

Group Work - Chapter 8

1 Researchers discover a new gene which, under the right circumstances, could lead to a mildly inconvenient, but chronic, disease. 10% of the general population have the gene. One of the researchers thinks that people with naturally red hair are more likely to have the gene.

- (a) The researchers would like to conduct a study which will estimate the prevalence of the gene in people with naturally red hair to within 2.5% at a 95% confidence level. What is the size of the sample they need to collect?

By hand:

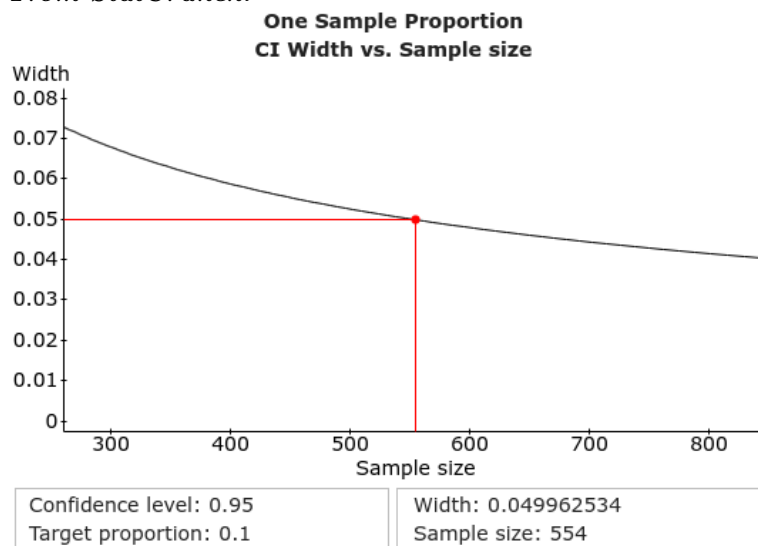
$$\alpha = 0.05, \quad z_{\alpha/2} = 1.96$$

$$ME = 0.025$$

$$\hat{p} = 0.1$$

$$n = \left(\frac{\sqrt{\hat{p} \times (1 - \hat{p})} \times z_{\alpha/2}}{ME} \right)^2 = \left(\frac{\sqrt{0.1 \times 0.9} \times 1.96}{0.025} \right)^2 = 553.19 \Rightarrow 554$$

From StatCrunch:



- (b) Genetic tests are conducted on a sample of 65 redheads and it is found that 11 of them have the gene. Construct a confidence interval at a 95% level of confidence. What conclusions can be drawn from the confidence interval?

$$\hat{p} = \frac{11}{65} = 0.169$$

$$\begin{aligned} CI &= \hat{p} \pm z_{\alpha/2} \times \sqrt{\frac{\hat{p} \times (1 - \hat{p})}{n}} = 0.169 \pm 1.96 \times \sqrt{\frac{0.169 \times 0.831}{65}} \\ &= 0.169 \pm 1.96 \times 0.0465 = 0.169 \pm 0.091 = (0.078, 0.260) \end{aligned}$$

10% is contained in the confidence interval. There is *not* evidence that the gene occurs at a greater rate among redheads.

- (c) What are the null and alternative hypotheses for a test on this claim? Is this a one-sided or two-sided test? Is the claim represented by the null or alternative hypothesis?

$$H_0 : p = 0.1$$

$$H_a : p > 0.1$$

This is a one-sided test.

The claim is represented by the alternative hypothesis.

2 Suppose the mean age of Metro State students is 32 years old with a standard deviation of 5.5. For a research project a stats student wants to examine the idea that students who take statistics classes are the same age as other Metro students.

- (a) What sample size is required to determine the mean age of stats students within plus or minus 3 years at a 99% confidence level?

By hand:

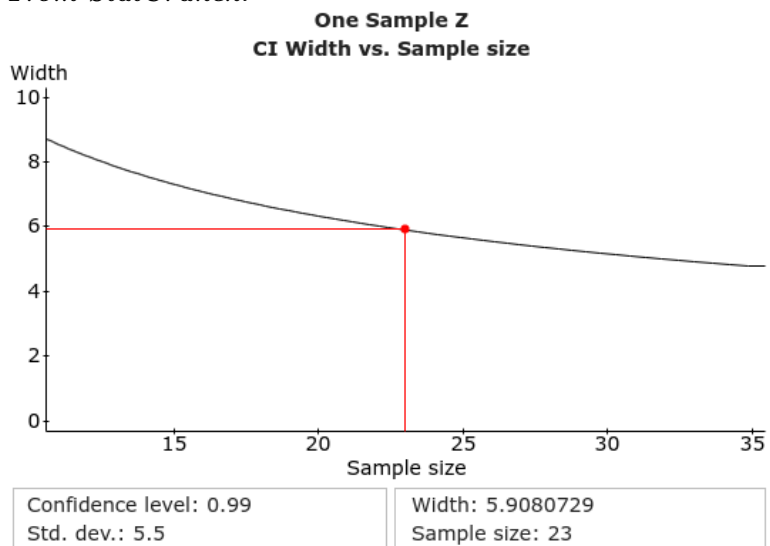
$$\alpha = 0.01, \quad z_{\alpha/2} = 2.58$$

$$ME = 3$$

$$s = 5.5$$

$$n = \left(\frac{s \times z_{\alpha/2}}{ME} \right)^2 = \left(\frac{5.5 \times 2.58}{3} \right)^2 = 22.37 \Rightarrow 23$$

From StatCrunch:



- (b) A random sample of 25 statistics students has a mean age of 35.4 with a standard deviation of 4.6. Construct a confidence interval at a 99% level of confidence. What conclusions can be drawn from the confidence interval?

By hand (normal approximation):

$$CI = \bar{x} \pm \frac{z_{\alpha/2} \times s}{\sqrt{n}} = 35.4 \pm \frac{2.58 \times 4.6}{5} = 35.6 \pm 2.37 = (33.23, 37.97)$$

From StatCrunch (T Stats):

$$CI = (32.83, 37.97)$$

32 is not contained in the confidence interval. There is evidence stats students are older than Metro State students in general.

- (c) What are the null and alternative hypotheses for a test on this claim? Is this a one-sided or two-sided test? Is the claim represented by the null or alternative hypothesis?

$$H_0 : \mu = 32$$

$$H_a : \mu \neq 32$$

This is a two-sided test.

The claim is represented by the null hypothesis.

3 The data file “bears.csv” on D2L contains measurements of a random sample of bears from a national park. Park officials are concerned that the bear population is underweight and thus not prepared for the long winter. A healthy bear population has a mean weight of 200 lbs.

- (a) A small pilot study the previous year determined the standard deviation for the weight of bears was about 100 lbs. What sample size is required to determine the mean weight of bears within plus or minus 30 lbs. at a 95% confidence level?

By hand:

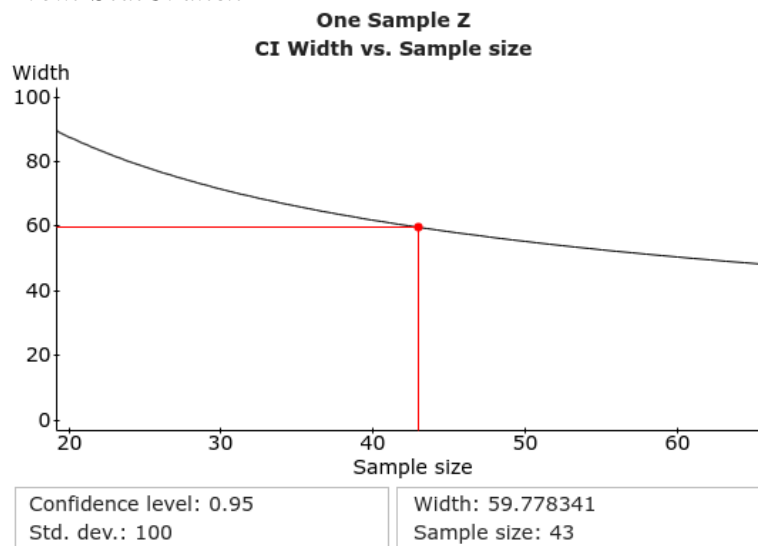
$$\alpha = 0.05, \quad z_{\alpha/2} = 1.96$$

$$ME = 30$$

$$s = 100$$

$$n = \left(\frac{s \times z_{\alpha/2}}{ME} \right)^2 = \left(\frac{100 \times 1.96}{30} \right)^2 = 42.68 \Rightarrow 43$$

From StatCrunch:



- (b) Using the data from the “bears.csv” construct a 95% confidence interval. What conclusions can be drawn from the confidence interval?

From StatCrunch (T Stats):

$CI = (149.64, 216.13)$

200 is contained in the confidence interval. There is no evidence that the bears are underweight.

- (c) What are the null and alternative hypotheses for a test on this claim? Is this a one-sided or two-sided test? Is the claim represented by the null or alternative hypothesis?

$H_0 : \mu = 200$

$H_a : \mu < 200$

This is a one-sided test.

The claim is represented by the alternative hypothesis.