NOTES: Confidence intervals from sample proportions

Suppose that we are estimating an unknown population proportion \( p \). We do this by first finding a sample proportion \( \hat{p} \) and then calculating its confidence interval. In theory, the formula for the confidence interval would be

\[
\hat{p} \pm z^* \left( \frac{SD(\hat{p})}{\sqrt{n}} \right)
\]

or

\[
\hat{p} \pm z^* \left( \frac{p(1-p)}{n} \right)
\]

where \( z^* \) depends on the Confidence Level.

But this formula involves the parameter \( p \) that we are trying to estimate! We therefore use the Standard Error for \( \hat{p} \) (in other words, replace \( p \) with \( \hat{p} \) in the standard deviation formula) instead of the Standard Deviation for \( \hat{p} \). The formula for the confidence interval then becomes

\[
\hat{p} \pm z^* SE(\hat{p})
\]

or

\[
\hat{p} \pm z^* \left( \frac{\hat{p}(1-\hat{p})}{n} \right)
\]

\( z^* \) can be found with the following, where \( C \) is the given Confidence Level inputted in decimal form.

\[
z^* = \text{InvNorm} \left( \frac{C}{2} + 0.5 \right)
\]

Let \( L \) denote the left endpoint of the confidence interval and let \( R \) denote the right endpoint of the confidence interval. Therefore, the confidence interval written in interval notation would be \( (L, R) \). The margin of error \( (m) \) is the distance that our confidence interval branches out in either direction from \( \hat{p} \).

Think of this as \( L \leftarrow m \rightarrow \hat{p} \rightarrow R \). Thus, the confidence interval formula can be thought of as

\[
\hat{p} \pm m
\]

It follows that the margin of error for these confidence intervals is

\[
z^* \left( \frac{\hat{p}(1-\hat{p})}{n} \right)
\]

In other words \( m = z^* \left( \frac{\hat{p}(1-\hat{p})}{n} \right) \). Also, \( m = \frac{(R-L)}{2} \).

INTERPRETATION: “We are ___% confident that \( p \) is between \( L \) and \( R \).”

Suppose that the confidence level is 95%. Then there is a 95% probability that choosing any confidence interval will give us \( p \) in it, because \( p \) is in 95% of the confidence intervals that we can create. Once we fix \( \hat{p} \) and the confidence interval it generates, then \( p \) is either in that particular confidence interval or it is not. Once \( \hat{p} \) and its confidence interval is fixed then we say that “we are 95% confident” that \( p \) is in that particular confidence interval. That is, in repeated trials, we expect that 95% of the confidence intervals in the long run, to contain \( p \).

CONDITIONS/REQUIREMENTS: Before we are allowed to use the confidence interval formula certain conditions must be met:

1. The sample proportion \( \hat{p} \) must be obtained from a Simple Random Sample (SRS).
2. The number of successes in the sample is at least 5, preferably at least 10, i.e., \( n\hat{p} \geq 10 \).
3. The number of failures in the sample is at least 5, preferably at least 10, i.e., \( n(1-\hat{p}) \geq 10 \).
4. The population size is at least ten times the sample size \( (n) \).
1. As part of a quality improvement program, your mail-order company is studying the process of filling customer orders. According to company standards, an order is shipped on time if it is sent within 3 working days of the time it is received. You select an SRS of 100 of the 5000 orders received in the past month for an audit. The audit reveals that 86 of these orders were shipped on time. Find a 95% confidence interval for the true proportion of the month’s orders that were shipped on time. Also, interpret the confidence interval.

2. A student organization wants to start a nightclub for students under the age of 21. To assess support for this proposal, they will select an SRS of students and ask each respondent if he or she would patronize this type of establishment. In the time available, they were able to obtain 356 responses of which 70% said “yes.” Calculate and interpret a 90% confidence interval for this situation. What is the margin of error?

"When you say you're 95% confident... just what are you inferring?"

"I got the instructions from my Statistics Professor. He was 80% confident that the true location of the restaurant was in this neighborhood."

Taken from CAUSEweb.org and cartoonist J.B. Landers ©