GPCO 453: Quantitative Methods I

Additional Notes on Comparing Two Means

Shane Xinyang Xuan[1]
ShaneXuan.com

November 27, 2017

[1]Department of Political Science, UC San Diego, 9500 Gilman Drive #0521.

Comparing Two Means

- Test the hypothesis at $\alpha = 0.05$ that the average temperature in San Diego is higher than the average temperature in San Francisco. A random sample of 33 days and 37 days is obtained from San Diego and San Francisco. Note that $\overline{x}_{SD} = 72, s_{SD} = 10, \overline{x}_{SF} = 65,\ s_{SF} = 12.$

Comparing Two Means

- Test the hypothesis at $\alpha = 0.05$ that the average temperature in San Diego is higher than the average temperature in San Francisco. A random sample of 33 days and 37 days is obtained from San Diego and San Francisco. Note that $\overline{x}_{SD} = 72, s_{SD} = 10, \overline{x}_{SF} = 65,\ s_{SF} = 12.$
- First, we want to write down the null and alternative hypotheses. By definition, the null hypothesis contains the equal sign, and the alternative hypothesis does not contain the equal sign. The claim in the problem does not have an equal sign, so we treat it as the alternative hypothesis:

$$H_0 : \mu_{SD} - \mu_{SF} \leq 0; \tag{1}$$

$$H_A : \mu_{SD} - \mu_{SF} > 0 \tag{2}$$

Comparing Two Means

- Test the hypothesis at $\alpha = 0.05$ that the average temperature in San Diego is higher than the average temperature in San Francisco. A random sample of 33 days and 37 days is obtained from San Diego and San Francisco. Note that $\overline{x}_{SD} = 72, s_{SD} = 10, \overline{x}_{SF} = 65,\ s_{SF} = 12.$
- First, we want to write down the null and alternative hypotheses. By definition, the null hypothesis contains the equal sign, and the alternative hypothesis does not contain the equal sign. The claim in the problem does not have an equal sign, so we treat it as the alternative hypothesis:

$$H_0 : \mu_{SD} - \mu_{SF} \leq 0; \tag{1}$$

$$H_A : \mu_{SD} - \mu_{SF} > 0 \tag{2}$$

- Note that this is a right-tailed test.

Comparing Two Means: Cont.

- Next, we want to calculate the t-statistic:

$$t = \frac{\overline{x}_1 - \overline{x}_2 - D_0}{\sigma_{\overline{x}_1 - \overline{x}_2}} \tag{3}$$

$$= \frac{72 - 65 - 0}{\sqrt{\frac{10^2}{33} + \frac{12^2}{37}}} = 2.66 \tag{4}$$

Comparing Two Means: Cont.

- Next, we want to calculate the t-statistic:

$$t = \frac{\overline{x}_1 - \overline{x}_2 - D_0}{\sigma_{\overline{x}_1 - \overline{x}_2}} \tag{3}$$

$$= \frac{72 - 65 - 0}{\sqrt{\frac{10^2}{33} + \frac{12^2}{37}}} = 2.66 \tag{4}$$

- Now, we want to find the critical value. We need degree of freedom and α for this. We first find degree of freedom using the formula available in lecture/section slides. We find that the degree of freedom is 68.

Comparing Two Means: Cont.

- Next, we want to calculate the t-statistic:

$$t = \frac{\overline{x}_1 - \overline{x}_2 - D_0}{\sigma_{\overline{x}_1 - \overline{x}_2}} \tag{3}$$

$$= \frac{72 - 65 - 0}{\sqrt{\frac{10^2}{33} + \frac{12^2}{37}}} = 2.66 \tag{4}$$

- Now, we want to find the critical value. We need degree of freedom and α for this. We first find degree of freedom using the formula available in lecture/section slides. We find that the degree of freedom is 68.
- We want to conduct the test at 95% confidence level. Since this is a right-tailed test, the critical value is simply $t_\alpha = t_{0.05}$.

Comparing Two Means: Cont.

- Find the critical value:

Degree of Freedom	...	**Area in Upper Tail = 0.05**
⋮	⋱	⋮
68	...	1.668

Comparing Two Means: Cont.

- Find the critical value:

Degree of Freedom	...	Area in Upper Tail = 0.05
⋮	⋱	⋮
68	...	1.668

- Compare t-statistic to the critical value:

$$2.66 > 1.668 \tag{5}$$

Note that t-statistic falls in the rejection area.

Comparing Two Means: Cont.

- Find the critical value:

Degree of Freedom	...	Area in Upper Tail = 0.05
⋮	⋱	⋮
68	...	1.668

- Compare t-statistic to the critical value:

$$2.66 > 1.668 \tag{5}$$

  Note that t-statistic falls in the rejection area.

- We conclude that we can reject the null hypothesis and the average SD temperature is higher than the average SF temperature

Comparing Two Means: Tweaks

- Now, consider a different case. Test the hypothesis at $\alpha = 0.05$ that the average temperature in San Diego is higher or equal to the average temperature in San Francisco. Note that $n_{SD} = 33, n_{SF} = 37, \overline{x}_{SD} = 72, s_{SD} = 10, \overline{x}_{SF} = 65$, $s_{SF} = 12$.

Comparing Two Means: Tweaks

- Now, consider a different case. Test the hypothesis at $\alpha = 0.05$ that the average temperature in San Diego is higher or equal to the average temperature in San Francisco. Note that $n_{SD} = 33, n_{SF} = 37, \overline{x}_{SD} = 72, s_{SD} = 10, \overline{x}_{SF} = 65$, $s_{SF} = 12$.
- The claim in the problem does have an equal sign, so we treat it as the null hypothesis:

$$H_0 : \mu_{SD} - \mu_{SF} \geq 0; \quad (6)$$

$$H_A : \mu_{SD} - \mu_{SF} < 0 \quad (7)$$

Comparing Two Means: Tweaks

- Now, consider a different case. Test the hypothesis at $\alpha = 0.05$ that the average temperature in San Diego is higher or equal to the average temperature in San Francisco. Note that $n_{SD} = 33, n_{SF} = 37, \overline{x}_{SD} = 72, s_{SD} = 10, \overline{x}_{SF} = 65$, $s_{SF} = 12$.
- The claim in the problem does have an equal sign, so we treat it as the null hypothesis:

$$H_0 : \mu_{SD} - \mu_{SF} \geq 0; \quad (6)$$
$$H_A : \mu_{SD} - \mu_{SF} < 0 \quad (7)$$

- Note that this is a left-tailed test.

Comparing Two Means: Tweaks Cont.

- As before, $t = \frac{\overline{x}_1 - \overline{x}_2 - D_0}{\sigma_{\overline{x}_1 - \overline{x}_2}} = \frac{72-65-0}{\sqrt{\frac{10^2}{33} + \frac{12^2}{37}}} = 2.66$, and $d.f. = 68$

Comparing Two Means: Tweaks Cont.

- As before, $t = \frac{\overline{x}_1 - \overline{x}_2 - D_0}{\sigma_{\overline{x}_1 - \overline{x}_2}} = \frac{72 - 65 - 0}{\sqrt{\frac{10^2}{33} + \frac{12^2}{37}}} = 2.66$, and $d.f. = 68$
- Since this is a left-tailed test, we want to find the critical value when the area on the left is 0.05, which means that the area in the upper tail is 0.95.

Comparing Two Means: Tweaks Cont.

- As before, $t = \frac{\overline{x}_1 - \overline{x}_2 - D_0}{\sigma_{\overline{x}_1 - \overline{x}_2}} = \frac{72-65-0}{\sqrt{\frac{10^2}{33} + \frac{12^2}{37}}} = 2.66$, and $d.f. = 68$
- Since this is a left-tailed test, we want to find the critical value when the area on the left is 0.05, which means that the area in the upper tail is 0.95.
- Since the critical value when the area in the upper tail is 0.05 is 1.668, we know that the critical value when the area in the upper tail is 0.95 should be -1.668.

Comparing Two Means: Tweaks Cont.

- As before, $t = \frac{\overline{x}_1 - \overline{x}_2 - D_0}{\sigma_{\overline{x}_1 - \overline{x}_2}} = \frac{72-65-0}{\sqrt{\frac{10^2}{33} + \frac{12^2}{37}}} = 2.66$, and $d.f. = 68$
- Since this is a left-tailed test, we want to find the critical value when the area on the left is 0.05, which means that the area in the upper tail is 0.95.
- Since the critical value when the area in the upper tail is 0.05 is 1.668, we know that the critical value when the area in the upper tail is 0.95 should be -1.668.
- Compare t-statistic to the critical value:

$$2.66 > -1.668 \tag{8}$$

Note that t-statistic does not fall in the rejection area.

Comparing Two Means: Tweaks Cont.

- As before, $t = \frac{\overline{x}_1 - \overline{x}_2 - D_0}{\sigma_{\overline{x}_1 - \overline{x}_2}} = \frac{72-65-0}{\sqrt{\frac{10^2}{33} + \frac{12^2}{37}}} = 2.66$, and $d.f. = 68$
- Since this is a left-tailed test, we want to find the critical value when the area on the left is 0.05, which means that the area in the upper tail is 0.95.
- Since the critical value when the area in the upper tail is 0.05 is 1.668, we know that the critical value when the area in the upper tail is 0.95 should be -1.668.
- Compare t-statistic to the critical value:

$$2.66 > -1.668 \tag{8}$$

  Note that t-statistic does not fall in the rejection area.
- We conclude that we fail to reject the null hypothesis that San Diego is hotter than or as hot as San Francisco.