

7.2.2 A confidence Interval for a Population Proportion

Consider a population whose members can be divided into 2 separate groups. Let p be the population proportion or the true proportion, which is unknown. To estimate p , we use

$$\hat{p} = \frac{X}{n}$$

where X = the number of people in the sample who have a given characteristic, n = sample size.

- X follows a $Bm(n, p)$.
- Furthermore, if both $np \geq 10$ and $n(1 - p) \geq 10$, X has approximately a $N(np, np(1-p))$ distribution.
- Since $\hat{p}$ is just X multiplied by the constant $\frac{1}{n}$, $\hat{p}$ also has approximately a $N(p, \frac{p(1-p)}{n})$ distribution with mean

$$E(\hat{p}) = p$$

and variance

$$\text{Var}(\hat{p}) = \frac{p(1-p)}{n}$$

- If $n > \frac{30}{p}$, then a CI for the proportion p is

$$\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Since in general, for all CI:

estimate $\pm$ critical value $\cdot \sigma$ estimate

$$\left. \begin{array}{l} \text{For } \mu: \\ \bar{X} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \\ \uparrow \\ \hat{\mu} \end{array} \right\}$$

Lec 27

Example 83. (Exercise 21 on textbook page 294) In a sample of 1000 randomly selected consumers who had opportunities to send in a rebate claim form after purchasing a product, 250 of these people said they never did so. Calculate an upper confidence bound at the 95% confidence level for the true proportion of such consumers who never apply for a rebate. Based on this bound, is there compelling evidence that the true proportion of such consumers is smaller than 1/3? Explain your reasoning.

Solution.

$$P(p \leq ?) \geq 0.95$$

LCB

$$P(p \geq ?) \geq 0.95$$

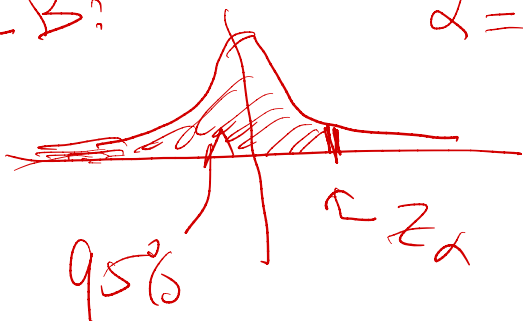
CI:

$$P(? \leq p \leq ?') \geq 0.95$$

UCB:

$$\alpha = 0.05$$

LCB



Conf. upper bound for μ :

$$\bar{X} + z_{\alpha} \frac{\sigma}{\sqrt{n}} \leftarrow \text{replace by } S \text{ if } \sigma \text{ is unknown}$$

Lower conf. bound:

$$\bar{X} - z_{\alpha} \frac{\sigma}{\sqrt{n}} \leftarrow \text{or } S$$

proportion:

$$CUB: \hat{p} + z_{\alpha} \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$$

$$CLB: \hat{p} - z_{\alpha} \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$$

$$\hat{p} = \frac{250}{1000} = 0.25$$

Estimate for σ :

$$\sqrt{\hat{p}(1-\hat{p})} = \sqrt{0.25 \times 0.75}$$

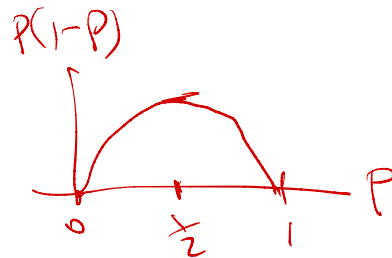
$$\left[0.25 + \overset{\downarrow z_{0.05}}{1.645} \times \frac{\sqrt{0.25 \times 0.75}}{\sqrt{1000}} \right] = 0.273$$

We are 95% confident that $p \leq 0.273$
 true proportion

$$\text{since } 0.273 < \frac{1}{3}$$

we are inclined to believe the claim.

7.2.3 Sample Size Consideration



Again, if we specify our confidence level (i.e. the reliability) and the width w of the CI (i.e. the precision), then the “smallest” sample size can be found as

$$2 z_{\alpha/2} \frac{\sqrt{p(1-p)}}{\sqrt{n}} \leq w$$

$$n \geq \frac{4 z_{\alpha/2}^2 \hat{p}(1-\hat{p})}{w^2}$$

Since $\hat{p}$ unknown,
use $\frac{1}{2}$

Example 84. (Exercise 25 on textbook page 294) A state legislator wishes to survey residents of her district to see what proportion of the electorate is aware of her position on using state funds to pay for abortions.

a. What sample size is necessary if the 95% CI for p is to have a width of at most 0.1 irrespective of p ?

b. If the legislator has strong reason to believe that at least $2/3$ of the electorate know of her position, how large a sample size would you recommend?

Solution.

a)
$$\frac{4 z_{\frac{0.05}{2}}^2 \cdot \frac{1}{2} \times (1 - \frac{1}{2})}{0.1^2} = 384.16 \quad n \geq \underline{\underline{385}}$$

$z_{\frac{0.05}{2}} = 1.96$

b)
$$\frac{4 z_{\frac{0.05}{2}}^2 \cdot \frac{2}{3} (1 - \frac{2}{3})}{0.1^2} = 341.408 \quad n \geq 342$$