Chapter 641

Confidence Intervals for One Standard Deviation with Tolerance Probability

Introduction

This procedure calculates the sample size necessary to achieve a specified width (or in the case of one-sided intervals, the distance from the standard deviation to the confidence limit) with a given tolerance probability at a stated confidence level for a confidence interval about a single standard deviation when the underlying data distribution is normal.

Technical Details

For a single standard deviation from a normal distribution with unknown mean, a two-sided, $100(1 - \alpha)\%$ confidence interval is calculated by

$$\left[s \left\{ \frac{n-1}{\chi^2_{1-\alpha/2,n-1}} \right\}^{1/2} , s \left\{ \frac{n-1}{\chi^2_{\alpha/2,n-1}} \right\}^{1/2} \right]$$

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

$$s \left\{ \frac{n-1}{\chi^2_{\alpha,n-1}} \right\}^{1/2}$$

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

$$s \left\{ \frac{n-1}{\chi^2_{1-\alpha,n-1}} \right\}^{1/2}$$

For two-sided intervals, the distance from the standard deviation to each of the limits is different. Thus, instead of specifying the distance to the limits we specify the width of the interval, W.

The basic equation for determining sample size for a two-sided interval when *W* has been specified is

$$W = s\left\{\frac{n-1}{\chi^2_{\alpha/2,n-1}}\right\}^{1/2} - s\left\{\frac{n-1}{\chi^2_{1-\alpha/2,n-1}}\right\}^{1/2}$$

For one-sided intervals, the distance from the standard deviation to limits, *D*, is specified.

The basic equation for determining sample size for a one-sided upper limit when *D* has been specified is

$$D = s\left\{\frac{n-1}{\chi^2_{\alpha/2,n-1}}\right\}^{1/2} - s$$

The basic equation for determining sample size for a one-sided lower limit when *D* has been specified is

$$D = s - s\left\{\frac{n-1}{\chi^2_{1-\alpha/2,n-1}}\right\}^{1/2}$$

These equations can be solved for any of the unknown quantities in terms of the others.

There is an additional subtlety that arises when the standard deviation is to be chosen for estimating sample size. The sample sizes determined from the formulas above produce confidence intervals with the specified widths only when the future sample has a sample standard deviation that is no greater than the value specified.

As an example, suppose that 15 individuals are sampled in a pilot study, and a standard deviation estimate of 3.5 is obtained from the sample. The purpose of a later study is to estimate the standard deviation with a confidence interval with width no greater than 1.3 units. Suppose further that the sample size needed is calculated to be 105 using the formula above with 3.5 as the estimate for the standard deviation. The sample of size 105 is then obtained from the population, but the standard deviation of the 105 individuals turns out to be 3.9 rather than 3.5. The confidence interval is computed and the width of the interval is greater than 1.3 units.

This example illustrates the need for an adjustment to adjust the sample size such that the width or distance from the standard deviation to the confidence limits will be below the specified value with known probability.

Such an adjustment for situations where a previous sample is used to estimate the standard deviation is derived for the case of confidence intervals for a mean by Harris, Horvitz, and Mood (1948) and discussed in Zar (1984) and Hahn and Meeker (1991). The adjustment is made by replacing *s* with

$$s = \sigma\sqrt{F_{1-\gamma;n-1,m-1}}$$

where $1 - \gamma$ is the probability that the width or distance from the standard deviation to the confidence limit will be below the specified value, and *m* is the sample size in the previous sample that was used to estimate the standard deviation.

The corresponding adjustment when no previous sample is available is discussed in Kupper and Hafner (1989) and Hahn and Meeker (1991). The adjustment in this case is

$$s = \sigma\left(\frac{\chi^2_{1-\gamma,n-1}}{n-1}\right)^{1/2}$$

where, again, $1 - \gamma$ is the probability that the width or distance from the standard deviation to the confidence limit will be below the specified value.

Each of these adjustments accounts for the variability in a future estimate of the standard deviation. In the first adjustment formula (Harris, Horvitz, and Mood, 1948), the distribution of the standard deviation is based on the estimate from a previous sample. In the second adjustment formula, the distribution of the standard deviation is based on a specified value that is assumed to be the population standard deviation.

For this procedure, both adjustments are adapted from the case of a one-sample confidence interval for a single mean to the case of a one-sample confidence interval for a single standard deviation.

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 standard deviation 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 two-sided confidence interval, an interval that has only an upper limit, or an interval that has only a lower limit.

Confidence and Tolerance

Confidence Level (1 – Alpha)

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 standard deviation 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*.

Tolerance Probability

This is the probability that a future interval with sample size *N* and the specified confidence level will have a width or distance from the standard deviation to the limit that is less than or equal to the width or distance specified.

If a tolerance probability is not used, as in the 'Confidence Intervals for One Standard Deviation using Standard Deviation' procedure, the sample size is calculated for the expected width or distance from the SD to the limit, which assumes that the future standard deviation will also be the one specified.

Using a tolerance probability implies that the standard deviation of the future sample will not be known in advance, and therefore, an adjustment is made to the sample size formula to account for the variability in the standard deviation. Use of a tolerance probability is similar to using an upper bound for the standard deviation in the 'Confidence Intervals for One Standard Deviation using Standard Deviation' procedure.

Values between 0 and 1 can be entered. The choice of the tolerance probability depends upon how important it is that the width distance from the interval limits to the standard deviation is at most the value specified.

You can enter a range of values such as *0.70 0.80 0.90* or *0.70 to 0.95 by 0.05*.

Sample Size

N (Sample Size)

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

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

Precision

Confidence Interval Width (Two-Sided)

This is the distance from the lower confidence limit to the upper confidence limit. The distance from the standard deviation to the lower and upper limits is not equal.

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

Distance from SD to Limit (One-Sided)

This is the distance from the standard deviation to the lower or upper limit of the confidence interval, depending on whether the Interval Type is set to Lower Limit or Upper Limit.

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

Standard Deviation

Standard Deviation Source

This procedure permits two sources for estimates of the standard deviation:

- **S is a Population Standard Deviation**

  This option should be selected if there is no previous sample that can be used to obtain an estimate of the standard deviation. In this case, the algorithm assumes that future sample obtained will be from a population with standard deviation S.

- **S from a Previous Sample**

  This option should be selected if the estimate of the standard deviation is obtained from a previous random sample from the same distribution as the one to be sampled. The sample size of the previous sample must also be entered under 'Sample Size of Previous Sample'.

Standard Deviation – Population Standard Deviation

S (Standard Deviation)

Enter an estimate of the standard deviation (must be positive). In this case, the algorithm assumes that future samples obtained will be from a population with standard deviation S.

One common method for estimating the standard deviation is the range divided by 4, 5, or 6.

You can enter a range of values such as *1 2 3* or *1 to 10 by 1*.

Press the Standard Deviation Estimator button to load the Standard Deviation Estimator window.

Standard Deviation – Standard Deviation from Previous Sample

S (SD Estimated from a Previous Sample)

Enter an estimate of the standard deviation from a previous (or pilot) study. This value must be positive.

A range of values may be entered.

Press the Standard Deviation Estimator button to load the Standard Deviation Estimator window.

Sample Size of Previous Sample

Enter the sample size that was used to estimate the standard deviation entered in S (SD Estimated from a Previous Sample).

This value is entered only when 'Standard Deviation Source:' is set to 'S from a Previous Sample'.

Example 1 – Calculating Sample Size

A researcher would like to estimate the standard deviation of a population with 95% confidence. It is very important that the interval width is less than 5 grams. Data available from a previous study are used to provide an estimate of the standard deviation. The estimate of the standard deviation is 25.4 grams, from a sample of size 12.

The goal is to determine the sample size necessary to obtain a two-sided confidence interval such that the width of the interval is less than 5 grams. Tolerance probabilities of 0.70 to 0.95 will be examined.

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 One Standard Deviation with Tolerance Probability** procedure window by expanding **Variances**, then clicking on **One Standard Deviation**, and then clicking on **Confidence Intervals for One Standard Deviation with Tolerance Probability**. 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**.

Option	**Value**
Design Tab	
Solve For	**Sample Size**
Interval Type	**Two-Sided**
Confidence Level	**0.95**
Tolerance Probability	**0.70 to 0.95 by 0.05**
Confidence Interval Width	**5**
Standard Deviation Source	**S from a Previous Sample**
S...	**25.4**
Sample Size of Previous Sample	**12**

Annotated Output

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

Numeric Results

Numeric Results for Two-Sided Confidence Intervals

Confidence Level	Sample Size (N)	Target Width	Actual Width	Standard Deviation (S)	Tolerance Probability
0.950	273	5.000	4.997	25.400	0.700
0.950	293	5.000	4.998	25.400	0.750
0.950	318	5.000	4.997	25.400	0.800
0.950	351	5.000	4.993	25.400	0.850
0.950	398	5.000	4.995	25.400	0.900
0.950	484	5.000	4.997	25.400	0.950

Sample size for estimate of S from previous sample = 12.

References
Hahn, G. J. and Meeker, W.Q. 1991. Statistical Intervals. John Wiley & Sons. New York.
Zar, J. H. 1984. Biostatistical Analysis. Second Edition. Prentice-Hall. Englewood Cliffs, New Jersey.
Harris, M., Horvitz, D. J., and Mood, A. M. 1948. 'On the Determination of Sample Sizes in Designing
    Experiments', Journal of the American Statistical Association, Volume 43, No. 243, pp. 391-402.

Report Definitions
Confidence Level is the proportion of confidence intervals (constructed with this same confidence level,
    sample size, etc.) that would contain the population standard deviation.
N is the size of the sample drawn from the population.
Width is the distance from the lower limit to the upper limit.
Target Width is the value of the width that is entered into the procedure.
Actual Width is the value of the width that is obtained from the procedure.
Standard Deviation (S) is the estimated standard deviation based on a previous sample.
Tolerance Probability is the probability that a future interval with sample size N and corresponding
   confidence level will have a width that is less than or equal to the specified width.

Summary Statements
The probability is 0.700 that a sample size of 273 will produce a two-sided 95% confidence
interval with a width that is less than or equal to 4.997 if the population standard deviation
is estimated to be 25.400 by a previous sample of size 12.

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

Plots Section
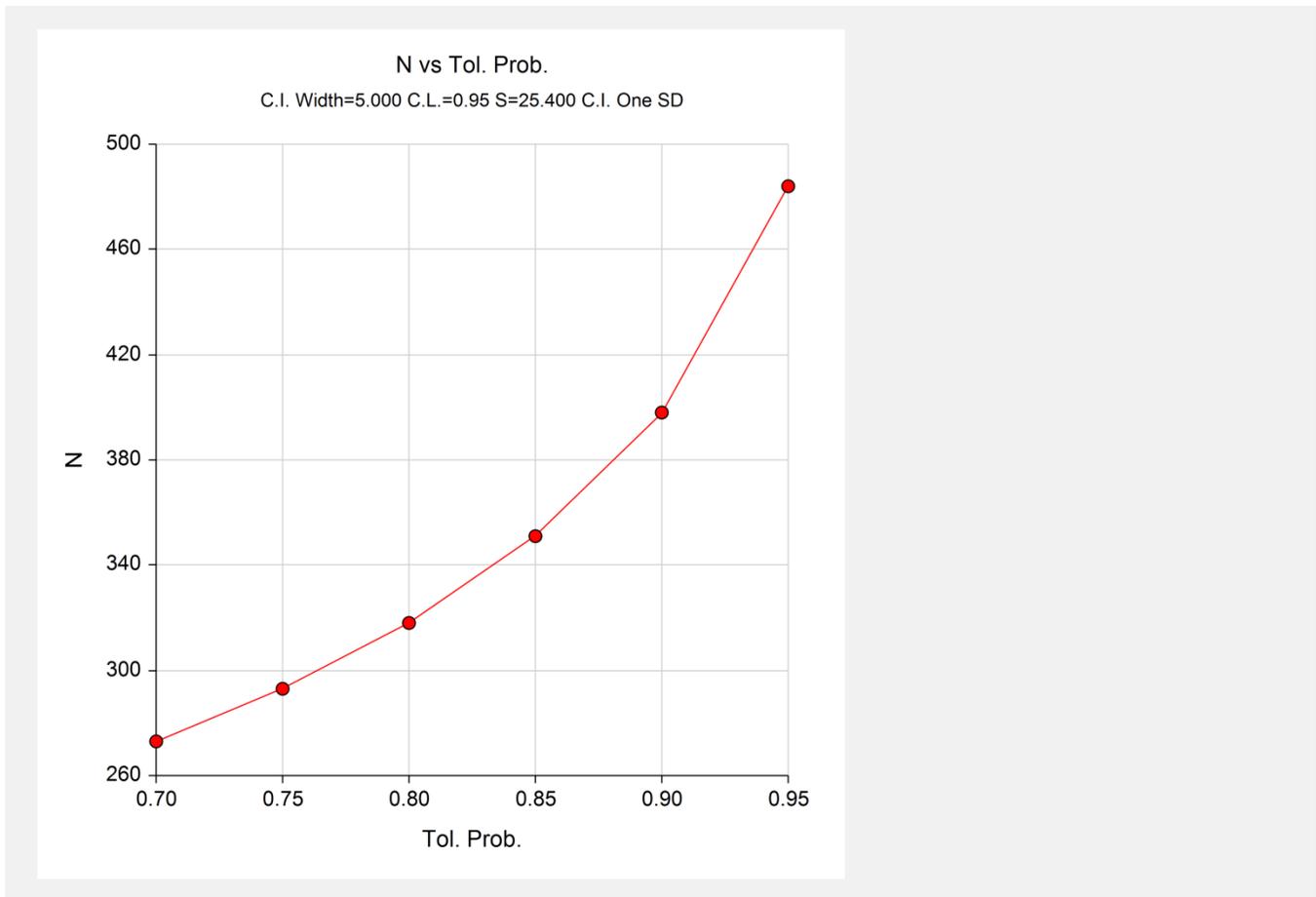
This plot shows the sample size versus the tolerance probability.

Example 2 – Validation using Simulation

We could not find a published result for confidence intervals for a standard deviation with tolerance probability. This procedure is validated using a simulation. A simulation was run with a confidence level of 95%, sample size of 100, a specified confidence interval width of 6, and assumed population standard deviation of 20.

The number of simulations was 100,000. The proportion of intervals with width less than 6 (tolerance probability) was 0.80168.

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 One Standard Deviation with Tolerance Probability** procedure window by expanding **Variances**, then clicking on **One Standard Deviation**, and then clicking on **Confidence Intervals for One Standard Deviation with Tolerance Probability**. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

Option   **Value**

Design Tab
Solve For .... **Tolerance Probability**
Interval Type .... **Two-Sided**
Confidence Level **0.95**
N (Sample Size) **100**
Confidence Interval Width **6**
Standard Deviation Source .... **S is a Population Standard Deviation**
S **20**

Output

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

Numeric Results

Numeric Results for Two-Sided Confidence Intervals

Confidence Level	Sample Size (N)	Target Width	Actual Width	Standard Deviation (S)	Tolerance Probability
0.950	100	6.000	6.000	20.000	0.802

PASS calculated the tolerance probability to be 0.802, which is well within the simulation error of the simulation estimate of 0.80168.