STAT 111

Recitation 11

Mo Huang

Email: mohuang@wharton.upenn.edu

Office Hours: Wednesdays 3:00 - 4:00 pm, JMHH F96

Slides: github.com/mohuangx/STAT111-Spring2019

April 19, 2019

Hypothesis testing outline

- ▶ Test of binomial/proportion

- ▶ Test of means/regression
  - ▶ One-sample t test
  - ▶ Two-sample t test (unpaired)
  - ▶ Paired two-sample t test
  - ▶ Regression t test

- ▶ Test of equality of two binomial parameters (two-by-two table)

Regression t test

▶ Linear Regression Model: for the ith observation,

  ▶ Mean of $Y_i = \alpha + \beta x_i$, Variance of $Y_i = \sigma^2$.

  ▶ β is estimated by $b = s_{xy}/s_{xx}$
  ▶ α is estimated by $a = \bar{y} - b\bar{x}$
  ▶ σ^2 is estimated by $s_r^2 = \frac{s_{yy} - b^2 s_{xx}}{n-2}$

▶ We want to test $H_0 : \beta = 0$ (x has no effect on Y).

▶ The test statistic is

$$t = \frac{b}{s_r/\sqrt{s_{xx}}} \quad \text{with } n-2 \text{ degrees of freedom.}$$

Example

- We observe
  $\overline{x} = 21, \overline{y} = 79.7, s_{xx} = 80, s_{yy} = 83.54, s_{xy} = 25, n = 20.$

Step 1 $H_0 : \beta = 0$ vs. $H_1 : \beta \neq 0$. Two-sided test.

Step 2 Choose $\alpha = 0.05$.

Step 3 Test-statistic is

$$t = \frac{b}{s_r / \sqrt{s_{xx}}}$$

$$b = \frac{s_{xy}}{s_{xx}} = \frac{25}{80} = 0.3125$$

$$s_r = \sqrt{\frac{s_{yy} - b^2 s_{xx}}{n - 2}} = \sqrt{\frac{83.54 - 0.3125^2(80)}{20 - 2}} = 2.051$$

$$t = \frac{b}{s_r / \sqrt{s_{xx}}} = \frac{0.3125}{2.051 / \sqrt{80}} = 1.363$$

Example

▶ We observe
$\overline{x} = 21, \overline{y} = 79.7, s_{xx} = 80, s_{yy} = 83.54, s_{xy} = 25, n = 20.$

Step 1 $H_0 : \beta = 0$ vs. $H_1 : \beta \neq 0$. Two-sided test.

Step 2 Choose $\alpha = 0.05$.

Step 3 Test-statistic is $t = 1.363$.

Step 4 Find the critical region.

How many degrees of freedom do we have? $n - 2 = 18$

So we need to look at t_{18}. What is the critical region?

$t \geq t_{18,0.025} = 2.101$ and $t \leq -t_{18,0.025} = -2.101$.

Example

- We observe
  $\overline{x} = 21, \overline{y} = 79.7, s_{xx} = 80, s_{yy} = 83.54, s_{xy} = 25, n = 20.$

Step 1 $H_0 : \beta = 0$ vs. $H_1 : \beta \neq 0$. Two-sided test.

Step 2 Choose $\alpha = 0.05$.

Step 3 Test-statistic is $t = 1.363$.

Step 4 Find the critical region: $t \geq 2.101$ and $t \leq -2.101$

Step 5 Do we reject H_0? No, $t = 1.363$ is not in the critical region.

Testing for equality of two binomial parameters using two-by-two tables

- ▶ Suppose we have two binomial parameters θ_1 and θ_2 and we want to test if they are equal.

- ▶ For example, we want to see if there is a difference in voter turnout between men and women. Let θ_1 be the voter turnout for men and θ_2 be the voter turnout for women. The two-by-two table would be

	voted	did not vote	total
men	o_{11}	o_{12}	r_1
women	o_{21}	o_{22}	r_2
total	c_1	c_2	n

- ▶ We want to test $H_0 : \theta_1 = \theta_2$.

- ▶ The test statistic is

$$z = \frac{(o_{11} \times o_{22} - o_{21} \times o_{12})\sqrt{n}}{\sqrt{r_1 \times r_2 \times c_1 \times c_2}}$$

Example

▶ Suppose we have the following data on voter turnout:

	voted	did not vote	total
men	170	140	310
women	120	110	230
total	290	250	540

Step 1 $H_0 : \theta_1 = \theta_2$ vs. $H_1 : \theta_1 \neq \theta_2$. Two-sided test.

Step 2 Choose $\alpha = 0.05$.

Step 3 Test-statistic is

$$z = \frac{(o_{11} \times o_{22} - o_{21} \times o_{12})\sqrt{n}}{\sqrt{r_1 \times r_2 \times c_1 \times c_2}}$$

$$z = \frac{(170 \times 110 - 120 \times 140)\sqrt{540}}{\sqrt{310 \times 230 \times 290 \times 250}} = 0.6141$$

Example

▶ Suppose we have the following data on voter turnout:

	voted	did not vote	total
men	170	140	310
women	120	110	230
total	290	250	540

Step 1 $H_0 : \theta_1 = \theta_2$ vs. $H_1 : \theta_1 \neq \theta_2$. Two-sided test.

Step 2 Choose $\alpha = 0.05$.

Step 3 Test-statistic is $z = 0.6141$.

Step 4 Find the critical region and p-value.

$z \leq z_{0.025} = -1.96$ and $z \geq z_{0.975} = 1.96$.

p-value $= 2P(Z \geq |0.6141|) \approx 2(0.27) = 0.54$

Example

▶ Suppose we have the following data on voter turnout:

	voted	did not vote	total
men	170	140	310
women	120	110	230
total	290	250	540

Step 1 $H_0 : \theta_1 = \theta_2$ vs. $H_1 : \theta_1 \neq \theta_2$. Two-sided test.

Step 2 Choose $\alpha = 0.05$.

Step 3 Test-statistic is $z = 0.6141$.

Step 4 Find the critical region p-value: $z \leq -1.96$ and $z \geq 1.96$.
p-value $= 0.54$.

Step 5 Do we reject H_0? No, $z = 0.6141$ is not in the critical region and the p-value is greater than 0.05.