Primer on Confidence Intervals & Margin of Error

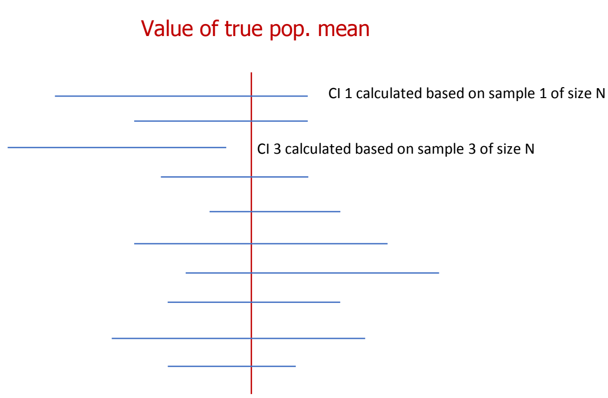
BC Office of Patient-Centred Measurement & Improvement

What is a Confidence Interval (CI)?

A confidence interval (CI) is an *interval estimate* of a population value, such as the population mean. This can be contrasted with *point estimates* of population values. For example, a sample mean can be used as a *point estimate* of the population mean and a CI built around a sample mean can be used as an *interval estimate* of the population mean. A two-sided CI will have a lower limit and an upper limit. We estimate that it is likely that the value of the true population mean falls somewhere within these limits. But, in reality, CIs may or may not contain the true population values.

We won't know with complete certainty whether a CI captures a true population value; however, we can talk about the likelihood of capturing true population values using *confidence levels***.** These are used to quantify the likelihood that the true population value lies inside the interval. Confidence levels are set based on the amount of confidence that a researcher or analyst desires for their estimates and is not set by the data. Confidence levels are typically set at 90%, 95%, or 99%.

It is important to note that likelihoods associated with confidence levels are provided in terms of *long-run probabilities.* To understand long-run probabilities, imagine that you were to draw an infinite number of random samples of the same size from a population of interest. Now, imagine that a 95% CI for the population mean were calculated for each of these samples drawn. This might look something like the figure below:



In the figure above, the red vertical line represents the value of the true target population mean (an unknown). Each of the horizontal lines represents a 95% CI calculated on different samples of data of the same size drawn repeatedly and infinitely from the target population. We can say that we are confident that 95% of the CIs calculated based on samples drawn from the target population will capture the true population mean, but 5% won't (e.g., CI 3 does not capture the true population mean). We hope that ours will be one of the CIs that falls within the 95%!

What is Margin of Error?

Margin of Error (MOE) is the maximum expected difference between the true population value and a sample estimate of that value due to sampling error. It describes the range that a true population value likely falls between if we had surveyed everyone in that target population (instead of just a sample). In other words, it describes the width of our CI. There is a direct relationship between CIs, confidence levels, and MOE. As confidence levels increase, margin of error increases, which will also result in a wider CI. MOE therefore provides an estimate of the imprecision inherent in survey data. The larger the MOE, the less confidence one should have that the reported results are close to "true" population figures.

Each response and estimate has its own margin of error. The overall margin of errors in the reports are calculated using the total number of respondents who answered a certain percentage of questions (the exact cut-off vary by sectors) and assume a percent response of 50%.

What is Considered "Good" Margin of Error?

A margin of error of ±5% is considered good while ±8% is acceptable.

Factors that affect Margin of Error

MOE is affected by three factors:

- It is positively related to the width of the confidence interval
- It is inversely related to sample size
- It varies by response variation (e.g., the MOE of a percentage is largest when the percent response to a yes/no item is at 50% and decreases as the percent of "yes" responses moves towards 0% or 100%)

How to calculate Margin of Error?

Standard formulas are available to calculate two commonly reported margin of errors:

- the margin of error of the mean and;
- the margin of error of the percentage.

A margin of error can be computed with or without the finite population correction factor (fpc). The fpc factor assumes a finite population size and reduces the MOE when more of the population is included in the sample. If the population size is unknown, it can be assumed as infinite and the fpc would not be calculated.

General Form	Critical value	×	Standard error	×	(fpc)
Percent	1.96	×	$\frac{\sqrt{p-(1-p)}}{n}$	×	$\sqrt{\frac{(N-n)}{(N-1)}}$
Mean	1.96	×	$\frac{\sigma}{\sqrt{n}}$		

Example Calculation and Statement

Example Calculation: Consider a survey with 19,500 respondents sampled from a patient population of 40,000, with a 95% confidence level and assuming a top-box score of 50%. The MOE with the finite population correction factor applied is 0.5%.

$$MOE = 1.96\, x\, \frac{\sqrt{.5-(1-.5)}}{19500}\, x \sqrt{\frac{(40000-19500)}{(19500-1)}}$$
$$= 0.00005 * 100 = 0.5\%$$

Example Statement: Approximately 67.4% of respondents indicated they would "definitely" recommended the hospital they stayed at to their friends or family, with a margin of error of ±0.5% at 95% confidence level.

Interpretation: If we repeated the survey many times with sample of the same size using the same method, 95% of the time, the actual percentage of acute care patients in BC who would "definitely" recommended their hospital would lies between 66.9% to 67.9%.

The true value is expected to be outside of the interval, 5% of the time. But, the true value is either in or outside of the interval