Chapter 18

CONFIDENCE INTERVALS FOR PROPORTIONS

STANDARD ERROR

- Estimates the theoretical standard deviation of the sampling distribution for sample proportions based on a single sample:

\[ SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}} \]

A CONFIDENCE INTERVAL

- By the 68-95-99.7% Rule, we know:
  - About 68% of all samples will have \( \hat{p} \) within 1 SE of \( p \)
  - So we are 68% sure \( p \) lies within one SE of \( \hat{p} \)
  - About 95% of all samples will have \( \hat{p} \) within 2 SEs of \( p \)
  - So we are 95% sure \( p \) lies within two SEs of \( \hat{p} \)
  - About 99.7% of all samples will have \( \hat{p} \) within 3 SEs of \( p \)
  - So we are about 99.7% sure \( p \) lies within three SEs of \( \hat{p} \)
- These are confidence intervals
CONFIDENCE INTERVALS

- An interval of values that is fairly certain to contain the true value of the population parameter of interest
- The degree of confidence reflects the frequency of times that the confidence interval actually does contain the population parameter, assuming that the estimation process is repeated a large number of times.

VISUALIZING CONFIDENCE INTERVALS


MARGIN OF ERROR: CERTAINTY VS. PRECISION

- We can claim, with 95% confidence, that the interval \( \hat{p} \pm 2SE(\hat{p}) \) contains the true population proportion.
  - The extent of the interval on either side of \( \hat{p} \) is called the margin of error (ME).
- In general, confidence intervals have the form estimate ± ME.
- The more confident we want to be, the larger our ME needs to be.
CRITICAL VALUES

- The ‘2’ in $\hat{p} \pm 2SE(\hat{p})$ (our 95% confidence interval) came from the 68-95-99.7% Rule.
- Using a table or technology, we find that a more exact value for our 95% confidence interval is 1.96 instead of 2.
  - We call 1.96 the critical value and denote it $z^*$. 
- For any confidence level, we can find the corresponding critical value.

CRITICAL VALUES (CONT.)

- Example: For a 90% confidence interval, the critical value is 1.645:
ONE-PROPORTION Z-INTERVAL

- The confidence interval for the population proportion \( p \) is
  \[
  \hat{p} \pm z^* \times SE(\hat{p})
  \]
  where
  \[
  SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}
  \]
- The critical value, \( z^* \), depends on the particular confidence level that you specify.

INTERPRETING THE INTERVAL

Don’t Misstate What the Interval Means:
- Don’t suggest that the parameter varies.
- Don’t claim that other samples will agree with yours.
- Don’t be certain about the parameter.
- Don’t forget: It’s the parameter (not the statistic).
- Don’t claim to know too much.
- Do take responsibility (for the uncertainty).

CHOOSING YOUR SAMPLE SIZE

- In general, the sample size needed to produce a confidence interval with a given margin of error at a given confidence level is:
  \[
  n = \left( \frac{z^*}{ME} \right)^2 \hat{p}\hat{q}
  \]
  where \( z^* \) is the critical value for your confidence level.
  - To be safe, round up the sample size you obtain.