GPCO 453: Quantitative Methods I

Review: Hypothesis Testing

Shane Xinyang Xuan[1]
ShaneXuan.com

December 6, 2017

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

Thought Process

- One population or two populations?

Thought Process

- One population or two populations?
- Mean, proportion, or variance?

Thought Process

- One population or two populations?
- Mean, proportion, or variance?
- Right, left, or two tails?

Thought Process

- One population or two populations?
- Mean, proportion, or variance?
- Right, left, or two tails?
- Use z-score or t-score?

Thought Process

- One population or two populations?
- Mean, proportion, or variance?
- Right, left, or two tails?
- Use z-score or t-score?
- Here is what we will cover in the next 30 minutes

Thought Process

- One population or two populations?
- Mean, proportion, or variance?
- Right, left, or two tails?
- Use z-score or t-score?
- Here is what we will cover in the next 30 minutes
  - One population mean, known $\sigma \rightarrow z; s.e. = \frac{\sigma}{\sqrt{n}}$

Thought Process

- One population or two populations?
- Mean, proportion, or variance?
- Right, left, or two tails?
- Use z-score or t-score?
- Here is what we will cover in the next 30 minutes
  - One population mean, known $\sigma \to z; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population mean, unknown $\sigma \to t; s.e. = \frac{\sigma}{\sqrt{n}}$

Thought Process

- One population or two populations?
- Mean, proportion, or variance?
- Right, left, or two tails?
- Use z-score or t-score?
- Here is what we will cover in the next 30 minutes
  - One population mean, known $\sigma \rightarrow z; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population mean, unknown $\sigma \rightarrow t; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population proportion $\rightarrow z; s.e. = \sqrt{\frac{p_0(1-p_0)}{n}}$

Thought Process

- One population or two populations?
- Mean, proportion, or variance?
- Right, left, or two tails?
- Use z-score or t-score?
- Here is what we will cover in the next 30 minutes
  - One population mean, known $\sigma \rightarrow z; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population mean, unknown $\sigma \rightarrow t; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population proportion $\rightarrow z; s.e. = \sqrt{\frac{p_0(1-p_0)}{n}}$
  - One population variance $\rightarrow \chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$; critical value approach

Thought Process

- One population or two populations?
- Mean, proportion, or variance?
- Right, left, or two tails?
- Use z-score or t-score?
- Here is what we will cover in the next 30 minutes
  - One population mean, known $\sigma \to z; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population mean, unknown $\sigma \to t; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population proportion $\to z; s.e. = \sqrt{\frac{p_0(1-p_0)}{n}}$
  - One population variance $\to \chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$; critical value approach
  - Two population variances $\to F = \frac{s_1^2}{s_2^2}$; critical value approach

Thought Process

- One population or two populations?
- Mean, proportion, or variance?
- Right, left, or two tails?
- Use z-score or t-score?
- Here is what we will cover in the next 30 minutes
  - One population mean, known $\sigma \rightarrow z; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population mean, unknown $\sigma \rightarrow t; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population proportion $\rightarrow z; s.e. = \sqrt{\frac{p_0(1-p_0)}{n}}$
  - One population variance $\rightarrow \chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$; critical value approach
  - Two population variances $\rightarrow F = \frac{s_1^2}{s_2^2}$; critical value approach
  - Two population means, known $\sigma \rightarrow z; s.e. = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Thought Process

- One population or two populations?
- Mean, proportion, or variance?
- Right, left, or two tails?
- Use z-score or t-score?
- Here is what we will cover in the next 30 minutes
  - One population mean, known $\sigma \to z; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population mean, unknown $\sigma \to t; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population proportion $\to z; s.e. = \sqrt{\frac{p_0(1-p_0)}{n}}$
  - One population variance $\to \chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$; critical value approach
  - Two population variances $\to F = \frac{s_1^2}{s_2^2}$; critical value approach
  - Two population means, known $\sigma \to z; s.e. = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
  - Two population means, unknown $\sigma \to t; s.e. = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

Thought Process

- One population or two populations?
- Mean, proportion, or variance?
- Right, left, or two tails?
- Use z-score or t-score?
- Here is what we will cover in the next 30 minutes
  - One population mean, known $\sigma \rightarrow z; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population mean, unknown $\sigma \rightarrow t; s.e. = \frac{\sigma}{\sqrt{n}}$
  - One population proportion $\rightarrow z; s.e. = \sqrt{\frac{p_0(1-p_0)}{n}}$
  - One population variance $\rightarrow \chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$; critical value approach
  - Two population variances $\rightarrow F = \frac{s_1^2}{s_2^2}$; critical value approach
  - Two population means, known $\sigma \rightarrow z; s.e. = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
  - Two population means, unknown $\sigma \rightarrow t; s.e. = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
  - Two population proportions $\rightarrow z; s.e. = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

What We Won't Cover

- Important! You should be able to read z table, t table, χ^2 table, and F table; and you should be able to compute degree of freedoms for all the above cases

What We Won't Cover

- Important! You should be able to read z table, t table, χ^2 table, and F table; and you should be able to compute degree of freedoms for all the above cases
- For each case, you should do at least one example – We suggest that you go to the respective section in the textbook, and follow the example to make sure that you *get* it

One Population Mean, known σ

	Lower Tail Test	Upper Tail Test	Two-Tailed Test
Hypotheses	$H_0: \mu \geq \mu_0$ $H_a: \mu < \mu_0$	$H_0: \mu \leq \mu_0$ $H_a: \mu > \mu_0$	$H_0: \mu = \mu_0$ $H_a: \mu \neq \mu_0$
Test Statistic	$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$
Rejection Rule: *p*-Value Approach	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$
Rejection Rule: Critical Value Approach	Reject H_0 if $z \leq -z_\alpha$	Reject H_0 if $z \geq z_\alpha$	Reject H_0 if $z \leq -z_{\alpha/2}$ or if $z \geq z_{\alpha/2}$

One Population Mean, unknown σ

	Lower Tail Test	Upper Tail Test	Two-Tailed Test
Hypotheses	$H_0\colon \mu \geq \mu_0$ $H_a\colon \mu < \mu_0$	$H_0\colon \mu \leq \mu_0$ $H_a\colon \mu > \mu_0$	$H_0\colon \mu = \mu_0$ $H_a\colon \mu \neq \mu_0$
Test Statistic	$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$	$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$	$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$
Rejection Rule: *p*-Value Approach	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$
Rejection Rule: Critical Value Approach	Reject H_0 if $t \leq -t_\alpha$	Reject H_0 if $t \geq t_\alpha$	Reject H_0 if $t \leq -t_{\alpha/2}$ or if $t \geq t_{\alpha/2}$

One Population Proportion

	Lower Tail Test	Upper Tail Test	Two-Tailed Test
Hypotheses	$H_0: p \geq p_0$ $H_a: p < p_0$	$H_0: p \leq p_0$ $H_a: p > p_0$	$H_0: p = p_0$ $H_a: p \neq p_0$
Test Statistic	$z = \dfrac{\bar{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$	$z = \dfrac{\bar{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$	$z = \dfrac{\bar{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$
Rejection Rule: *p*-Value Approach	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$
Rejection Rule: Critical Value Approach	Reject H_0 if $z \leq -z_\alpha$	Reject H_0 if $z \geq z_\alpha$	Reject H_0 if $z \leq -z_{\alpha/2}$ or if $z \geq z_{\alpha/2}$

One Population Variance

TABLE 11.2 SUMMARY OF HYPOTHESIS TESTS ABOUT A POPULATION VARIANCE

	Lower Tail Test	Upper Tail Test	Two-Tailed Test
Hypotheses	$H_0\colon \sigma^2 \geq \sigma_0^2$ $H_a\colon \sigma^2 < \sigma_0^2$	$H_0\colon \sigma^2 \leq \sigma_0^2$ $H_a\colon \sigma^2 > \sigma_0^2$	$H_0\colon \sigma^2 = \sigma_0^2$ $H_a\colon \sigma^2 \neq \sigma_0^2$
Test Statistic	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$
Rejection Rule: *p*-Value Approach	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$
Rejection Rule: Critical Value Approach	Reject H_0 if $\chi^2 \leq \chi^2_{(1-\alpha)}$	Reject H_0 if $\chi^2 \geq \chi^2_{\alpha}$	Reject H_0 if $\chi^2 \leq \chi^2_{(1-\alpha/2)}$ or if $\chi^2 \geq \chi^2_{\alpha/2}$

Two Population Variances

	Upper Tail Test	Two-Tailed Test
Hypotheses	$H_0: \sigma_1^2 \le \sigma_2^2$ $H_a: \sigma_1^2 > \sigma_2^2$	$H_0: \sigma_1^2 = \sigma_2^2$ $H_a: \sigma_1^2 \ne \sigma_2^2$ Note: Population 1 has the larger sample variance
Test Statistic	$F = \frac{s_1^2}{s_2^2}$	$F = \frac{s_1^2}{s_2^2}$
Rejection Rule: *p*-Value Approach	Reject H_0 if p-value $\le \alpha$	Reject H_0 if p-value $\le \alpha$
Rejection Rule: Critical Value Approach	Reject H_0 if $F \ge F_\alpha$	Reject H_0 if $F \ge F_{\alpha/2}$

Comparing Two Populations

- Standard error

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \tag{1}$$

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \tag{2}$$

Comparing Two Populations

- Standard error

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \tag{1}$$

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \tag{2}$$

- Confidence interval

$$(\bar{x}_1 - \bar{x}_2) \pm z \times \sigma_{\bar{x}_1-\bar{x}_2} \tag{3}$$

$$(\hat{p}_1 - \hat{p}_2) \pm z \times \sigma_{\hat{p}_1-\hat{p}_2} \tag{4}$$

Some “Tweaks”

- What if σ is unknown?

Some "Tweaks"

- What if σ is unknown?

  Use s to replace σ, and calculate t-score instead of z-score

Some "Tweaks"

- What if σ is unknown?
  Use s to replace σ, and calculate t-score instead of z-score
- How to calculate t-score again?

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - D_0}{\sigma_{\overline{x}_1 - \overline{x}_2}} \tag{5}$$

Some "Tweaks"

- What if σ is unknown?
  Use s to replace σ, and calculate t-score instead of z-score
- How to calculate t-score again?

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - D_0}{\sigma_{\overline{x}_1 - \overline{x}_2}} \tag{5}$$

- Degree of freedom is calculated by

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} \tag{6}$$

Some "Tweaks"

- What if σ is unknown?
  Use s to replace σ, and calculate t-score instead of z-score
- How to calculate t-score again?

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - D_0}{\sigma_{\overline{x}_1 - \overline{x}_2}} \tag{5}$$

- Degree of freedom is calculated by

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} \tag{6}$$

- If we do not know p, we use $\hat{p}$ instead:

$$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}} \tag{7}$$

Concluding Remarks

– It might sound really easy, but you need to practice

Concluding Remarks

- It might sound really easy, but you need to practice
- If you have not done teaching evaluations for the TAs yet, please consider doing so before the deadline (Monday)

Concluding Remarks

- It might sound really easy, but you need to practice
- If you have not done teaching evaluations for the TAs yet, please consider doing so before the deadline (Monday)
- All section materials available at `ShaneXuan.com/teaching`