

## Chapter 574

# Confidence Intervals for a Percentile of a Normal Distribution using Assurance Probability

---

### Introduction

This procedure calculates the sample size necessary to achieve an exact, equal error per tail confidence interval width for a percentile (often called the reference limit) at a stated confidence level when the underlying data distribution is normal. It bases the optimal sample size on the assurance probability.

---

### Technical Details

The methods used in this procedure come from Shieh (2018). This article provides useful insights into why one should use an alternative to the classic Bland-Altman interval.

---

### Confidence Interval

Assume that you are planning a study of a population whose distribution is reasonably close to  $N(\mu, \sigma^2)$  to create a confidence interval about the 100 $p$ th percentile of this population.

Let  $X$  represent the variable of interest. Suppose that a random sample of  $N$  observations will be obtained and the usual mean and standard deviation  $(\bar{X}, S)$  will be used to estimate the 100 $p$ th percentile of  $X$  which is denoted by  $\theta$ , where

$$\theta = \mu + z_p \sigma$$

Here,  $z_p$  is the 100 $p$ th percentile of the standard normal distribution  $N(0,1)$ .

## Confidence Intervals for a Percentile of a Normal Distribution using Assurance Probability

An exact two-sided,  $100(1 - \alpha)\%$  confidence interval is calculated using

$$\left( \bar{X} + t'_{\frac{\alpha}{2}, [N-1, (\sqrt{N}Z_p)]} \frac{S}{\sqrt{N}}, \bar{X} + t'_{1-\frac{\alpha}{2}, [N-1, (\sqrt{N}Z_p)]} \frac{S}{\sqrt{N}} \right)$$

with one-sided limits calculated by replacing  $\frac{\alpha}{2}$  with  $\alpha$ .  $t'$  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

$$W = \left( t'_{1-\frac{\alpha}{2}, [N-1, (\sqrt{N}Z_p)]} \frac{S}{\sqrt{N}} - t'_{\frac{\alpha}{2}, [N-1, (\sqrt{N}Z_p)]} \frac{S}{\sqrt{N}} \right)$$

---

## Sample Size Based on the Assurance Probability

The method of determining a useful sample size used in this procedure is determine an  $N$  that guarantees with specified assurance probability no less than  $(1 - \gamma)$  that the actual width of the confidence interval no greater than a boundary value,  $\omega$ . That is, we select  $N$  so that  $\Pr(W \leq \omega) \geq 1 - \gamma$ . This leads to the following expression from which the desired  $N$  can be determined by a simple search.

$$k = \frac{N(N-1) \left( \frac{\omega}{\sigma} \right)^2}{\left\{ t'_{1-\frac{\alpha}{2}, [N-1, (\sqrt{N}Z_p)]} - t'_{\frac{\alpha}{2}, [N-1, (\sqrt{N}Z_p)]} \right\}^2}$$

Now, we find the smallest  $N$  for which

$$X_{N-1}^2(k) \geq 1 - \gamma$$

where  $X_{N-1}^2(k)$  is the cumulative chi-square distribution with degrees of freedom  $N - 1$ .

## 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 90<sup>th</sup> percentile such that the interval width is guaranteed to be no wider than 6 units with probability 0.90, 0.95, or 0.99. The standard deviation is between 21 and 23.

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 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**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size (N)</b>
Confidence Level (1 - $\alpha$ ) .....	<b>0.95</b>
1 - $\gamma$ (Lower Bound of Assurance Probability) .....	<b>0.9 0.95 0.99</b>
$\omega$ (Upper Bound of the C.I. Width) .....	<b>6</b>
p (Percentile Proportion).....	<b>0.9</b>
$\sigma$ (Standard Deviation) .....	<b>21 22 23</b>

### Annotated Output

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

#### Numeric Results

<b>Numeric Results</b>							
Solve For:		Sample Size (N)					
Assurance Requirement:		$\Pr(W \leq \omega) \geq 1 - \gamma$					
Interval Type:		Two-Sided Confidence Interval					
Confidence Level <b>1 - <math>\alpha</math></b>	Sample Size <b>N</b>	Percentile Proportion <b>p</b>	Upper Bound of the C.I. Width <b><math>\omega</math></b>	Lower Bound of Assurance Probability		Standard Deviation <b><math>\sigma</math></b>	
				Target <b>1 - <math>\gamma_T</math></b>	Actual <b>1 - <math>\gamma_A</math></b>		
0.95	379	0.9	6	0.90	0.9038	21	
0.95	414	0.9	6	0.90	0.9035	22	
0.95	451	0.9	6	0.90	0.9059	23	
0.95	388	0.9	6	0.95	0.9515	21	
0.95	424	0.9	6	0.95	0.9534	22	
0.95	461	0.9	6	0.95	0.9531	23	
0.95	405	0.9	6	0.99	0.9903	21	
0.95	442	0.9	6	0.99	0.9909	22	
0.95	479	0.9	6	0.99	0.9901	23	

## Confidence Intervals for a Percentile of a Normal Distribution using Assurance Probability

### Report Definitions

$1 - \alpha$  is the confidence level of the confidence interval of the percentile.

$N$  is the sample size of the study.

$p$  is the percentile proportion. It is the proportion of observations that fall at or below the 100 $p$ th percentile value.

For example, a value of 0.7 indicates the 70th percentile.

$\omega$  is an upper bound on the c.i. width. The sample size guarantees (assures) that  $100(1 - \gamma)\%$  of interval widths will be less than this value.

$1 - \gamma_T$  is a lower bound on the target (planned) assurance probability. The assurance probability is the probability that the confidence interval width will be less than  $\omega$ .

$1 - \gamma_A$  is a lower bound on the actual assurance probability achieved by this sample size. It may be different from the target value because of the discrete nature of  $N$ .

$\sigma$  is the standard deviation of the population about which the percentile is being constructed.

### References

Shieh, Gwopen. 2018. 'The appropriateness of Bland-Altman's approximate confidence intervals for limits of agreement.' BMC Medical Research Methodology. 18,45,1. <https://doi.org/10.1186/s12874-018-0505-y>

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

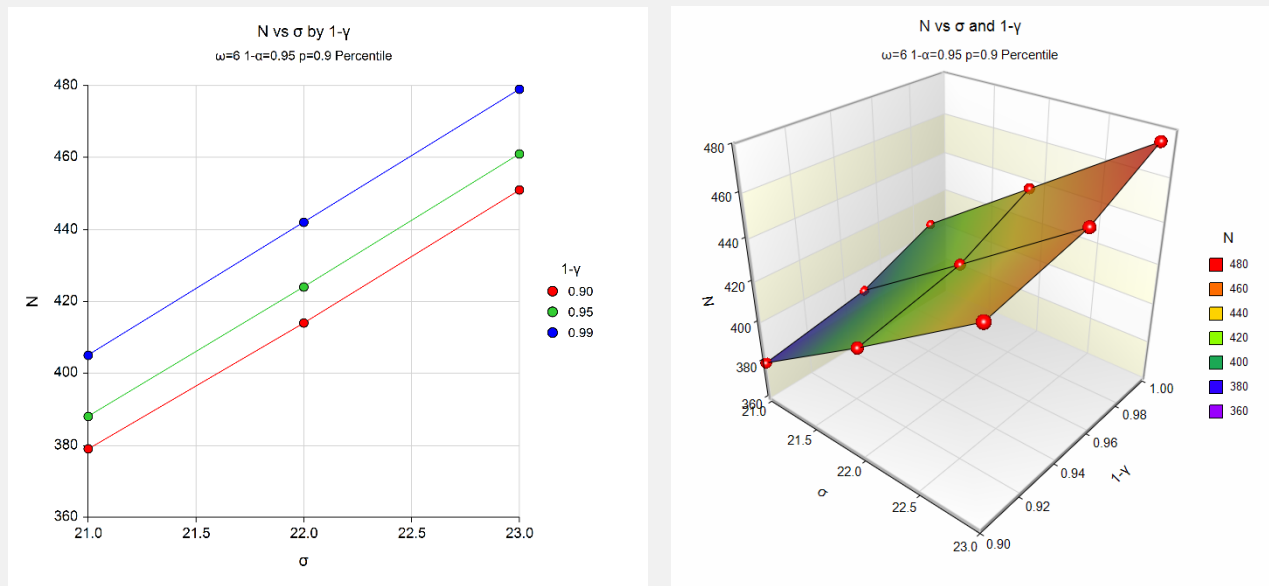
### Summary Statements

A sample size of 379 produces a two-sided 95% confidence interval of the 90th percentile with a width no more than 6 with probability 0.9. The standard deviation is 21.

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

## Chart Section

### Chart Section



These plots show the sample sizes required for the various scenarios.

## Example 2 – Validation using Shieh (2018)

Shieh (2018) page 7 gives an example of an exact interval calculation for a confidence interval for the 97.5<sup>th</sup> percentile when the confidence coefficient is 95%, the standard deviation is 19.61, the assurance level is 0.90, and the confidence interval width boundary is 9.805. The sample size is 207.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. 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**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size (N)</b>
Confidence Level (1 - $\alpha$ ) .....	<b>0.95</b>
1 - $\gamma$ (Lower Bound of Assurance Probability) .....	<b>0.9</b>
$\omega$ (Upper Bound of the C.I. Width) .....	<b>9.805</b>
p (Percentile Proportion).....	<b>0.975</b>
$\sigma$ (Standard Deviation) .....	<b>19.61</b>

### Output

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

#### Numeric Results

Numeric Results						
Solve For:		Sample Size (N)				
Assurance Requirement:		$\Pr(W \leq \omega) \geq 1 - \gamma$				
Interval Type:		Two-Sided Confidence Interval				
Confidence Level $1 - \alpha$	Sample Size N	Percentile Proportion p	Upper Bound of the C.I. Width $\omega$	Lower Bound of Assurance Probability		Standard Deviation $\sigma$
				Target $1 - \gamma_T$	Actual $1 - \gamma_A$	
0.95	207	0.975	9.805	0.9	0.9065	19.61

The sample size computed by **PASS** is also 207. This validates the procedure.