Unit 6 - CI and hypothesis tests for proportions

Confidence Intervals - Categorical

Confidence Interval means:

we are ___ % confident that the true proportion of population is from ___ to ___.

It is the specified probability of a range including parameter.

Conditions

- random
- Ind $n < .1N$
- normal $n \cdot p \geq 10$ $n(1-p) \geq 10$

Main idea: We observe a sample. To guess the parameter, we make a range with the sample with degree of certainty because sample is likely close to the true proportion.

formula

$$CI = \underbrace{\hat{p}}_{\text{sample proportion}} \pm \overbrace{\underbrace{z^*}_{\text{critical value}} \overbrace{\sqrt{\frac{\hat{p}(1-\hat{p})}{\underbrace{n}_{\text{sample size}}}}}^{\text{standard error}}}^{\text{margin error}}$$

Z^*: Convert degree certainty into z^* on Calc:
- 2nd vars, Invnorm
- % is area, center cause CI, enter.

♡ Our C.I. rather catches pop parameter or not, with repeated samples+tests, our CI would capture it % certainty times.

Hypothesis tests - categorical

Main idea: We observe a sample. If the chances of getting that sample are significantly small with the null hypothesis being true, we reject the null.

Conditions
- random
- $n \cdot p \geq 10$ $n(1-p) \geq 10$
- $n < .1N$

α - alpha
- Set alpha value: the point at which a smaller chance would be considered significant.

Hypothesis
- Set hypothesises to know what rejecting null means.
- Ho: Has = sign, previous claim
- Ha: >, <, or ≠, what we are inferring

C.I. difference in proportions

$$\hat{p}_1 - \hat{p}_2 \pm \underbrace{z^* \underbrace{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}_{\text{standard error}}}_{\text{margin Error}}$$

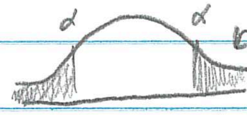
If 0 is in the CI, it is not Signifigantly different.

Solving P-value

formula: $= \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = $ Z-star

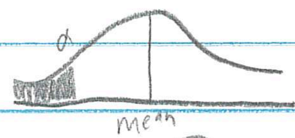
plug into 2nd vars, norm cdf, draw picture to know which side to put on.

Ha: $\neq$

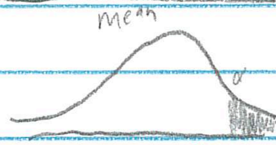


Split % ex) 95% → 97.25%

Ha: $\bar{x} < \#$

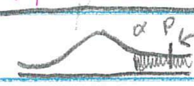


Ha: $\bar{x} > \#$



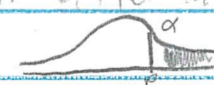
P-Value Interpretation

1. When $\alpha >$ p-val



reject null. When null is rejected, you can accept alternate

2. When $\alpha <$ p-val



Don't reject null, not Signifigant, but don't accept null or alternate.

Logic: When null is rejected, the only other option is the alternate is true. When null is not rejected, we don't have enough information to accept it but we can't confirm that its false either.

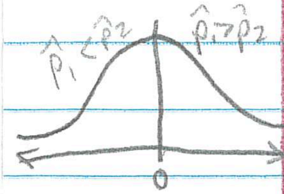
Error types

- E1: Reject a true null hypothesis
  probability: α

Ex) $\alpha = .05$ means you are accepting a 5% chance you reject true

- E2: fail to reject false null hypothesis.
  probability: 1-power
- ☆ power of test is increased by
- Sample size increasing, or the
- α value increasing.

Differences hypothesis test



$$\frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}_c(1-\hat{p}_c)} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

→ Plug into 2nd vars, norm cdf, compare α and p-value.

$$\hat{p}_c = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}$$

Ho: $\hat{p}_1 - \hat{p}_2 = 0$ no differ

Ha. $\hat{p}_1 - \hat{p}_2 <, >, \neq 0$

CI: If 0 fits in Interval, it is not Signifigant diff