Chapter 420

Confidence Intervals for One Mean

Introduction

This routine calculates the sample size necessary to achieve a specified distance from the mean to the confidence limit(s) at a stated confidence level for a confidence interval about the mean 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 sample or represents the population standard deviation, the Confidence Intervals for One Mean with Tolerance Probability procedure should be considered. That procedure controls the probability that the distance from the mean to the confidence limits will be less than or equal to the value specified.

Technical Details

For a single mean from a normal distribution with known variance, a two-sided, 100(1 – α)% confidence interval is calculated by

$$\bar{X} \pm \frac{z_{1-\alpha/2}\sigma}{\sqrt{n}}$$

A one-sided 100(1 – α)% upper confidence limit is calculated by

$$\bar{X} + \frac{z_{1-\alpha}\sigma}{\sqrt{n}}$$

Similarly, the one-sided 100(1 – α)% lower confidence limit is

$$\bar{X} - \frac{z_{1-\alpha}\sigma}{\sqrt{n}}$$

For a single mean from a normal distribution with unknown variance, a two-sided, 100(1 – α)% confidence interval is calculated by

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

A one-sided 100(1 – α)% upper confidence limit is calculated by

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

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

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

Each confidence interval is calculated using an estimate of the mean plus and/or minus a quantity that represents the distance from the mean 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

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

when the standard deviation is known, and

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

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:

$$\sigma_{finite} = \sigma\sqrt{1 - \frac{n}{N}}$$

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 is $1 - \alpha$.

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 mean such that the width of the interval is no wider than 14 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 28. Instead of examining only the interval half-width of 7, a series of half-widths from 5 to 9 will also be considered.

The goal is to determine the necessary sample size.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For ....**Sample Size**
Interval Type ....**Two-Sided**
Population Size ....**Infinite**
Confidence Level (1 - Alpha) **0.95 0.99**
Distance from Mean to Limit(s) ....**5 to 9 by 1**
S (Standard Deviation)..................................**28**
Known Standard Deviation ....**Unchecked**

Output

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

Numeric Reports

Numeric Results
―――

Solve For: Sample Size
Interval Type: Two-Sided
Standard Deviation: Unknown
―――

Confidence Level	Sample Size N	Distance from Mean to Limits		Standard Deviation S
		Target	Actual	
0.95	123	5	4.998	28
0.99	212	5	4.999	28
0.95	87	6	5.968	28
0.99	149	6	5.986	28
0.95	64	7	6.994	28
0.99	110	7	6.999	28
0.95	50	8	7.958	28
0.99	86	8	7.956	28
0.95	40	9	8.955	28
0.99	69	9	8.933	28

―――

Confidence Level   The proportion of confidence intervals (constructed with this same confidence level, sample
                                 size, etc.) that would contain the population mean.
N   The size of the sample drawn from the population.
Distance from Mean to Limits The distance from the confidence limit(s) to the mean. For two-sided intervals, it is also known
                                 as the precision, half-width, or margin of error.
Target Distance   The value of the distance that is entered into the procedure.
Actual Distance   The value of the distance that is obtained from the procedure.
S   The standard deviation of the population measures the variability in the population.

Summary Statements
―――

A single-group design will be used to obtain a two-sided 95% confidence interval for a single mean. The standard t
formula will be used to calculate the confidence interval. The estimated standard deviation is assumed to be 28. To
produce a confidence interval with a distance of no more than 5 from the sample mean to either limit, 123 subjects
will be needed.
―――

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	123	154	31
20%	212	265	53
20%	87	109	22
20%	149	187	38
20%	64	80	16
20%	110	138	28
20%	50	63	13
20%	86	108	22
20%	40	50	10
20%	69	87	18

Dropout Rate	The percentage of subjects (or items) that are expected to be lost at random during the course of the study and for whom no response data will be collected (i.e., will be treated as "missing"). Abbreviated as DR.
N	The evaluable sample size at which the confidence interval is computed. If N subjects are evaluated out of the N' subjects that are enrolled in the study, the design will achieve the stated confidence interval.
N'	The total number of subjects that should be enrolled in the study in order to obtain N evaluable subjects, based on the assumed dropout rate. After solving for N, N' is calculated by inflating N using the formula N' = N / (1 - DR), with N' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. (2018) pages 32-33.)
D	The expected number of dropouts. D = N' - N.

Dropout Summary Statements

Anticipating a 20% dropout rate, 154 subjects should be enrolled to obtain a final sample size of 123 subjects.
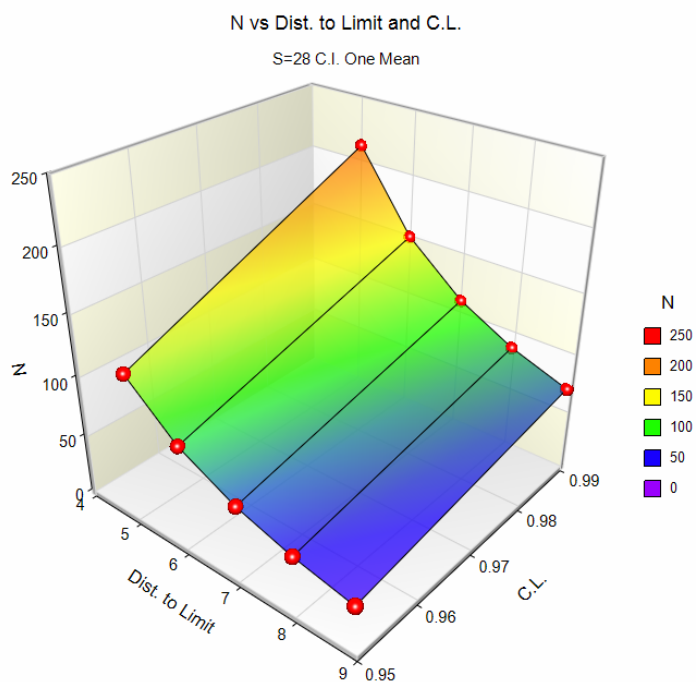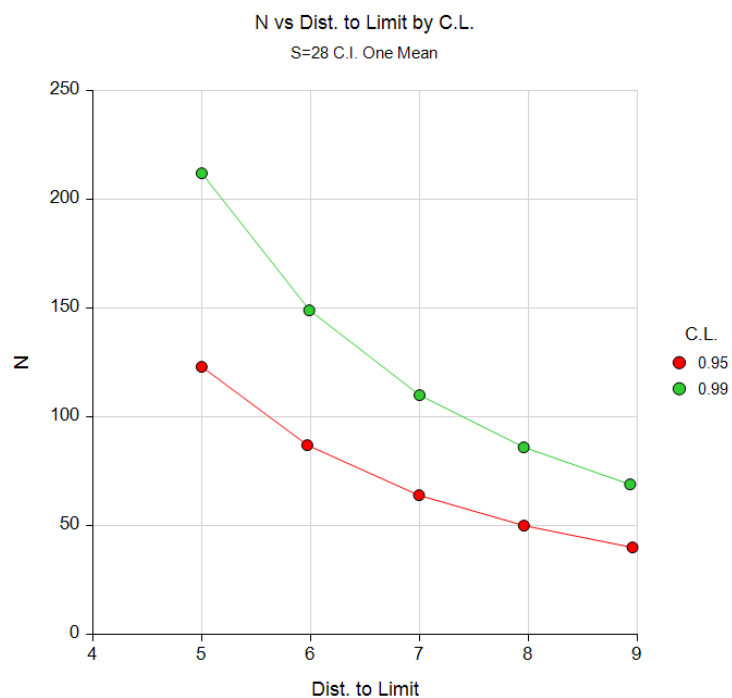
References

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

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

Plots Section

Plots
_____

These plots show the sample size versus the distance from the mean to the limits (precision) for the two confidence levels.

Example 2 – Validation using Moore and McCabe (1999)

Moore and McCabe (1999) page 443 give an example of a sample size calculation for a confidence interval on the mean when the confidence coefficient is 95%, the standard deviation is known to be 3, and the margin of error is 2. The necessary sample size is 9.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For ....**Sample Size**
Interval Type ....**Two-Sided**
Population Size...**Infinite**
Confidence Level (1 - Alpha) **0.95**
Distance from Mean to Limit(s) **2**
S (Standard Deviation)...................................**3**
Known Standard Deviation ....**Checked**

Output

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

Numeric Results
_____

Solve For: Sample Size
Interval Type: Two-Sided
Standard Deviation: Known
_____

Confidence Level	Sample Size N	Distance from Mean to Limits		Standard Deviation S
		Target	Actual	
0.95	9	2	1.96	3

PASS also calculated the necessary sample size to be 9.

Example 3 – Validation using Ostle and Malone (1988)

Ostle and Malone (1988) page 536 give an example of a sample size calculation for a confidence interval on the mean when the confidence coefficient is 95%, the standard deviation is known to be 7, and the margin of error is 5. The necessary sample size is 8.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab
_____

Solve For ....**Sample Size**
Interval Type ....**Two-Sided**
Population Size..**Infinite**
Confidence Level (1 - Alpha) **0.95**
Distance from Mean to Limit(s) **5**
S (Standard Deviation)..................................**7**
Known Standard Deviation ....**Checked**

Output

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

Numeric Results
_____

Solve For: Sample Size
Interval Type: Two-Sided
Standard Deviation: Known
_____

Confidence Level	Sample Size N	Distance from Mean to Limits Target	Distance from Mean to Limits Actual	Standard Deviation S
0.95	8	5	4.851	7

PASS also calculated the necessary sample size to be 8.