V31 - Hypothesis Testing – Part 7

- Use of Z distribution for inference on difference in two population proportions

Course: Statistical Testing & Regression
Dr. Renee Clark
Swanson School of Engineering
Industrial Engineering
University of Pittsburgh

Inference for Difference in Two Proportions ($p_1 - p_2$)

Example: explore sedans vs. full sized trucks on reliability

p_1 is proportion of all sedans that break down before accruing 50K miles

p_2 is proportion of all full-sized trucks that break down before accruing 50K miles

∴ Do these population proportions differ?

Inference for Difference in Two Proportions ($p_1 - p_2$)

CI:

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$p_1 - p_2$

$\sigma_1^2 = \sigma^2$ "pooled"

$H_o: p_1 = p_2$

$H_o: p_1 - p_2 = \phi$

Define $\hat{p}$ as pooled estimate of p:

$$p \approx \hat{p} = \frac{Y_1 + Y_2}{n_1 + n_2},$$

→ binomial RV's

*

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

→ 0

Inference for Difference in Two Proportions ($p_1 - p_2$)

Key Assumptions:

1. $\hat{p}(n_1+n_2) \geq 5$
2. $(1-\hat{p})(n_1+n_2) \geq 5$

pooled estimate

$$\hat{p} = \frac{y_1+y_2}{n_1+n_2}$$

$$\hat{p}_1 = \frac{y_1}{n_1}$$

$$\hat{p}_2 = \frac{y_2}{n_2}$$

Inference for Difference in Two Proportions ($p_1 - p_2$)

Relationship to Confidence Interval

$H_o: p_1 = p_2$ → Ho: $p_1 - p_2 = \phi$

*(-0.6 , 0.2) ϕ is plausible for $p_1 - p_2$

*(0.3 , 0.4) ϕ is not plausible for $p_1 - p_2$

Reject ϕ as plausible for $p_1 - p_2$

Acknowledgement

This material is based upon work partially supported by the National Science Foundation under Grant# 2335802. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

You are free

- to **Share** – to copy, distribute, display and perform the work
- to **Remix** – to make derivative works

Under the following conditions

- **Attribution** — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- **Noncommercial** — You may not use this work for commercial purposes.
- **Share Alike** — If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one.

THE END