AP Statistics: Confidence Intervals (Means & Proportions)

From Simple Studies, https://simplestudies.edublogs.org & @simplestudiesinc on Instagram

Confidence Interval Definition

- A confidence interval is a range of values obtained from a sample that is likely to contain the parameter
- The basic formula for a confidence interval is: *Statistic* ± *(Critical Value)(Standard Deviation of Statistic)*
- The margin of error is defined by the Critical Value (Standard Deviation of Statistic)
- The critical values for each confidence level can be found on Table B of the AP Statistics formula sheets (p. 5 of the 2020 version)

Three Steps

1. Conditions/Assumptions
2. Formula
3. Interpretation

Step 1: Conditions/Assumptions

- Random Sample
    - "The stem of the problem states that [*sample*] was chosen at random"
    - "The stem of the problem states that [*participants*] were randomly assigned to the groups"
- Approximately Normal Distribution
    - "The stem of the problem states that the distribution is approximately normal"
    - "Since n = _ ≥ 30, by the Central Limit Theorem, we can assume the distribution is approximately normal"
- For two samples, both n_1 and n_2 must be ≥ 30

    - "Since np = _ ≥ 10 and nq = _ ≥ 10, we can assume the distribution is approximately normal"
- For two samples, $n_1\hat{p}_1$, $n_1\hat{q}_1$, $n_2\hat{p}_2$, and $n_2\hat{q}_2$ must all be ≥ 10
    - "Since the [*graphical display*] shows no outliers or strong skewness, we can assume the distribution is approximately normal"
- You must provide a graphical display (preferably a box plot) if normality cannot be assumed by the other three ways

Step 2: Formula

- List the formula, your substitution, degrees of freedom (if using t*) and your unrounded answer
- One-Sample Means $\quad \overline{x} \pm \frac{Z_c \sigma}{\sqrt{n}} \qquad \overline{x} \pm \frac{t^* S}{\sqrt{n}}$
- Two-Sample Means

$$\overline{x} \pm \frac{Z_c \sigma}{\sqrt{n}} \qquad \overline{x} \pm \frac{t^* S}{\sqrt{n}}$$

- Proportions

$$(\overline{x_1} - \overline{x_2}) \pm Z_c \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad (\overline{x_1} - \overline{x_2}) \pm t^* \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$\widehat{p} \pm Z_c \sqrt{\frac{\widehat{p}\widehat{q}}{n}} \qquad (\widehat{p_1} - \widehat{p_2}) \pm Z_c \sqrt{\frac{\widehat{p_1}\widehat{q_1}}{n_1} + \frac{\widehat{p_2}\widehat{q_2}}{n_2}}$$

Step 3: Interpretation

- "Based on these samples, we are _% confident that the true [*context of problem*] is between [*lower value*] and [*upper value]*"
- "Since zero [*is/is not*] in the _% confidence interval [*interval*], we [*do not/do*] have sufficient evidence to suggest there is a difference at the α = [*1 - %]* level"
- Always put the % as a decimal (e.g. 95% = .95)
- Use this second statement for two-sample confidence intervals only

Z vs t Distribution

- Means
    - Use a Z distribution when you have σ
    - Use a t distribution when you do not have σ (i.e. you have S)
- Proportions
    - ALWAYS use a Z distribution

Degrees of Freedom

- Only applies to t-distributions
    - The t-distribution varies with degrees of freedom
- df = n - 1
- For a Z-distribution, df = ∞

Finding Critical Values

- AP Statistics formula sheet table B
- TI-84 Plus
    - 2nd → Vars
        - 3 = Z-distribution
        - 4 = t-distribution
        - Area = $^{1}/_{2}$(1 - %)
            - Always put the % as a decimal (e.g. 95% = .95)

Calculator (TI-84 Plus)

- Stat → Test
    - Means
        - 7 = 1-sample Z-Interval
        - 8 = 1-sample t-Interval
        - 9 = 2-sample Z-Interval
        - 10 = 2-sample t-Interval

- Proportions
    - A = 1-Sample
    - B = 2-Sample