significance testing). Since this spreadsheet does not make provisions for introduction of the unweighted percentage, use of the weighted percentage throughout the internal computations of this spread sheet will not reproduce the (correct) WinCross computation.

SPSS approach: For each of the samples, fill in the percent box with the weighted percent and fill in the sample size box with the sum of the weights.