STAT 111

Recitation 10

Mo Huang

Email: mohuang@wharton.upenn.edu

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

Slides (adapted from Gemma Moran): `github.com/mohuangx/STAT111-Fall2018`

November 16, 2018

Two-sample t-test

- Example: Suppose we have two classes, let's call 201 and 202, and we want to see if there is a difference in height between these two classes. How do we test this?
- Answer: Let $X_{11}, \ldots, X_{1m}$ represent the heights of the m students in 201 with mean μ_1 and let $X_{21}, \ldots, X_{2n}$ represent the heights of the n students in 202 with mean μ_2. We can test $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$.
- This is an example of a two-sample t-test!
- Two-sample t-test: Let $X_{11}, \ldots, X_{1m}$ be i.i.d random variables with (unknown) mean μ_1 and (unknown) variance σ^2. Let $X_{21}, \ldots, X_{2n}$ be i.i.d random variables with (unknown) mean μ_2 and (unknown) variance σ^2. We want to test whether or not $\mu_1 = \mu_2$.

Example

- Suppose we want to test whether there is a difference in height between class 201 $(X_{11}, \ldots, X_{1m})$ and 202 $(X_{21}, \ldots, X_{2n})$. We observe that $\bar{x}_1 = 66.7$, $s_1^2 = 10.5$, $m = 28$, and $\bar{x}_2 = 65.6$, $s_2^2 = 12.3$, $n = 34$.

Step 1 $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$. Two-sided test.

Step 2 Choose $\alpha = 0.05$.

Step 3 Test-statistic is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{m} + \frac{1}{n}}}, \text{ where } s = \sqrt{\frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2}}$$

$$t = \frac{66.7 - 65.6}{3.390\sqrt{\frac{1}{28} + \frac{1}{34}}} = 1.254$$

Example

- Suppose we want to test whether there is a difference in height between class 201 ($X_{11}, \ldots, X_{1m}$) and 202 ($X_{21}, \ldots, X_{2n}$). We observe that $\bar{x}_1 = 66.7$, $s_1^2 = 10.5$, $m = 28$, and $\bar{x}_2 = 65.6$, $s_2^2 = 12.3$, $n = 34$.

Step 1 $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$. Two-sided test.

Step 2 Choose $\alpha = 0.05$.

Step 3 Test-statistic is $t = 1.254$.

Step 4 Find the critical region.

How many degrees of freedom do we have? $m + n - 2 = 60$

So we need to look at t_{60}. What is the critical region?

$t \geq t_{60,0.025} = 2.000$ and $t \leq -t_{60,0.025} = -2.000$.

Example

- Suppose we want to test whether there is a difference in height between class 201 $(X_{11}, \ldots, X_{1m})$ and 202 $(X_{21}, \ldots, X_{2n})$. We observe that $\bar{x}_1 = 66.7$, $s_1^2 = 10.5$, $m = 28$, and $\bar{x}_2 = 65.6$, $s_2^2 = 12.3$, $n = 34$.

Step 1 $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$. Two-sided test.

Step 2 Choose $\alpha = 0.05$.

Step 3 Test-statistic is $t = 1.254$.

Step 4 Find the critical region: $t \geq 2.000$ and $t \leq -2.000$

Step 5 Do we reject H_0? No, $t = 1.254$ is not in the critical region.