

Chapter 496

Confidence Intervals for Paired Means

Introduction

This routine calculates the sample size necessary to achieve a specified distance from the paired sample mean difference to the confidence limit(s) at a stated confidence level for a confidence interval about the mean difference when the underlying data distribution is normal.

Caution: This procedure assumes that the standard deviation of the future sample will be the same as the standard deviation that is specified. If the standard deviation to be used in the procedure is estimated from a previous paired sample or represents the population standard deviation, the Confidence Intervals for Paired Means with Tolerance Probability procedure should be considered. That procedure controls the probability that the distance from the mean paired difference to the confidence limits will be less than or equal to the value specified.

Technical Details

For a paired sample mean difference from a normal distribution with known variance, a two-sided, $100(1 - \alpha)\%$ confidence interval is calculated by

$$\bar{X}_{Diff} \pm \frac{z_{1-\alpha/2} \sigma_{diff}}{\sqrt{n}}$$

where \bar{X}_{Diff} is the mean of the paired differences of the sample, and σ_{diff} is the known standard deviation of paired sample differences.

A one-sided $100(1 - \alpha)\%$ upper confidence limit is calculated by

$$\bar{X}_{Diff} + \frac{z_{1-\alpha} \sigma_{diff}}{\sqrt{n}}$$

Similarly, the one-sided $100(1 - \alpha)\%$ lower confidence limit is

$$\bar{X}_{Diff} - \frac{z_{1-\alpha} \sigma_{diff}}{\sqrt{n}}$$

For a paired sample mean difference from a normal distribution with unknown variance, a two-sided, $100(1 - \alpha)\%$ confidence interval is calculated by

$$\bar{X}_{Diff} \pm \frac{t_{1-\alpha/2, n-1} \hat{\sigma}_{diff}}{\sqrt{n}}$$

Confidence Intervals for Paired Means

where \bar{X}_{Diff} is the mean of the paired differences of the sample, and $\hat{\sigma}_{diff}$ is the estimated standard deviation of paired sample differences.

A one-sided $100(1 - \alpha)\%$ upper confidence limit is calculated by

$$\bar{X}_{Diff} + \frac{t_{1-\alpha, n-1} \hat{\sigma}_{diff}}{\sqrt{n}}$$

Similarly, the one-sided $100(1 - \alpha)\%$ lower confidence limit is

$$\bar{X}_{Diff} - \frac{t_{1-\alpha, n-1} \hat{\sigma}_{diff}}{\sqrt{n}}$$

Each confidence interval is calculated using an estimate of the mean difference plus and/or minus a quantity that represents the distance from the mean difference to the edge of the interval. For two-sided confidence intervals, this distance is sometimes called the precision, margin of error, or half-width. We will label this distance, D .

The basic equation for determining sample size when D has been specified is

$$D = \frac{z_{1-\alpha/2} \sigma_{diff}}{\sqrt{n}}$$

when the standard deviation is known, and

$$D = \frac{t_{1-\alpha/2, n-1} \hat{\sigma}_{diff}}{\sqrt{n}}$$

when the standard deviation is unknown. These equations can be solved for any of the unknown quantities in terms of the others. The value $\alpha/2$ is replaced by α when a one-sided interval is used.

Finite Population Size

The above calculations assume that samples are being drawn from a large (infinite) population. When the population is of finite size (N), an adjustment must be made. The adjustment reduces the standard deviation as follows:

$$\sigma_{finite} = \sigma \sqrt{\left(1 - \frac{n}{N}\right)}$$

This new standard deviation replaces the regular standard deviation in the above formulas.

Confidence Level

The confidence level, $1 - \alpha$, has the following interpretation. If thousands of samples of n items are drawn from a population using simple random sampling and a confidence interval is calculated for each sample, the proportion of those intervals that will include the true population mean difference is $1 - \alpha$.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab

The Design tab contains most of the parameters and options that you will be concerned with.

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters.

One-Sided or Two-Sided Interval

Interval Type

Specify whether the interval to be used will be a one-sided or a two-sided confidence interval.

Population

Population Size

This is the number of pairs in the population. Usually, you assume that samples are drawn from a very large (infinite) population. Occasionally, however, situations arise in which the population of interest is of limited size. In these cases, appropriate adjustments must be made. This option sets the population size.

Confidence

Confidence Level

The confidence level, $1 - \alpha$, has the following interpretation. If thousands of samples of n items are drawn from a population using simple random sampling and a confidence interval is calculated for each sample, the proportion of those intervals that will include the true population mean difference is $1 - \alpha$.

Often, the values 0.95 or 0.99 are used. You can enter single values or a range of values such as *0.90, 0.95 or 0.90 to 0.99 by 0.01*.

Sample Size (Number of Pairs)

N (Sample Size)

Enter one or more values for the sample size. This is the number of pairs selected at random from the population to be in the study.

You can enter a single value or a range of values.

Precision

Distance from Mean Difference to Limit(s)

This is the distance from the confidence limit(s) to the mean paired difference. For two-sided intervals, it is also known as the precision, half-width, or margin of error.

You can enter a single value or a list of values. The value(s) must be greater than zero.

Confidence Intervals for Paired Means

Standard Deviation of Paired Differences

S (Standard Deviation)

Enter a value (or range of values) for the standard deviation. You can use the results of a pilot study, a previous study, or a ball park estimate based on the range (e.g., Range/4) to estimate this parameter.

Know Standard Deviation

Check this box when you want to base your results on the normal distribution. When the box is not checked, calculations are based on the t-distribution. The difference between the two distributions is negligible when the sample sizes are large (>50).

Example 1 – Calculating Sample Size

A researcher would like to estimate the mean difference in weight following a specific diet using a two-sided 95% confidence interval. The confidence level is set at 0.95, but 0.99 is included for comparative purposes. The standard deviation estimate, based on the range of paired differences, is 9.6 lbs. The researcher would like the interval to be no wider than 10 lbs. (half-width = 5 lbs.), but will examine half-widths of 3, 4, 5, 6, and 7 lbs.

The goal is to determine the necessary sample size.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Confidence Intervals for Paired Means** procedure window by expanding **Means**, then expanding **Paired Means**, and then clicking on **Confidence Interval**, and then clicking on **Confidence Intervals for Paired Means**. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

| <u>Option</u> | <u>Value</u> |
|--------------------------------------|--------------------|
| Design Tab | |
| Solve For | Sample Size |
| Interval Type | Two-Sided |
| Population Size | Infinite |
| Confidence Level | 0.95 0.99 |
| Distance from Mean to Limit(s) | 3 to 7 by 1 |
| S (Standard Deviation) | 9.6 |

Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results for Two-Sided Confidence Intervals with Unknown Standard Deviation

| Confidence Level | Sample Size (N) | Target Dist from Mean Diff to Limits | Actual Dist from Mean Diff to Limits | Standard Deviation (S) |
|--------------------|-----------------|--------------------------------------|--------------------------------------|------------------------|
| 0.95000 | 42 | 3.000 | 2.992 | 9.600 |
| 0.99000 | 72 | 3.000 | 2.995 | 9.600 |
| 0.95000 | 25 | 4.000 | 3.963 | 9.600 |
| 0.99000 | 43 | 4.000 | 3.950 | 9.600 |
| (report continues) | | | | |

Confidence Intervals for Paired Means

References

Hahn, G. J. and Meeker, W.Q. 1991. Statistical Intervals. John Wiley & Sons. New York.

Report Definitions

Confidence level is the proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that would contain the population mean difference.

N is the size of the sample (or number of pairs) drawn from the population.

Dist from Mean Diff to Limit is the distance from the confidence limit(s) to the mean paired difference. For two-sided intervals, it is also known as the precision, half-width, or margin of error.

Target Dist from Mean Diff to Limit is the value of the distance that is entered into the procedure.

Actual Dist from Mean Diff to Limit is the value of the distance that is obtained from the procedure.

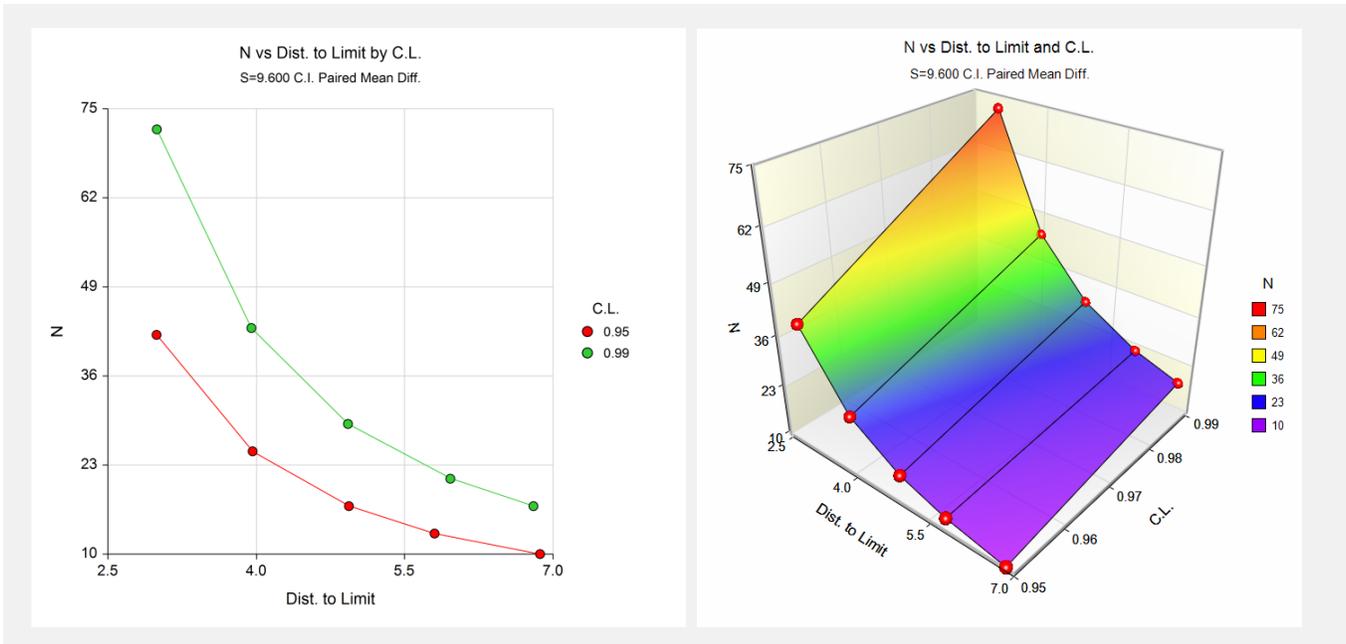
The standard deviation (S) is the standard deviation of the paired differences.

Summary Statements

A sample size of 42 produces a two-sided 95% confidence interval with a distance from the mean paired difference to the limits that is equal to 2.992 when the estimated standard deviation of the paired differences is 9.600.

This report shows the calculated sample size for each of the scenarios.

Plots Section



These plots show the sample size versus the precision for the two confidence limits.

Example 2 – Validation

This procedure uses the same mechanics as the Confidence Intervals for One Mean procedure. The validation of this procedure is given in Examples 2 and 3 of the Confidence Intervals for One Mean procedure.