

Documentation for Confidence Interval for a sample mean

Minn M. Soe, MD, MPH, MCTM : msoe@sph.emory.edu
 Kevin M. Sullivan, PhD, MPH, MHA: cdckms@sph.emory.edu

This module calculates the confidence interval for a sample mean. First, the user is prompted to enter the sample mean, standard deviation, sample size, population size from which the sample was taken (if known) and desired confidence interval. Although default values are provided, the user can change these parameter values as required.

As the confidence interval of sample mean takes into account "correction factor for finite population", population size needs to be entered. If population size is unknown or can not be estimated, set the value at 999,999,999. Please note that the sample size should be less than or equal to population size. Likewise, confidence intervals ranging from 20% - 99.99% can be entered, specifically, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99, 99.5, 99.8, 99.9, 99.95, 99.98 & 99.99.

Calculate	Confidence Intervals for a Sample Mean						
Clear	Sample Mean	150					
	Sample Std. Deviation	27.36	Std. Error		Variance		<i>Enter only one</i>
	Sample Size	268					
	Population Size	999999999					
	Confidence Level (%)	95					

The output from the default example is as below:

Confidence Intervals for a Sample Mean					
Input Data					
Sample Mean	150				
Sample Std. Deviation	27.36	Std. Error	1.67128	Variance	748.57
Sample size	268				
Population size	999999999				
Confidence Interval	95	% (100)			
95% Confidence Limits for the Mean of 150					
	Based on:		Lower Limit		Upper Limit
	z-test		146.725		153.275
	t-test		146.709		153.291
Results from OpenEpi open source calculator--CIMean					
file:///C:/OpenEpi%20March%2011%202005/Mean/CIMean.htm					
Source file last modified on 04/14/2005 19:11:33					
Print from the browser, or select all or part of the text and then copy and paste to other programs.					
Many browsers have an optional setting to print background colors.					

The interpretation is that we are 95% confident that the true mean of cholesterol in the population is captured between 146.725 mg% and 153.275 mg%. The confidence interval values shown in the first row of results are based on z-score according to large sample size theory (>30). If the sample size is small, confidence intervals based on t-test should be chosen. In our example, there is not much difference between the results based on z-test and t-test, because of large sample size (n=268). Currently, all confidence intervals calculated are two-sided confidence intervals. The formulae for the methods are provided below.

Formulae

The notation for the formulae are:

SE = standard error

σ = standard deviation of a sample

N = population size

n = sample size

CI = confidence limit

$Mean$ = sample mean

$Z_{1-\alpha/2}$ = the two-sided Z value (eg. $Z=1.96$ for 95% confidence interval).

$t_{n-1,\alpha/2}$ = the two-sided t value with $df = n-1$

$$SE = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{(N - n)}{(N - 1)}}$$

$$CI = Mean \pm Z_{1-\alpha/2} * SE$$

$$CI = Mean \pm T_{n-1, \alpha/2} * SE$$

Acknowledgement:

Default values are derived from an example of EpiTable in EPI6 (DOS version).

References:

"Normal approximation" method described in Bernard PM, Lapointe C, Measures Statistiques en e'pide'miologie, Presses de L' Universite' du Que'bec, 1987, pp 277-282.