Confidence intervals for proportions:

For a sample proportion $p = \hat{p} = \frac{x}{n}$, where $x$ is the number of successes in $n$ trials, and $N$ is the population size, the standard error (SE) is given by:

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \times \sqrt{\frac{N-n}{N-1}}$$

The margin of error (ME) is:

$$ME = 1.96 \times SE$$

For a 95% confidence interval (CI), the CI is:

$$CI = \hat{p} \pm ME$$

For a sample mean $x = \bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$, where $x_i$ are the individual observations, and $N$ is the population size, the standard error is:

$$SE = \frac{S.D.}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$$

The margin of error (ME) is:

$$ME = 1.96 \times SE$$

For a 95% confidence interval (CI), the CI is:

$$CI = \bar{x} \pm ME$$

95% CI for $p_1 - p_2$:

$$\left(\hat{p}_1 - \hat{p}_2\right) \pm 1.96 \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \times \sqrt{\frac{N_1-n_1}{N_1-1} + \frac{N_2-n_2}{N_2-1}}$$

95% CI for $\mu_1 - \mu_2$:

$$\left(x_1 - x_2\right) \pm 1.96 \sqrt{\frac{SD_1^2}{n_1} \times \frac{N_1-n_1}{N_1-1} + \frac{SD_2^2}{n_2} \times \frac{N_2-n_2}{N_2-1}}$$

Every CI is a balance between certainty and precision.

Claim: 95% of samples of this size will produce confidence intervals that capture the true proportion.

* $Z = \text{critical value, ex. 1.96 for 95% and 1.00 for 68'}$

* For $N$ large relative to $n$, $\frac{N-n}{N-1} \approx 1$

* For $p \approx 0.5$, $\sqrt{p(1-p)} \approx 0.5$
\( \hat{p} \) is top \( p \), so
\( \hat{x} \approx \hat{p} \) to \( \mu \)

solutions:

2. \( N = \) whole population = 500 pages
\( n = \) sample set = 6
\( * = \) special ones = 3
\( \{226, 498, 713, 226, 554, 370\} \)

\( \checkmark \) \( \checkmark \) \( \checkmark \)

cross off duplicate

3. \( \{2, 4, 12\} \)
\( \bar{x} = mean = \frac{2 + 4 + 12}{3} = 6 \)
standard deviation = S.D.

\begin{tabular}{|c|c|c|}
\hline
original value & deviations & squared deviation \\
\hline
2 & 2 - 6 = -4 & (-4)^2 = 16 \\
4 & 4 - 6 = -2 & (-2)^2 = 4 \\
12 & 12 - 6 = 6 & (6)^2 = 36 \\
\hline
\end{tabular}

1. add up squared deviations: \( 16 + 4 + 36 = 56 \)
2. divide by \( n - 1 \): \( \frac{56}{3 - 1} = \frac{56}{2} = 28 \) (\( n = \# \) in set)
3. take square root: \( \sqrt{28} = 5.2915 = S.D. \)
   or \( SD = \sqrt{\left(2 - 6\right)^2 + \left(4 - 6\right)^2 + \left(12 - 6\right)^2} = \sqrt{28} = 5.2915 \)

*calculator - state, edit, enter 2 4 \& 12, state,
calc, 1-var state, 2nd \( \bar{x} \), \( \sigma \), \text{SE} = S.D.

men

\( n_1 = 200 \)
\( \hat{x}_1 = 41 \)

women

\( n_2 = 316 \)
\( \hat{x}_2 = 17 \)

a. \( \hat{p}_1 = \frac{41}{200} \)
\( \hat{p}_2 = \frac{17}{316} \)

\( \hat{p}_1 = 0.205 \)
\( \hat{p}_2 = 0.054 \)

\( \text{ESE} \)

1. \( \hat{p}_1 - \hat{p}_2 = 0.205 - 0.054 = 0.151 = \text{point estimate of } p_1 - p_2 \)

2. \( \text{SE} = \sqrt{\left(\frac{0.0008}{200}\right) + \left(\frac{0.0001}{316}\right)} \approx 0.031 \)

3. \( \text{ME} = 1.96 \times 0.031 = 0.060 = 6\% \)

4. 95\% CI = 0.151 (which is \( \hat{p}_1 - \hat{p}_2 \)) \( \pm \) 0.060 = [0.091, 0.211]
b. 95% CI does not contain 0

c. tend to think p1 is larger, p1 - p2 > 0 (so R then)

d. it could have been left off because it was approx.

5. a. men | women
   \( \bar{x}_1 = 2.78 \) | \( \bar{x}_2 = 1.84 \)
   \( SD_1 = 3.42 \) | \( SD_2 = 3.91 \)

\( 1. \ \bar{x}_1 - \bar{x}_2 = 0.94 \)

\( 2. \ SD = \sqrt{\frac{S_1^2}{n_1} \cdot \frac{n_1 - n_1}{N_1 - 1} + \frac{S_2^2}{n_2} \cdot \frac{n_2 - n_2}{N_2 - 1}} = \sqrt{0.058482 \times 1 + 0.04838 \times 1} = \sqrt{0.10728} = 0.3269 \)

\( 3. \ EME = 1.96 \times 0.3269 = 0.6407 \)

\( 4. \ 95\% \ CI = 0.94 (\text{which is } \bar{x}_1 - \bar{x}_2) \pm 0.6407 = [0.2993, 1.5807] \)

b. 95% CI does not contain 0

c. tend to think p1 is larger