

Some Misconceptions about Confidence Intervals

By Keith M. Bower

Reprinted with permission from the American Society for Quality

Six Sigma practitioners often require a reasonable range of values for some characteristic of a process (e.g., the process mean, μ). The confidence interval (CI) procedure, developed by Jerzy Neyman in the early 1930s,¹ is typically employed. Confidence intervals (CIs) for the population mean (μ) or standard deviation (σ) are widely used in manufacturing and business.

This article discusses three misconceptions associated with confidence intervals:

1. Overlapping confidence intervals imply “no difference.”
2. Confidence intervals are always appropriate.
3. Confidence intervals address the spread of individual values.

The CI Procedure

Consider a population that is well modeled by a normal (Gaussian) distribution. To obtain a reasonable estimate of the arithmetic mean for the population, μ , assume a sample of size n is randomly drawn from the population. The sample mean (\bar{x}) and sample standard deviation (s) are calculated. As discussed by Robert V. Hogg and Eliot Tanis,² a CI for μ may be constructed using the formula:

$$\left\{ \bar{x} - t_{\frac{\alpha}{2}}(n-1) \left(\frac{s}{\sqrt{n}} \right), \bar{x} + t_{\frac{\alpha}{2}}(n-1) \left(\frac{s}{\sqrt{n}} \right) \right\}$$

where $t_{\frac{\alpha}{2}}(n-1)$ is the inverse cumulative probability of a t-distribution with $n-1$ degrees

of freedom at $1 - \alpha/2$. This value may be obtained from a t-distribution table or by using a statistical software package. Clearly, μ either will or will not be covered in the interval; the theory is that with repeated samples, μ would be contained in $100*(1-\alpha)\%$ of all intervals *in the long run*. Note that the value of α typically used is 0.05, leading to 95% CIs.

Importantly, however, many studies conducted by Six Sigma practitioners are analytic, a term discussed by W. Edwards Deming.⁴ Analytic studies are designed to assess the underlying causal system of a process. As such, the population under inspection therefore includes future production. Even with 100% inspection, CIs may therefore have some legitimacy (in particular, provided the process exhibits stability).

Misconception 3:

Confidence intervals address the spread of individual values.

Occasionally, practitioners are led to believe that a CI covers a particular proportion of the population (e.g., 95%). However, this is not the case for the most widely used CIs. In particular, the CI is a region one may expect a parameter (e.g., the population mean) to occupy. For the situation in which a reasonable upper and lower bound is required to cover a specific proportion of a population, tolerance intervals (TIs) may be appropriate.

Tolerance Intervals

TIs are constructed such that a specific proportion (p) of the population will be contained with a prescribed confidence level ($1-\alpha$). For example, consider a manufacturer of machine parts that have a required diameter of 0.42 cm and specification limits of 0.42 ± 0.02 cm. Assume 20 machine parts are randomly sampled from a stable process that is well modeled by a normal (Gaussian) distribution. Figure 2 shows the 95% CI for μ .

Fig. 2

One-Sample T: Diameter

Variable	N	Mean	StDev	SE Mean	95.0% CI
Diameter	20	0.42328	0.01776	0.00397	(0.41497, 0.43160)

Since values between 0.40 cm and 0.44 cm are within specification limits, the results obtained from the sample (0.4150 cm to 0.4316 cm) appear to indicate that the process would be acceptable. **However, it is crucial to note that the CI obtained is assessing the mean of the population – not the individual values.** The results from a 95% confidence interval that covers 99% of the population are shown in Figure 3.

Figure 3

Tolerance Interval for Diameter

Tolerance (Confidence) Level:	95%		
Proportion of Population Covered:	99%		
N	Mean	StDev	Tolerance Interval
20	0.423285	0.0177577	(0.359099, 0.487470)

These results (0.3591 cm to 0.4875 cm) indicate that the process itself is actually doing poorly; i.e., a high proportion falls outside the specification limits.

Implications

Though CIs are very useful tools for gaining process knowledge, it is crucial to consider drawbacks in their interpretation. For more information on the relationship between CIs and TIs, see Robert V. Hogg and Johannes Ledolter.⁵

About the Author

Keith M. Bower is a technical training specialist with Minitab Inc. He received a bachelor's degree in mathematics with economics from Strathclyde University in Great Britain and a master's degree in quality management and productivity from the University of Iowa in Iowa City, USA. Bower has conducted training courses for firms including GE Capital, American Express, and Motorola. He is a member of ASQ and the Six Sigma Forum.

Technical Information: For more information on the macro for tolerance intervals, refer to Answer ID #1216 at www.minitab.com/support/answers

References

1. Jerzy Neyman, "On the Two Different Aspects of the Representative Method," *Journal of the Royal Statistical Society A*, no.97 (1934): 558-606.
2. Robert V. Hogg and Eliot Tanis, *Probability and Statistical Inference*, 5th ed. (New Jersey: Prentice-Hall, Inc., 1997), 300.
3. Nathaniel Schenker and Jane F. Gentleman, "On Judging the Significance of Differences by Examining the Overlap Between Confidence Intervals," *The American Statistician* 55, no.3 (2001): 182-186.
4. W. Edwards Deming, *Some Theory of Sampling* (New York: Dover Publications, Inc., 1966), 247-261.
5. Robert V. Hogg and Johannes Ledolter, *Applied Statistics for Engineers and Physical Scientists*, 2nd ed. (New York: Macmillan Publishing Company, 1992), 169.