Chapter 8 Confidence Intervals (TI)

When we don’t know the population mean, \( \mu \) (or population proportion, \( \hat{p} \)) the best estimate is the sample mean, \( \bar{x} \) (or the sample proportion, \( \hat{p} \) (p hat)). But, just how good is it?

A confidence interval is a range of values that is likely to contain the true value of the population parameter (\( \mu \) in this case).

The degree (or level) of confidence is the probability \( 1 - \alpha \) that the confidence interval contains the true value of the population parameter. It is also the percentage of times that the confidence interval contains the population parameter.

### Confidence Intervals for Population Means

We can construct a confidence interval for the population mean using the normal distribution if the following requirements are met:

- Simple Random Sample
- Normal Distribution or \( n > 30 \)
- \( \sigma \) known

In most “real life” situations the population standard deviation is not known. If it is known then most likely the population mean would also be known and there would be no need to estimate it. If the population standard deviation is not known we can still estimate the population by using the \( t \)-distribution. The \( t \)-distribution is a bell–shaped curve similar to the Normal distribution.

The degrees of freedom is the number of sample values that can vary after certain restrictions have been imposed on all the data values. For this section the degrees of freedom: \( df = n - 1 \)

Requirements:
- Simple Random Sample
- Population is Normal or \( n > 30 \)

\( \alpha \) = the total area under both tails
\( \alpha/2 \) = the area under one tail
\( t_{\alpha/2} \) = the critical value which corresponds to the point that separates an area of \( \alpha/2 \) in the right tail

\( \bar{X} \sim t_{df} \)

Point Estimate for the population mean, \( \mu \) is \( \bar{x} \).

EBM (Error Bound for the Mean) = \( t_{\alpha/2} \frac{s}{\sqrt{n}} \)

Confidence Interval: (\( \bar{x} - EBM \), \( \bar{x} + EBM \))

Point Estimate = \( \frac{\text{lower bound} + \text{upper bound}}{2} \)

Error Bound = \( \frac{\text{upper bound} - \text{lower bound}}{2} \)

We will be using the calculator to compute the confidence interval.
Example 1:
A department store wishes to estimate the mean (average) number of customers who pass a particular location in their store each day. A sampling of the traffic on 18 randomly selected days at this location showed a mean of 1246 customers and a standard deviation of 145. Assume the distribution is approximately normal. Find a 90% confidence interval for the mean number of customers who daily pass this location.

\[ x = 1246 \]
\[ s = 145 \]
\[ n = 18 \]
\[ df = 17 \]
\[ CL = 90\% \]
\[ \alpha = 0.10 \]
\[ \alpha/2 = 0.05 \]

\( X \) is the number of customers who pass a particular location in this store in one day
\( \bar{X} \) is the mean number of customers who pass a particular location in this store in 18 days
\( \bar{X} \sim t_{17} \)

To construct a Confidence Interval using the TI calculator:

\[
\begin{array}{c}
[\text{STAT}] \{\text{TESTS}\} \{\text{TInterval}\} \\
\{\text{Stats}\} \\
\text{Enter the values as shown then} \\
\{\text{Calculate}\}
\end{array}
\]

The 90% confidence interval is (1186.5, 1305.5)

If many random samples were taken with sample size of 18, and 90% confidence intervals were constructed, we would expect 90% of them to contain the population mean number customers who pass a particular location in this store.

or

We are 90% confident this interval contains the true population mean number of customers who pass a particular location in this store in 18 days

The error bound is \( \frac{1305.5 - 1186.5}{2} = 59.5 \)

Additional questions: (know these for the exam)

a. Based on the confidence interval, does it appear that 1400 people could pass this location?
   \( \text{Answer: No, 1400 is not contained in the interval} \)

b. Based on the confidence interval, does it appear that 1200 people could pass this location?
   \( \text{Answer: Yes, 1200 is contained in the interval} \)

c. Based on the confidence interval, does it appear fewer than 1400 people passes this location?
   \( \text{Answer: Yes, all the confidence interval values are less than 1400} \)

d. Based on the confidence interval, does it appear more than 1400 people passes this location?
   \( \text{Answer: No, all the confidence interval values are less than 1400 (no credit for saying 1400 is not in the interval)} \)
Example 2:
Boxes of Eat Healthy cereal are labeled as containing 14 ounces. Following are the weights (in ounces) of a random sample of Eat Healthy cereal boxes. Find a 90% confidence interval for the mean weight of these cereal boxes. Assume the distribution is approximately normal. Based on this interval, does it appear the company is cheating customers?

14.3  14.1  13.4  13.6  13.6  13.3  13.6  14.1
14.2  14.1  13.5  13.6  13.1  13.7  13.2  14.0

\( \bar{x} = 13.7125 \)
\( s = 0.3757 \)
\( n = 16 \)
\( df = 15 \)
\( CL = 90\% \)
\( \alpha = 0.10 \)
\( \alpha/2 = 0.05 \)

\( X \) is the weight of one box of Eat Healthy cereal
\( \bar{X} \) is the mean weight of 16 boxes of Eat Healthy cereal
\( X \sim t_{14} \)

The 90% confidence interval is (13.548, 13.877).

Since all of the interval values are less than 14, it appears the weight of the cereal is less than 14 oz and the company is cheating their customers.
Confidence Intervals for Population Proportion

We can use the Normal distribution to estimate the population proportion if the data satisfies the requirements for a Binomial distribution, \( np' > 5 \), and \( nq' > 5 \).

Notation:
\[
p' = \frac{x}{n} \quad \text{the best point estimate}
\]
\[
q' = 1 - p'
\]
\[
\alpha = \text{the total area under both tails}
\]
\[
\alpha/2 = \text{the area under one tail}
\]
\[
z_{\alpha/2} = \text{the critical value which corresponds to the point that separates an area of } \alpha/2 \text{ in the right tail}
\]
\[
P' \sim N \left( p', \sqrt{\frac{p'q'}{n}} \right)
\]

Point Estimate for the population proportion, \( \mu \) is \( p' \).

EBP (Error Bound for the Proportion) = \( z_{\alpha/2} \sqrt{\frac{p'q'}{n}} \)

Confidence Interval: \( (p' - EBP, p' + EBP) \)

Point Estimate = \( \frac{\text{lower bound} + \text{upper bound}}{2} \)

Error Bound = \( \frac{\text{upper bound} - \text{lower bound}}{2} \)

We will be using the calculator to compute the confidence interval.
Example 3:
A random sample of 600 homes produced 318 home owners who admitted to owning at least one gun. Find a 98% confidence interval for the proportion of home owners who own a gun.

\[ x = 318 \]
\[ n = 600 \]
\[ p' = \frac{318}{600} = 0.53 \]
\[ q' = 1 - 0.53 = 0.47 \]
\[ \sqrt{\frac{(0.53)(0.47)}{600}} = 0.0204 \]
The distribution is \( p' \sim N(0.53, 0.0204) \)
CL = 98%
\[ \alpha = 1 - .98 = .02 \]
\[ \alpha/2 = .01 \]

X is the number of home owners who admit to owning at least one gun
\( p' \) is the proportion of home owners from the sample who admit to owning at least one gun

To construct a Confidence Interval using the TI calculator:

[STAT] {TESTS} {1–PropZInt}
Enter the values as shown then {Calculate}

The 98% confidence interval is (0.4826, 0.5774)

If many random samples were taken with sample size of 600, and 98% confidence intervals were constructed, we would expect 98% of them to contain the population proportion of home owners who admit to owning at least one gun.

or

We are 98% confident that this interval contains the true proportion or homeowners who own at least one gun.

Follow–up questions:
Based on the confidence interval, does it appear than 50% of home owners own a gun?
Answer: Yes because 0.50 is contained in the confidence interval.

Based on the confidence interval, does it appear that more than 60% of home owners own a gun?
Answer: No because 0.60 is greater than the confidence interval values. (Note that if you say no because 0.60 is not contained in the confidence interval you will not receive full credit.)
Comparing Confidence Intervals

When comparing intervals we want to determine if there is an overlap. If the intervals do overlap say there is no evidence of a significant difference in the population means (or proportions). If the intervals do not overlap we say there is evidence of a significant difference between the population means (or proportions). For this case you need to say which population appears to have the greater (or smaller) population mean (or proportion). Be sure to write the statement in the context of the problem.

Example 4:
Suppose we want to determine which brand of batteries last longest. A study is done to test the average time (in hours) each brand lasts and the following confidence intervals are produced.

Brand A: (33.5, 42.1) Brand B: (38.0, 45.6) Brand C: (19.6, 25.9)

Comparing Brand A and Brand B: The confidence intervals overlap so there is no evidence of a significant difference between the population average length of times for Brand A and Brand B

Comparing Brand A and Brand C: The confidence intervals do not overlap so there is evidence of a significant difference between the population average length of times for Brand A and Brand C. It appears that Brand A batteries last longer. (You need to include the word “appears”)

Sample Size

How large should the sample size be to determine an estimate for the population proportion?

EBP = \( z_{\alpha/2} \sqrt{\frac{p'q'}{n}} \) solving for n:

\[
  n = \frac{(Z_{\alpha/2})^2 p'q'}{EBP^2} \quad (p' \text{ known})
\]

\[
  n = \frac{(Z_{\alpha/2})^2 0.25}{EBP^2} \quad (p' \text{ not known})
\]

**Always round up**

Note: The reason for using 0.25 in the second formula is if \( p' \) is not known we want to use the largest possible value for \( p' \cdot q' = (.5)(.5) = 0.25 \)

Example 5:
An investment institution needs to know the percentage of working people who have a tax–sheltered annuity plan. How large a sample will ensure with 95% certainty that the error in the estimate of \( p \) by the sample proportion \( \hat{p} \) is at most 4%?

EBP = 0.04
CL = 0.95
\( \alpha = 0.05 \)
\( \alpha/2 = 0.025 \)
\( Z_{\alpha/2} = \text{InvNorm}(0.975, 0, 1) = 1.96 \)

\[
  \frac{(1.96)^2(0.25)}{0.04^2} = 600.25
\]

rounding up gives \( n = 601 \)