March 10 - An introduction to confidence intervals

1. Setting:
   (a) \( X \) is a random variable with unknown distribution.
   (b) \( X_1, \ldots, X_n \) is a random sample from \( X \) (independent and identically distributed random variables.)

2. Key fact used: for large \( n \), \( \bar{X} \) has a distribution that has mean \( \mu \), variance \( \sigma^2/n \), and that is approximately normal. Therefore the following random variable is approximately normal with mean 0 and standard deviation 1.

\[
Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}
\]

3. Using \( Z \) and algebra we have

\[
P\left(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}\right) \approx .95
\]

(The symbol \( \approx \) means approximately equal and is because of the central limit theorem. IF \( X \) is normal, then this probability statement exact.)

4. But \( \sigma \) is not known. If \( n \) is large, use \( S \) to approximate \( \sigma \). Then

\[
P\left(\bar{X} - 1.96 \frac{S}{\sqrt{n}} < \mu < \bar{X} + 1.96 \frac{S}{\sqrt{n}}\right) \approx .95
\]

There are now two approximations here, one using the CLT and the other using \( S \) to approximate \( \sigma \).

5. The interval \([\bar{X} - 1.96 \frac{S}{\sqrt{n}}, \bar{X} + 1.96 \frac{S}{\sqrt{n}}]\) is called a 95% confidence interval for \( \mu \). Important: our confidence is not in the interval but in the procedure for producing the interval. Approximately 95% of the 95% confidence intervals that we produce will successfully capture \( \mu \).

6. Other confidence intervals: other percentages; one sided.

7. Work to do: What’s with all these approximations?

Homework, Due Tuesday, March 22

1. Read Devore and Farnum, Section 7.2

2. Do problems 7.8,9,10,12,14 of Devore and Farnum.