PS 2010: 10. Estimation and Inference 2

Qing Chang

Department of Political Science
University of Pittsburgh

Fall 2023

Agenda

- Homework 6 Due Nov 21th
- Hypothesis Testing
- R Summarize Function

Hypothesis Testing

- A statistical hypothesis is an assertion or conjecture concerning one or more populations.
- We use data from random sample to test a hypothesis
- This is a probability issue
  - Rejection of a hypothesis implies that the sample evidence refutes it
  - We could be wrong, but the probability is very small
- Support for one theory is obtained by showing lack of support for its converse (in a sense, a proof by contradiction)
- Example: Vaccine, election prediction

Elements of a Statistical Test

- **Null Hypothesis (H_0):**
  - The hypothesis to be tested
  - The opposite of the H_1
  - We **reject or fail to reject** H_0
- **Alternative Hypothesis (H_1):**
  - The hypothesis to be accepted in case H_0 is rejected
  - Represents the question to be answered or the theory to be tested
- **Test statistic**: a function of sample which the statistical decision will be based
- **Rejection region**: specifies the values of the test statistic for which the null hypothesis is to be rejected in favor of the alternative hypothesis.
- Example: Jury trial, high school example again

Rejection Region

- Choosing RR is very important
- Two errors we may make
- **Type I error:** Rejection of the null hypothesis when it is true
  - P(Type I error) $= \alpha$ called the level of the test
- **Type II error:** Nonrejection of the null hypothesis when it is false
  - P(Type II error) $= \beta$

	H_0 is true	H_0 is false
Do not reject H_0	Correct decision	Type II error
Reject H_0	Type I error	Correct decision

Relation between Type I and II error

- A decrease in the probability of one generally results in an increase in the probability of the other.
- Adjust the RR is to adjust the type I error
- Increase sample size can decrease both errors

Test Concerning a Single Mean for Large Sample

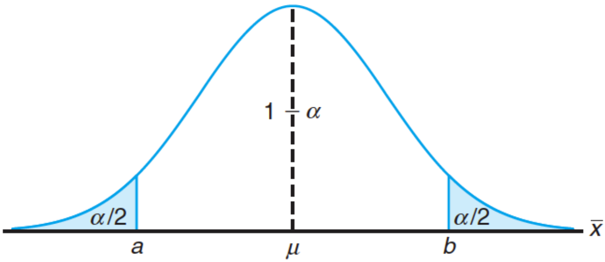
- Suppose X_1, X_2, X_3 ... X_n are randomly sample from a distribution with mean μ and variance σ^2
- We want to test if population $\mu = \mu_0$

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

- To estimate population mean, we could use the mean of sampling distribution based data
- Under large sample size (CLT), we know $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$
- Then, we could calculate $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$, given the null hypothesis

Test Concerning a Single Mean for Large Sample

- We also need to decide α
- For two tail test, we need $\frac{\alpha}{2}$ to determine RR
- For one tail test, we use α to determine RR
- Once we have RR, we also have critical values
- Reject H_0 falls into RR region



Examples

- A random sample of 100 recorded deaths in the United States during the past year showed an average life span of 71.8 years. Assuming a population standard deviation of 8.9 years, does this seem to indicate that the mean life span today is greater than 70 years? Use a 0.05 level of significance.
- A vice president in charge of sales claims that salespeople are averaging no more than 15 sales contacts per week. As a check on his claim, n $=$ 36 salespeople are selected at random, and the number of contacts made by each is recorded for a single randomly selected week. The mean and variance of the 36 measurements were 17 and 9, respectively. Does the evidence contradict the vice president's claim? Use a test with level $\alpha = .05$.

Example

- A manufacturer of sports equipment has developed a new synthetic fishing line that the company claims has a mean breaking strength of 8 kilograms. Test the hypothesis that $\mu = 8$ kilograms against the alternative that $\mu \neq 8$ kilograms if a random sample of 50 lines is tested and found to have a mean breaking strength of 7.8 kilograms with a standard deviation of 0.5 kilogram. Use a 0.01 level of significance.

Test Concerning a Single Mean for Small Sample

- The logic and procedures are the same as in the large sample
- Core difference is when replace population σ^2 with sample s^2, the test statistic becomes to t distribution
- That is: $T = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$
- Notice, we also need to make sure the population distribution is a normal

Example

- A study claimed that a vacuum cleaner uses an average of 46 kilowatt hours per year. If a random sample of 12 homes included in a planned study indicates that vacuum cleaners use an average of 42 kilowatt hours per year with a standard deviation of 11.9 kilowatt hours, does this suggest at the 0.05 level of significance that vacuum cleaners use, on average, less than 46 kilowatt hours annually? Assume the population of kilowatt hours to be normal.

P-value

Definition: If W is a test statistic, the p-value, or attained significance level, is the smallest level of significance α for which the observed data indicate that the null hypothesis should be rejected.

- In all the above hypothesis testing, we pre-select α
- Instead, we can calculate the test statistic first and then calculate the P-value
- We then make decisions
- Previous examples

Examples

- Just prior to an important election, in a random sample of 749 voters, 397 preferred Candidate Y over Candidate Z.
  - Construct a 90% confidence interval for the overall proportion of voters who prefer Candidate Y over Candidate Z.
  - Test if the candidate is going to win less than 58% of vote, and calculate P value.
- In a random sample of 25 direct flights from New York to Boston by Delta Airline, the sample mean flight time was 56 minutes and the sample standard deviation was 8 minutes.
  - Construct a 99% confidence interval for the overall mean flight time on this route. (Assume the flight times are approximately normally distributed.)
  - Test the claim that the mean flight time is more than 50 mins, and calculate P value