Math 1070-2: Spring 2008

Lecture 11

Davar Khoshnevisan

Department of Mathematics
University of Utah
http://www.math.utah.edu/~davar

April 9, 2008

This week's reading

Only 2 weeks left

- Complete the material of chapter 8:
  - §8.3–§8.4
  - Skip §8.5–§8.6
  - §9.1

Testing for means

- $H_0: \mu = 0$ versus $H_a: \mu \neq 0$, $H_a: \mu > 0$, or $H_a: \mu < 0$
  [or more generally]
- $H_0: \mu = \mu_0$ versus $H_a: \mu \neq \mu_0$ [μ_0 known, etc.]
- Always check:
  - Assumptions
    [quant. var.; random sample; population $\approx$ normal and/or n large]
  - Hypotheses [of above type]
  - Test statistic:

$$\frac{\bar{x} - \mu_0}{\text{SE}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

[$\approx t$ with $n - 1$ df, if H_0 true]

  - P-value:
    - Two-tail for $H_a: \mu \neq \mu_0$
    - Right-tail for $H_a: \mu > \mu_0$
    - Left-tail for $H_a: \mu < \mu_0$
  - Conclusion

Anorexia example

Example 7, pp. 388–389

- Study different therapies for teenage girls with anorexia
- Weight measured before & after therapy
- Variable: Weight change = wt after $-$ wt before therapy
- $n = 29$ girls
- $\bar{x} = 3.00$ pounds, $s = 7.32$ pounds

Anorexia example

Continued

- Assumptions:
  - Quant. var.
  - Sampling *not* random [interpret results with care]
  - Normal population [could be … tentative]
- Hypotheses: [Is the test effective?]
  - $H_0 : \mu = 0$ versus $H_a : \mu > 0$
- Test statistic:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{3 - 0}{7.32/\sqrt{29}} \approx 2.21.$$

- P-value: [Use t with $n - 1 = 28$ df]

$$2.5\% = 0.025 > P\text{-value} > 0.01 = 1\% \quad \text{[middle]}$$

- Conclusion: P-value $\ll 0.05 \rightarrow$ reject H_0 [even for $H_a : \mu \neq 0$]

Robustness

- If $H_a: \mu \neq \mu_0$ is two-sided then normality not that important
- Normality *important* for one-sided tests. Therefore testing for means OK if:
  - n large;
  - n small, but population normal; or
  - $H_a: \mu \neq \mu_0$ is two-sided
  - Anorexia example: We are 95% confident that $\mu \neq 0$

CI for $p_1 - p_2$

- Point estimate for $p_1 - p_2$ is $\hat{p}_1 - \hat{p}_2$
- Fact:

$$\mathrm{SE} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

- CI for $p_1 - p_2$: [n_1 and n_2 large]

$$(\hat{p}_1 - \hat{p}_2) \pm z\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

An example

pp. 429–431

- Want difference between placebo & aspirin
- $n_1 = 11034$, $n_2 = 11037$ [large ✓]
- $\hat{p}_1 = 0.017$, $\hat{p}_2 = 0.009$
- $\hat{p}_1 - \hat{p}_2 = 0.017 - 0.009 = 0.008$
- $$\begin{aligned} \text{SE} &= \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \\ &= \sqrt{\frac{0.017 \times 0.983}{11034} + \frac{0.009 \times 0.991}{11037}} \approx 0.0015 \end{aligned}$$
- 95% CI for $p_1 - p_2$:

$$0.008 \pm (1.96 \times 0.0015) = (0.005\,, 0.011)$$