CHAPTER 8

Confidence Intervals
Chapter 8 Objectives
The student will be able to

- Calculate and interpret confidence intervals for one population average and one population proportion.
- Interpret the student-t probability distribution as the sample size changes.
- Discriminate between problems applying the normal and the student-t distributions.
Putting a boundary around where we think the population mean or proportion is
- recall that up until now we have calculated a point estimate of the population parameter
- point estimate ± margin of error gives the “confidence interval”
- An example
CONFIDENCE INTERVAL, SINGLE POPULATION MEAN, STANDARD DEVIATION KNOWN

To calculate the confidence interval we need sample average and a margin of error

- use EBM for margin of error
- EBM depends on confidence level (CL)
  - 95% CL implies \( z = 1.96 \)
  - 90% CL implies \( z = 1.645 \)

\[
EBM = z^* \frac{\sigma}{\sqrt{n}}
\]
WHAT THE CALCULATOR CAN DO FOR US

- TI-83, 86 simplify this whole procedure
  - go in through STAT and get to TESTS
    - find Zinterval or Zint1
    - gives the confidence interval (CI)
    - if we want the EBM - (high - low)/2
    - Note: (high + low)/2 = point estimate (i.e. the sample statistic)

- Let’s try one
CONFIDENCE INTERVAL, SINGLE POPULATION MEAN, STANDARD DEVIATION UNKNOWN

- Problems when population standard deviation is unknown and small sample size
  - we will always use $t$ when population standard deviation is unknown
    - New random variable $\bar{X} \sim t_{df}$
  - use Tint1 or Tinterval on calculator
    - data entry the same as for Zinterval

- Let’s try one

\[
EBM = t^* \frac{s}{\sqrt{n}}
\]
P’ ~ N(p, \sqrt{\frac{pq}{n}})

Error Bound of Proportion

\[EBP = z_{\frac{\alpha}{2}} \sqrt{\frac{p'q'}{n}}\]

- Calculator ZPin1 or 1-PropZInt

And now to try one.
Interpret CI in two ways
- We are CL% confident that the true population mean (proportion) is between the CI numbers
- CL% of all CI constructed in this way contain the true mean (proportion)
For population average

- Solve the EBM equation for n
  - $n = \left(\frac{z\sigma}{EBM}\right)^2$

For population proportion

- Solve the EBP equation of n

$$n = \frac{p'q'}{\left(\frac{EBP}{z}\right)^2}$$

- Use ‘conservative’ estimate for $p'$ and $q'$ as 0.50
CHAPTER 9

Hypothesis Testing:
Single Mean and
Single Proportion
The student will be able to

- Conduct hypothesis tests for
  - Single Mean
    - standard normal - $\sigma$ known
    - student-t - $\sigma$ unknown
  - Single Proportion
- Write statements defining
  - Type I error
  - Type II error
Making a decision about a population parameter
- based on two contradictory hypotheses or statements

Conducting a hypothesis test is very structured
- Set up two contradictory hypotheses.
- Determine the correct distribution to perform the hypothesis test.
- Perform the calculations.
- Make a decision and write a conclusion.
Null and Alternative Hypotheses

- $H_o$ - Null hypothesis is what is believed or assumed to be true
  - Read the problem. An implied $=$ is the null hypothesis

- $H_a$ - Alternative hypothesis is contradictory to $H_o$.

- If $H_o$ has $=$, then $H_a$ has $\neq$, $>$, or $<$
- If $H_o$ has $>$, then $H_a$ has $<$
- If $H_o$ has $<$, then $H_a$ has $>$. 
An article in the SJ Mercury News stated that student in the CSU system take an average of 4.5 years to finish their undergraduate degrees. Suppose you believe that the average time is longer. You conduct a survey of 49 CSU students and obtain a sample mean of 5.1 years with a sample standard deviation of 1.2 years. Do the data support your claim at a 1% level of significance?

Set up 2 contradictory hypotheses.
- \( H_0: \mu = 4.5 \)
- \( H_a: \mu > 4.5 \)
HYPOTHESIS TESTING
SINGLE MEAN

- Determine distribution to perform test.
  - Single mean
    \[ \bar{X} \sim N(\mu, \sigma / \sqrt{n}) \]
    \[ \bar{X} \sim t_{df} \]
  - Single proportion
    \[ P' \sim N(p, \sqrt{pq/n}) \]

- What kind of test? If \( H_a \)
  - \( > \), right-tailed test
  - \( < \), left-tailed test
  - \( \neq \), two-tailed test

- Our problem
  \[ \bar{X} \sim t_{48} \]
HYPOTHESIS TESTING
SINGLE MEAN

- Perform Calculation (pvalue)
  - Ztest
  - Ttest
  - Our problem

\[ P(\overline{X} > 5.1) = 0.0005 \]

- Make a decision - Compare pvalue to level of significance \( \alpha \)
  - If \( \alpha > \text{pvalue} \), reject \( H_0 \)
  - If \( \alpha < \text{pvalue} \), do not reject \( H_0 \)
  - If \( \alpha \approx \text{pvalue} \), test inconclusive

- Our problem
  - \( \alpha = 0.01 > \text{pvalue of 0.0005 so reject null.} \)
Write a concluding statement.

- Not a statement of the decision but rather interprets the decision in light of what it means in the problem.

- Our Problem
  - Wrong concluding statement: We reject the null hypothesis.
  - Right concluding statement: There is sufficient evidence to believe that the average time to complete their undergraduate degree is greater than 4.5 years.
HYPOTHESIS TESTING
THE ISSUE OF ERROR

- It is possible that the conclusion reached in a hypothesis test is wrong. We try to minimize this possibility but still....

  - Type I error - rejecting the null when, in fact, the null is true
    - $P(\text{Type I error}) = \alpha$

  - Type II error - don’t reject the null when, in fact, the null is false
    - $P(\text{Type II error}) = \beta$
HYPOTHESIS TESTING
SINGLE PROPORTION

- Same steps but dealing with proportion
  - 1-PropZtest

- An example
CHAPTER 10

Hypothesis Testing:
Two means and
Two proportions
CHAPTER 10 OBJECTIVES

The student will be able to

- Conduct hypothesis test for
  - Two means
  - Two Proportions
  - Matched Pairs
Comparison of Two Populations

- Scientific research usually involves the comparison of two groups
  - placebo versus real drug
  - before and after intervention

- We’ll rely on the technology we have to do our calculations
  - 2-sampZtest
  - 2-sampTtest
  - 2-propZtest
Tests we will address

- Independent groups
  - two population means
  - two population proportions

- Matched or paired samples
  - measurements drawn from same pair of individuals or objects
  - before and after intervention
Two independent sample from two distinct populations.

Both populations are normally distributed with population means and standard deviations unknown.

New random variable
- \( \bar{X}_1 - \bar{X}_2 \) = difference of sample means

Hypotheses compare two means
The following random samples are measurements of the heat-producing capacity (in millions of calories per ton) of specimens of coal from two mines. Use 0.01 level of significance to test whether there is a difference between the two mines.

- **Mine 1:** 8260, 8130, 8350, 8070, 8340
- **Mine 2:** 7950, 7890, 7900, 8140, 7920, 7840
Two Population Means Example

- Ho: \( \mu_1 = \mu_2 \)  
  Ha: \( \mu_1 \neq \mu_2 \)

- Two-tailed test, population standard deviations unknown (don’t pool)

- Calculate p-value.
  - P-value = 0.0035

- Make decision
  - \( \alpha = 0.01 \)  
    p-value = 0.0035
  - reject null

- Concluding statement
  - The two mines do not have coal with the same heat-producing capacity.
Two independent samples from two distinct populations

Comparing the proportions

New random variable
- \( P'_A - P'_B = \text{difference in population proportions} \)

Hypotheses compare two proportions
- null generally states that the two proportions are equal
Want to determine if the percentage of smokers is less as a result of anti-smoking campaign. In 1985, 576/1500 people smoked. In 1990, 652/2000 smoked. Is the percentage of smokers in 1990 less than in 1985?

- **Ho:** $p_{1985} = p_{1990}$
- **Ha:** $p_{1985} > p_{1990}$

Perform calculations
- $p_{value} = 0.0002$
Make decision
- $\alpha = 0.05$  \ pvalue = 0.0002
- reject null

Concluding statement
- The percentage of smokers in 1990 is less than in 1985.
**Matched or Paired Samples**

- Measurements taken from same population before and after some type of intervention.

- New random variable
  - $\overline{X}_d$ = the average difference in measurements.

- Null hypothesis states that there is no improvement, alternate states that there is improvement.

- Always a t-test, use Ttest in calculator.
Example 10.11, page 542

- $H_0: \mu_d \geq 0$  $H_a: \mu_d < 0$
  - if the hypnotism helps (improvement) one would expect that the difference (after - before) to be negative.

- Perform calculations
  - need mean and standard deviation of the differences
    - $\overline{x}_d = -3.13$  $s_d = 2.91$
    - use these in Ttest
    - pvalue = 0.0094
Matched Samples Example

- **Make decision**
  - $\alpha = 0.05 \quad \text{pvalue} = 0.0094$
  - reject null

- **Concluding statement**
  - The sensory measurements are lower after hypnotism. Hypnotism appears to be effective in reducing pain.
What’s fair game
- Chapter 8
- Chapter 9
- Chapter 10

21 multiple choice questions
- The last 3 quarters’ exams

What to bring with you
- Scantron (#2052), pencil, eraser, calculator, 1 sheet of notes (8.5x11 inches, both sides)