§7.5 Determining the sample size

In the formula for approximating μ,

$$\bar{x} \pm (z_{\alpha/2})\left(\frac{\sigma}{\sqrt{n}}\right),$$

the variability of the confidence interval is called the sampling error. That is

$$(z_{\alpha/2})\left(\frac{\sigma}{\sqrt{n}}\right) = SE.$$

Now suppose we wish to build a confidence interval with a limit on the size of the sampling error. What we can control is n.

In order to estimate μ with a sampling error (SE) and with $100(1-\alpha)\%$ confidence level, the sample size (n) is

$$n \geq \frac{(z_{\alpha/2})^2 \sigma^2}{(SE)^2}$$

Note: Since σ is rarely known. It is common to run a pilot survey and approximate σ by s from the pilot survey or $R/4 \approx \sigma$ (one-forth of the range of the pilot survey). In any case n should ALWAYS be rounded up.

Example A pilot survey of statistics students exam scores generated a range of values between 57 and 91. How many students should be surveyed to find a 95% confidence interval for the average score on this exam with a sampling error of 2.5?

$\alpha = .05$, $z_{\alpha/2} = 1.96$, the estimate for σ is $(91-57)/4 \approx 8.5$, so

$$n \geq \frac{(z_{\alpha/2})^2 \sigma^2}{(SE)^2} = \frac{(1.96)^2(8.5)^2}{(2.5)^2} = 44.408896$$

So $n = 45$ is needed to produce the 95% confidence interval for μ with sampling error of 2.5.