

## Which WinCross Statistical/Significance Test Should I Use?

Statistical Testing: Banner 1

☒ Means

Means Tests

**T-Tests**

- ☒ WinCross selects T-Test (default)
- ☐ Independent (based on test for equal variances)
- ☐ Independent (assume unequal variances)
- ☐ Independent (assume equal variances)
- ☐ Dependent Paired/Overlap (Multi)
- ☐ Dependent Paired/Overlap (LOC+/VAR+)

☒ I-Test Options...

**One-Way ANOVA**

- ☐ Least-significant difference
- ☐ Student Newman Keuls
- ☐ Kramer-Tukey B
- ☐ Kramer-Tukey
- ☐ Scheffe

One-Way ANOVA Options...

☒ Percents

Proportions Tests

**Z-Tests**

- ☒ WinCross selects Z-Test (default)
- ☐ Independent (using unpooled proportions)
- ☐ Independent (using pooled proportions)
- ☐ Dependent Paired/Overlap (Multi)
- ☐ Dependent Paired/Overlap (LOC+/VAR+)

☒ Z-Test Options...

☐ Chi-Square

Chi-Square Options...

OK Cancel ? Help

Don't let this array of choices overwhelm you when you are about to do statistical/significance testing. The appropriate test choice is determined primarily by your data's characteristics, and also by the organization of your banner. Moreover, if you would like to have WinCross do so, WinCross will, with one caveat to be described below, select the appropriate test choice, by your clicking the "WinCross selects T-test" and "WinCross selects Z-test" buttons. In fact, we have made these the "default" option, in that if you do not check another option WinCross will determine whether the appropriate choice is independent, Multi, or LOC+/VAR+, and perform the corresponding test. WinCross knows from the structure of the table that a "Dependent Paired" test is to be performed and whether Multi or LOC+/VAR+ is the appropriate test, and so does not need to be informed of this; when this test is called for, the WinCross default test (T or Z) will automatically perform it. (We retain the Multi and LOC+/VAR+ buttons in case you want to select the test.)

The one thing the "default" will not do is perform the test for whether or not to assume equal variances in the independent T-test or whether or not to use pooled proportions in the independent Z-test. The "default" option assumes unequal variances in the independent T-test and unpooled proportions in the independent Z-test. If you would like to override these defaults, you must select the appropriate option from the list of choices. And if you would like WinCross to decide whether or not to use the T-test based on equal variances or unequal variances, you must select that option from the list of choices.

If you want to decide for yourself what the appropriate test choice is, following are the considerations you must make. Whether comparing means or proportions, first you must determine whether the means or proportions based on the data for each of the banner points are **independent** or **dependent**. The means and/or proportions are independent if the data corresponding to each of the banner points *are in no way correlated* with each other. The means and/or proportions are dependent if the data corresponding to each of the banner points *are correlated* with each other.

For example, if all the data are based on a random sample and the banner is based on demographic breaks on such variables as age, income, gender, where the banner breaks are mutually exclusive then, the means are independent. On the other hand, if the sample consists of responses by both husbands and wives, and the banner consists of various age breaks, then the means are dependent, due to the correlation of ages between husband and wife. If the banner points consist of the various answers to multiple response questions, such as “Which soft drinks have you consumed in the last 3 months; Coke, Pepsi, Sprite?” the means are dependent, because the response data to each the various items defining the banner are made by the same people.

### **Comparing means**

Once you have determined that your means qualify as being **independent**, you still have to decide which of the three alternatives you should select, depending on what you know about the variances of the two groups of data. *(The variance of a population is the square of its standard deviation. Even if the means of the two groups of data are identical, the standard deviations may not be. Since different statistical are appropriate for situations in which the standard deviations of the two groups are or are not identical, WinCross provides multiple options to cover various contingencies.)* If you are sure the variances are **unequal** select: **Independent (assume unequal variances)**. If you are sure the variances are **equal** select: **Independent (assume equal variances)**. If you don’t know or are unsure of whether the two variances are equal, you can perform a preliminary statistical test of whether the variances are equal, and then perform the t-test that corresponds to whatever this preliminary test has determined. WinCross can do this test automatically, if you select: **Independent (based on test for equal variances)**.

Once you have determined that your means qualify as being **dependent**, you still have to determine whether your table is structured as **Multi** or **LOC+/VAR+**. A table is **Multi** if its banner is a set of attributes which qualify the respondent for the respective banner points, and whose rows are some other data about the respondents. For example, we want to compare the mean age of those who drank Coke with the mean age of those who drank Pepsi, and we set “drank Coke” and “drank Pepsi” as our banner points and the age breaks as the rows of our table. Since respondents can drink both Coke and Pepsi, the same respondent can appear in both columns of the table; hence the means are dependent. If the table is **Multi**, select the **Dependent Paired/Overlap (Multi)** test.

A table is **LOC+/VAR+** if its banner is a set of attributes which qualify the respondent for the respective banner points, and whose rows are separate measurements relating to the banner points. For example, we want to compare the mean rating of Coke from those who rated Coke with the mean rating of Pepsi from those who rated Pepsi, and we set “rated Coke” and “rated Pepsi” as our banner points and the rating scale as the rows of our table. Since respondents can rate both Coke and Pepsi, the same respondent can appear in both columns of the table; hence the means are dependent. If the table is **LOC+/VAR+**, select the **Dependent Paired/Overlap (LOC+)** test.

Since you can only select one of these five choices, what does one do if in a given table one wants to compare means from a pair of banner points with independent data and means from a pair of banner points with dependent data? The **Dependent Paired/Overlap** tests provide for the possibility that there are both independent and dependent data being analyzed. This could happen, for example, if we had data from a sample of Coke-only raters, a sample of Pepsi-only raters, and a sample of respondents who rated both. When there are no dependent data in the samples, the t-test reduces to the **Independent (assume unequal variances)** t-test. All one needs to do in this situation is determine whether the table is a **Multi** or **LOC+/VAR+** table and choose the appropriate **Dependent Paired/Overlap** test. The independent t-test, assuming unequal variances, will be applied to the pair of banner points with independent data and the dependent t-test will be applied to the pair of banner points with dependent data.

### **Comparing proportions**

Once you have determined that your proportions qualify as being **independent**, you still have to decide which of the two alternative tests you should select. (*One of the tests, the one using “pooled” proportions, argues that, since the null hypothesis that we are testing is that the two proportions are equal, we should pool the data from the two samples to create a pooled estimate of the common population proportion and use that estimate as the best basis for determining the standard error of the difference between the two sample proportions when the null hypothesis holds. The other test, the one using “unpooled” proportions, argues that, since the standard error of the difference between the two sample proportions is valid even in the event that the null hypothesis does not hold whereas that based on the pooled proportions is not, it is more appropriate.*) We recommend **using unpooled proportions** because statistical theory has shown this is the more powerful and robust test. Since many statistics textbooks still recommend the procedure in which one uses pooled proportions in the test procedure, WinCross has implemented this as an option, which one invokes by **using pooled proportions**.

Once you have determined that your proportions qualify as being **dependent**, you still have to determine whether your table is structured as **Multi** or **LOC+**. This determination is done in the same way as is the determination about means described above.