Chapter 404

Confidence Intervals for a Percentile of a Normal Distribution

Introduction

This procedure calculates the sample size necessary to achieve a specified interval width or distance from a sample percentile to the confidence limit at a stated confidence level when the underlying data distribution is Normal.

Caution: This procedure assumes that the percentile and standard deviation of the future sample will be the same as the percentile and standard deviation that is specified. If the sample percentile or standard deviation is different from those specified when running this procedure, the interval width (or distance to limit) may be narrower or wider than specified.

Technical Details

Chakraborti and Li (2007) compare several methods of confidence interval estimation of a Normal percentile. Their simulation studies showed four of the methods to behave almost identically. One of those methods, which they called the Lawless method (Lawless, 2003, p. 231), is the method used in this procedure. Hahn and Meeker (1991) and Odeh and Owen (1980) provide additional insights into this method.

For a single sample percentile $\hat{Y}_p$ from a Normal distribution with sample standard deviation $\hat{\sigma}$, a two-sided, 100(1 – α)% confidence interval is calculated by

$$\left(\bar{X} - t'_{1-\frac{\alpha}{2},[n-1,(-\sqrt{n}z_p)]}\frac{\hat{\sigma}}{\sqrt{n}}, \bar{X} - t'_{\frac{\alpha}{2},[n-1,(-\sqrt{n}z_p)]}\frac{\hat{\sigma}}{\sqrt{n}}\right)$$

with one-sided limits calculated by replacing $\frac{\alpha}{2}$ with α. t' is the non-central t distribution, with corresponding degrees of freedom and noncentrality parameter.

The width of the two-sided confidence interval is calculated as the difference between the upper and lower limits, or

$$W = \bar{X} - t'_{\frac{\alpha}{2},[n-1,(-\sqrt{n}z_p)]}\frac{\hat{\sigma}}{\sqrt{n}} - \left(\bar{X} - t'_{1-\frac{\alpha}{2},[n-1,(-\sqrt{n}z_p)]}\frac{\hat{\sigma}}{\sqrt{n}}\right)$$

$$= -t'_{\frac{\alpha}{2},[n-1,(-\sqrt{n}z_p)]}\frac{\hat{\sigma}}{\sqrt{n}} + t'_{1-\frac{\alpha}{2},[n-1,(-\sqrt{n}z_p)]}\frac{\hat{\sigma}}{\sqrt{n}}$$

With the sample percentile $\hat{Y}_p$ calculated as

$$\hat{Y}_p = \bar{X} + Z_p\hat{\sigma}$$

the distance from $\hat{Y}_p$ to the lower and upper limits is given as

$$D_{Lower} = \bar{X} + Z_p\hat{\sigma} - \left(\bar{X} - t'_{1-\frac{\alpha}{2},[n-1,(-\sqrt{n}Z_p)]}\frac{\hat{\sigma}}{\sqrt{n}}\right) = Z_p\hat{\sigma} + t'_{1-\frac{\alpha}{2},[n-1,(-\sqrt{n}Z_p)]}\frac{\hat{\sigma}}{\sqrt{n}}$$

and

$$D_{Upper} = \bar{X} - t'_{\frac{\alpha}{2},[n-1,(-\sqrt{n}Z_p)]}\frac{\hat{\sigma}}{\sqrt{n}} - \left(\bar{X} + Z_p\hat{\sigma}\right) = -t'_{\frac{\alpha}{2},[n-1,(-\sqrt{n}Z_p)]}\frac{\hat{\sigma}}{\sqrt{n}} - Z_p\hat{\sigma}$$

respectively.

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 individuals 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

This is the proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that would contain the true population percentile.

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

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

Distance from Percentile to Limit

This precision measure is the distance from the sample percentile estimate to the lower confidence limit (LCL) or the upper confidence limit (UCL). It is calculated using |Percentile - LCL| or |UCL - Percentile|. This can be any value greater than 0. You can enter a single value such as *3* or a series of values such as *2 3 4* or *2 to 4 by 0.5*.

Width of Confidence Interval

This precision measure is the distance between the lower confidence limit (LCL) and the upper confidence limit (UCL). It is calculated using UCL - LCL. This can be any value greater than 0. You can enter a single value such as *3* or a series of values such as *2 3 4* or *2 to 4 by 0.5*.

Percentile

Percentile Percentage

Enter the value of the desired percentile percentage. For example, entering 75 here would indicate you are examining a confidence interval for the 75th percentile. The range is between 0 and 100. You can enter a single value such as *30* or a series of values such as *30 40 50* or *30 to 70 by 5*.

Standard Deviation

Standard Deviation

Enter an estimate of the sample standard deviation. The sample size and width calculations assume that the value entered here is the standard deviation estimate that is obtained from the sample. This is sometimes refered to as a planning estimate. If the sample standard deviation is different from the value specified here, the actual confidence interval width will be different to the degree of the difference. This sample standard deviation can be any value greater than 0. You can enter a single value such as *3* or a series of values such as *2 3 4* or *2 to 4 by 0.5*.

Example 1 – Calculating Sample Size

Suppose a study is planned in which the researcher wishes to construct a two-sided 95% confidence interval for the 10th Percentile such that the width of the interval is no wider than 6 units. 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 data values, is 22.4. Instead of examining only the interval width of 6, a series of widths from 4 to 8 will also be considered.

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 a Percentile of a Normal Distribution** procedure window. 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 0.99**
Width of Confidence Interval	**4 to 8 by 0.5**
Percentile Percentage	**10**
Standard Deviation	**22.4**

Annotated Output

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

Numeric Results

Numeric Results for Two-Sided Confidence Intervals for a Percentile of a Normal Distribution

Confidence Level	Sample Size N	Target Width	Actual Width	Percentile Percentage	Sample Standard Deviation
0.950	881	4.000	4.000	10	22.4
0.990	1521	4.000	3.999	10	22.4
0.950	697	4.500	4.499	10	22.4
0.990	1203	4.500	4.499	10	22.4
0.950	565	5.000	5.000	10	22.4
0.990	975	5.000	5.000	10	22.4
0.950	468	5.500	5.497	10	22.4
0.990	807	5.500	5.498	10	22.4
0.950	394	6.000	5.995	10	22.4
0.990	679	6.000	5.997	10	22.4
0.950	336	6.500	6.496	10	22.4
0.990	579	6.500	6.499	10	22.4
0.950	290	7.000	6.998	10	22.4
0.990	500	7.000	6.998	10	22.4
0.950	253	7.500	7.499	10	22.4
0.990	436	7.500	7.499	10	22.4
0.950	223	8.000	7.994	10	22.4
0.990	384	8.000	7.997	10	22.4

Confidence Intervals for a Percentile of a Normal Distribution

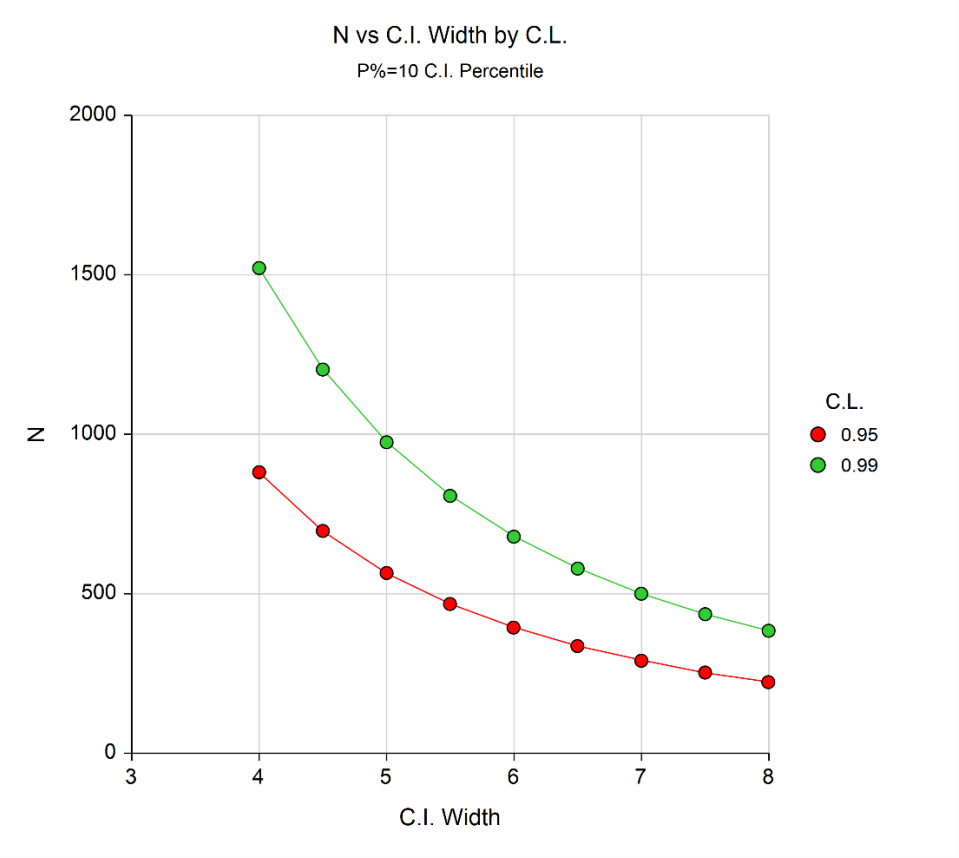
Report Definitions
Confidence level is the proportion of confidence intervals (constructed with this same confidence level, sample size, etc.) that would contain the true percentile.
Sample Size N is the size of the sample drawn from the population.
Width is the distance from the lower limit to the upper 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.
Percentile Percentage is the percent of the observations that falls below the percentile value. For example, a value of 70 indicates the 70th percentile is of interest.
Sample Standard Deviation is the anticipated sample standard deviation.

References
Chakraborti, S., and Li, J. 2007. 'Confidence Interval Estimation of a Normal Percentile.' The American Statistician, Volume 61, No. 4, pages 1-6.
Lawless, Jerald F. 2003. Statistical Models and Methods for Lifetime Data, 2nd Edition. John Wiley, New York.
Hahn, G. J. and Meeker, W.Q. 1991. Statistical Intervals. John Wiley & Sons. New York.
Odeh, R.E. and Owen, D.B. 1980. Tables for Normal Tolerance Limits, Sampling Plans, and Screening. Marcel Dekker, Inc. New York, NY.

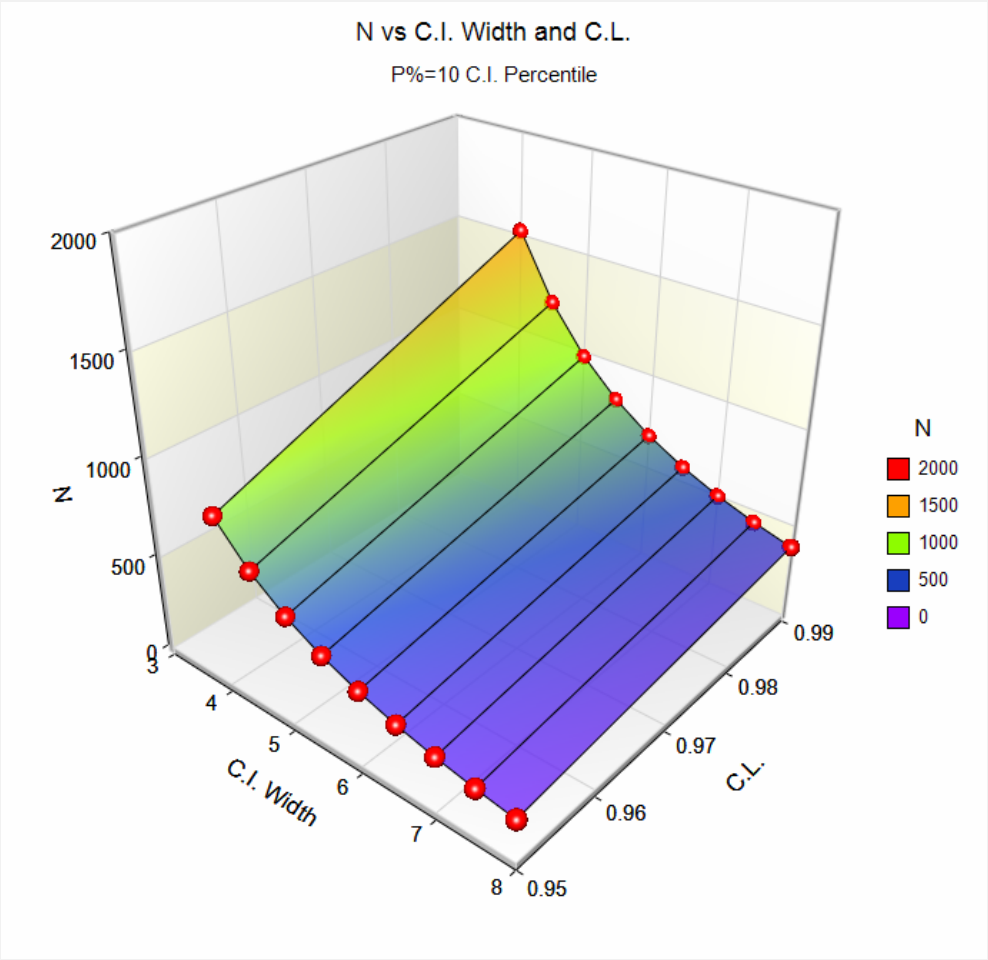
Summary Statements
A sample size of 881 produces a two-sided 95% confidence interval with a width equal to 4.000 when the percentile percentage is 10 and the sample standard deviation is 22.4.

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

Chart Section



Confidence Intervals for a Percentile of a Normal Distribution



These plots show the sample size versus the confidence interval widths for the two confidence levels.

Example 2 – Validation using Hahn and Meeker (1991)

Hahn and Meeker (1991) page 57 give an example of an interval calculation for a confidence interval for the 10th percentile when the confidence coefficient is 95%, the standard deviation is 1.31, and the confidence interval width is 4.95. The sample size given is 5.

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 a Percentile of a Normal Distribution** procedure window. 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**
Width of Confidence Interval	**4.95**
Percentile Percentage	**10**
Standard Deviation	**1.31**

Output

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

Numeric Results

Numeric Results for Two-Sided Confidence Intervals for a Percentile of a Normal Distribution

Confidence Level	Sample Size N	Target Width	Actual Width	Percentile Percentage	Sample Standard Deviation
0.950	5	4.950	4.949	10	1.31

PASS also calculated the sample size to be 5.