Set 23 Part B: Sample Size for Confidence Intervals of Means

Stat 260: July 19, 2024

Question: When designing an experiment, how large of a sample size n do we require?

Recall the $(1-\alpha)100\%$-Confidence Interval for μ:

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad \text{or} \quad \bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$$

point estimator

estimated standard error

$$\bar{x} - \underbrace{\boxed{z_{\alpha/2} \frac{\sigma}{\sqrt{n}}}}_{d} \qquad \bar{x} + \underbrace{\boxed{z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}}}_{d}$$

- **Length** of a confidence interval:

  $L = 2d = 2$ (critical value)(standard error or estimated standard error)

- **Margin of error:** the distance d from one endpoint of a (symmetric) CI to its centre.

  $d =$ (critical value)(standard error or estimated standard error)

Rearranging $d = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$$\to \frac{d}{z_{\alpha/2}} = \frac{\sigma}{\sqrt{n}}$$

$$\to \frac{z_{\alpha/2}}{d} = \frac{\sqrt{n}}{\sigma}$$

$$\to n = \left(\frac{z_{\alpha/2} \cdot \sigma}{d}\right)^2$$

$\to$ always round up to next whole num

Use σ when available, and s when it is not. But what if we don't know either?

- Estimate s by using a sample standard deviation from an older study.
- Run a small pilot study to find an estimate for s.

$\nu = n-1$

What about the Case 3 CI: $\bar{x} \pm t_{\nu,\alpha/2}\frac{s}{\sqrt{n}}$? $\rightarrow n = \left(\frac{t_{\nu,\alpha/2} \cdot s}{d}\right)^2$

CI's with a T-dist have $\nu = n-1$

(we cant find $t_{\nu,\alpha/2}$, unless we already know "n")

* always use $z_{\alpha/2}$ critical value in sample size calculations (never use t)

Example 1: The Texas Flame tulip is a "parrot tulip" featuring a brilliant yellow and red variegated blossom. A horticulturist wants to study the (population) mean blossom length of Texas Flame tulips. An older study from 1992 suggests that the variance of the blossom lengths is 0.49cm. Determine the sample size needed for the horticulturist to estimate μ within 0.08cm, with 98% confidence.

$s^2 = 0.49 \rightarrow s = \sqrt{0.49}$ $\quad d = 0.08\text{cm}$ $\quad$ 98% CI

98% CI $\Rightarrow 1-\alpha = 0.98 \rightarrow \alpha = 0.02 \Rightarrow \alpha/2 = 0.01$

we want $z_{0.01} = 2.33$

$$n = \left(\frac{z_{\alpha/2} \cdot s}{d}\right)^2 = \left(\frac{2.33 \cdot \sqrt{0.49}}{0.08}\right)^2 = 415.65$$

round up to $\boxed{n = 416}$ $\rightarrow$ final answer

Example 2: Orcas (*Orcinus orca*) are noted for having large retinal ganglion cells compared to terrestrial mammals. Determine the sample size need to estimate μ, the true mean retinal ganglion cell diameter to within 3μm with 82% confidence, if we assume from past data that $\sigma = 9\mu$m.