Econ 103: Statistics for Economists

Hypothesis Testing Quiz

Mallick Hossain

R Project Logistics

- When is the R project due?

R Project Logistics

- When is the R project due?
  - Today!

R Project Logistics

R Project Logistics

Final Exam Logistics

- When is the final exam?

Final Exam Logistics

- When is the final exam?
  - Wednesday, December 21.

Final Exam Logistics

- When is the final exam?
  - Wednesday, December 21.
  - Exam will start promptly at 5 and end at 7

Final Exam Logistics

- When is the final exam?
    - Wednesday, December 21.
    - Exam will start promptly at 5 and end at 7
    - In SHDH 1206!

Final Exam Logistics

- When is the final exam?
  - Wednesday, December 21.
  - Exam will start promptly at 5 and end at 7
  - In SHDH 1206!
  - I will not give any time extensions if you're late. You have been warned.

Final Exam Logistics

- When is the final exam?
  - Wednesday, December 21.
  - Exam will start promptly at 5 and end at 7
  - In SHDH 1206!
  - I will not give any time extensions if you're late. You have been warned.
  - Comprehensive. Will have slightly more emphasis on the most recent material, but because of the cumulative nature of the material, those questions will likely pull on knowledge covered earlier in the course.

Final Exam Logistics

- When is the final exam?
  - Wednesday, December 21.
  - Exam will start promptly at 5 and end at 7
  - In SHDH 1206!
  - I will not give any time extensions if you're late. You have been warned.
  - Comprehensive. Will have slightly more emphasis on the most recent material, but because of the cumulative nature of the material, those questions will likely pull on knowledge covered earlier in the course.
- Will you have extra office hours?

Final Exam Logistics

- When is the final exam?
    - Wednesday, December 21.
    - Exam will start promptly at 5 and end at 7
    - In SHDH 1206!
    - I will not give any time extensions if you're late. You have been warned.
    - Comprehensive. Will have slightly more emphasis on the most recent material, but because of the cumulative nature of the material, those questions will likely pull on knowledge covered earlier in the course.
- Will you have extra office hours?
    - Yes, I'll extend my normal office hours during reading week.

Final Exam Logistics

- When is the final exam?
  - Wednesday, December 21.
  - Exam will start promptly at 5 and end at 7
  - In SHDH 1206!
  - I will not give any time extensions if you're late. You have been warned.
  - Comprehensive. Will have slightly more emphasis on the most recent material, but because of the cumulative nature of the material, those questions will likely pull on knowledge covered earlier in the course.
- Will you have extra office hours?
  - Yes, I'll extend my normal office hours during reading week.
  - 8-10am on Tuesday of reading week

Final Exam Logistics

- When is the final exam?
  - Wednesday, December 21.
  - Exam will start promptly at 5 and end at 7
  - In SHDH 1206!
  - I will not give any time extensions if you're late. You have been warned.
  - Comprehensive. Will have slightly more emphasis on the most recent material, but because of the cumulative nature of the material, those questions will likely pull on knowledge covered earlier in the course.
- Will you have extra office hours?
  - Yes, I'll extend my normal office hours during reading week.
  - 8-10am on Tuesday of reading week
  - 3-5pm on Wednesday of reading week

Class Logistics

- I got this course evaluation email. I just trashed it. Is that okay?

Class Logistics

- I got this course evaluation email. I just trashed it. Is that okay?
  - No! Please fill it out and give me comments on how to improve the course.

Questions

1. What is power?

Questions

1. What is power?
    - The probability of rejecting the null when it is false

Questions

1. What is power?
    - The probability of rejecting the null when it is false
2. How are hypothesis testing and confidence intervals related?

Questions

1. What is power?
    - The probability of rejecting the null when it is false
2. How are hypothesis testing and confidence intervals related?
    - For a given significance level, if you reject the null hypothesis, that means your test statistic lies outside of the corresponding confidence interval. Alternatively, if you construct a confidence level for a given significance level, you will reject your null hypothesis if your test statistic lies outside of the confidence interval.

Questions

3. What is the p-value?

Questions

3. What is the p-value?
    - This is the probability of observing a test statistic at least as extreme as the one we actually observed. Alternatively, this is the smallest significance level at which we could have rejected the null.

Questions

3. What is the p-value?
    - This is the probability of observing a test statistic at least as extreme as the one we actually observed. Alternatively, this is the smallest significance level at which we could have rejected the null.
4. If we are testing whether our populations are the same, what can we do to improve our estimates?

Questions

3. What is the p-value?
   - This is the probability of observing a test statistic at least as extreme as the one we actually observed. Alternatively, this is the smallest significance level at which we could have rejected the null.
4. If we are testing whether our populations are the same, what can we do to improve our estimates?
   - You can pool your sample together. This only works if you are testing if the populations are the SAME. We were only able to do this in the Bernoulli case because a Bernoulli RV is completely determined by its parameter p. In cases where we do not know the underlying distribution, you cannot do this.

Old Exam Questions (Spring 2013 Question 7)

Let $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ and suppose we know that $\sigma^2 = 1$

1. Write down the sampling distribution of $\sqrt{n}(\bar{X}_n - \mu)$

Old Exam Questions (Spring 2013 Question 7)

Let $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ and suppose we know that $\sigma^2 = 1$

1. Write down the sampling distribution of $\sqrt{n}(\bar{X}_n - \mu)$
    - $N(0, 1)$

Old Exam Questions (Spring 2013 Question 7)

Let $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ and suppose we know that $\sigma^2 = 1$

1. Write down the sampling distribution of $\sqrt{n}(\bar{X}_n - \mu)$
   - $N(0, 1)$
2. Using your answer to (1), derive the sampling distribution of $\sqrt{n}\bar{X}_n$.

Old Exam Questions (Spring 2013 Question 7)

Let $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ and suppose we know that $\sigma^2 = 1$

1. Write down the sampling distribution of $\sqrt{n}(\bar{X}_n - \mu)$
   - $N(0, 1)$
2. Using your answer to (1), derive the sampling distribution of $\sqrt{n}\bar{X}_n$.
   - $\sqrt{n}(\bar{X}_n - \mu) = \sqrt{n}\bar{X}_n - \sqrt{n}\mu$

Old Exam Questions (Spring 2013 Question 7)

Let $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ and suppose we know that $\sigma^2 = 1$

1. Write down the sampling distribution of $\sqrt{n}(\bar{X}_n - \mu)$
   - $N(0, 1)$
2. Using your answer to (1), derive the sampling distribution of $\sqrt{n}\bar{X}_n$.
   - $\sqrt{n}(\bar{X}_n - \mu) = \sqrt{n}\bar{X}_n - \sqrt{n}\mu$
   - $\sqrt{n}\bar{X}_n = \sqrt{n}(\bar{X}_n - \mu) + \sqrt{n}\mu = N(0, 1) + \sqrt{n}\mu \sim N(\sqrt{n}\mu, 1)$

Old Exam Questions (Spring 2013 Question 7)

Let $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ and suppose we know that $\sigma^2 = 1$

1. Write down the sampling distribution of $\sqrt{n}(\bar{X}_n - \mu)$
   - $N(0, 1)$
2. Using your answer to (1), derive the sampling distribution of $\sqrt{n}\bar{X}_n$.
   - $\sqrt{n}(\bar{X}_n - \mu) = \sqrt{n}\bar{X}_n - \sqrt{n}\mu$
   - $\sqrt{n}\bar{X}_n = \sqrt{n}(\bar{X}_n - \mu) + \sqrt{n}\mu = N(0, 1) + \sqrt{n}\mu \sim N(\sqrt{n}\mu, 1)$
3. Suppose we wanted to test the null hypothesis that $\mu = 0$ against the two-sided alternative at the 5% level. What test statistic should we use and what should our decision rule be?

Old Exam Questions (Spring 2013 Question 7)

Let $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ and suppose we know that $\sigma^2 = 1$

1. Write down the sampling distribution of $\sqrt{n}(\bar{X}_n - \mu)$
   - $N(0,1)$
2. Using your answer to (1), derive the sampling distribution of $\sqrt{n}\bar{X}_n$.
   - $\sqrt{n}(\bar{X}_n - \mu) = \sqrt{n}\bar{X}_n - \sqrt{n}\mu$
   - $\sqrt{n}\bar{X}_n = \sqrt{n}(\bar{X}_n - \mu) + \sqrt{n}\mu = N(0,1) + \sqrt{n}\mu \sim N(\sqrt{n}\mu, 1)$
3. Suppose we wanted to test the null hypothesis that $\mu = 0$ against the two-sided alternative at the 5% level. What test statistic should we use and what should our decision rule be?
   - Under the null, $T = \sqrt{n}\bar{X}_n \sim N(0,1)$ and we reject if $|T| \geq \texttt{qnorm(0.975)} \approx 2$.

Questions

4. How would your answer to (3) change if we tested instead against the one-sided alternative that $\mu > 0$?

Questions

4. How would your answer to (3) change if we tested instead against the one-sided alternative that $\mu > 0$?
   - We still examine the test statistic $T = \sqrt{n}\bar{X}_n$ except in this case the decision rule is: reject if $T \geq$ `qnorm`(0.95). This critical value is *less* than 2.

Questions

5. Derive the power of the hypothesis test from (3) if $\sigma = 1$, $n = 25$ and $\mu = 1$. Your answer should be given in terms of the relevant R commands.

Questions

5. Derive the power of the hypothesis test from (3) if $\sigma = 1$, $n = 25$ and $\mu = 1$. Your answer should be given in terms of the relevant R commands.
    - Regardless of the true value of μ, we know from the solution to part (c) that $\sqrt{n}\bar{X}_n \sim N(\sqrt{n}\mu, 1)$. Power equals the probability of rejecting the null *when it is false*. Here we are asked to suppose that the true mean is one rather than zero, as assumed under the null, and that the sample size is 25. Hence,

      $$T = \sqrt{n}\bar{X}_n \sim N(5, 1)$$

      The decision rule from above was: reject if $|T| \geq 2$, so we simply need to calculate the probability that a $N(5, 1)$ random variable is greater than 2 or less than -2. Since these are mutually exclusive events, the probabilities sum. Hence, the R command to calculate the power are is follows:

      ```
      pnorm(-2, mean = 5, sd = 1) + (1 - pnorm(2, mean = 5, sd = 1))
      ```

Ready for the next lecture?

Let’s Do This!