MATH 4720 / MSSC 5720

Instructor: Mehdi Maadooliat

Chapter 5 (Part B)

Department of Mathematical and Statistical Sciences

CHAPTER 5 (PART B)

- Confidence Interval (CI)
- CI for μ, when σ is known
- Choosing Sample Size for Estimating μ
- A Statistical Test for μ
- Type I and Type II Errors
- Power of a Test
- Choosing Sample Size for Testing μ
- Level of Significance
- P-value
- Inference about μ, when σ is unknown

HYPOTHESIS TESTING

MARQUETTE UNIVERSITY
Be The Difference.

- **Hypothesis testing is the most used statistical methodology of statistics. Whenever you collect data to test a hypothesis, you would most likely use this methodology. In application, this is how it works,**
  - **1. First build or assume a statistical model for the data.**
  - **2. Write the hypothesis of interest in terms of the parameters of the model.**
  - **3. Use a statistical decision rule (that minimizes the rate of false discovery) to draw conclusion.**

A GENERAL FRAMEWORK:

- **Elements of the hypothesis testing**

1. **Null Hypothesis H_0 (in terms of the parameters)**
2. **Alternative Hypothesis H_a**
   - (This is actually the research hypothesis).
3. **Test Statistics**
4. **Decision Rule (some call it "rejection region")**
5. **Conclusion**

EXAMPLE 1

- A person comes into court charged with a crime. A jury must decide whether the person is innocent (null hypothesis) or guilty (alternative hypothesis). Even though the person is charged with the crime, at the beginning of the trial (and until the jury declares otherwise) the accused is assumed to be innocent. Only if overwhelming evidence of the person's guilt can be shown is the jury expected to declare the person guilty--otherwise the person is considered innocent.
  - H_0: The person is innocent
  - H_a: The person is guilty
  - Test Statistics: Evidence
  - Decision Rule: Jury's decision
  - Conclusion: "guilty" OR "NOT enough evidence to convict"

EXAMPLE 2

- A new diet is developed for weight loss.

- A simple research question: Does this diet work?

- You conduct a clinical trial on 50 subjects with similar characteristics. Among them, a randomly selected 25 subjects went through the new diet for eight weeks and the other 25 received Placebo for eight weeks. At the end of the trial, you determined how many of them lost weight.

EXAMPLE 2 CONT'D

- **Data:**

	Placebo	**New Diet**
	$n_1 = 25$	$n_2 = 25$
Number of subjects who lost weight	$x_1 = 4$	$x_2 = 6$

- **The probability distribution?**

$X_1 \sim$ **Binomial**$(25, \pi_1)$ $X_2 \sim$ **Binomial**$(25, \pi_2)$

- **Null Hypothesis** $H_0: \pi_1 = \pi_2$

  Alternative Hypothesis: $H_a: \pi_1 < \pi_2$

- **Test Statistics and Decision rule for this problem will be discussed later.**

CONCEPTS OF HYPOTHESIS TESTING

- H_0: **Null Hypothesis**
  H_a: **Alternative Hypotheses (Research Hypothesis)**
- **Based on the evidence from the data, either we reject H_0 in favor of H_a or we accept H_0.**

Decision	H_0 is True	H_a is true
Reject H_0	Type-I Error	Correct Decision
Accept H_0	Correct Decision	Type-II Error

- **Back to Example 1:**

Decision	Truth is Person Innocent	Truth is Person Guilty
Jury Decides Person Guilty	Type-I Error	Correct Decision
Jury Decides Person Innocent	Correct Decision	Type-II Error

TYPES OF ERROR

- **Type-I Error:**
  - **Falsely reject H_0 in favor of the research hypothesis H_a**

- **Type-II Error:**
  - **Falsely accept H_0**

- $\alpha = P(\text{Type}-\text{I Error})$
  $= P(\text{Falsely reject } H_0 \text{ in favor of } H_a)$

- $\beta = P(\text{Type}-\text{II Error})$
  $= P(\text{Falsely accept } H_0)$

- **We like to have a decision rule that has both α and β very small, but that is not possible.**

TYPES OF ERROR CONT'D

- **We would of course prefer α to be very small since we would not like to conclude in favor of the research hypothesis falsely.**
  - **As a consequence, $\beta = P(\text{Falsely accept } H_0)$ can be very large.**

- $\alpha = P(\text{Falsely reject } H_0)$ **is usually chosen before hand depending upon how much error one is willing to accept.**
  - **Mostly,** $\alpha = 0.10,\ 0.05,\ 0.01,\ 0.001$.
  - **Most frequently used** $\alpha = 0.05$.

- **Having α too small would most likely result in no discovery (Failing to reject H_0).**

HYPOTHESIS TESTING FOR THE MEAN

Marquette University
Be The Difference.

- **Examples:**
  - **A teacher claims her method of teaching will increase test scores by 10 points on average. You randomly sample 25 students to receive her method of teaching and find their test scores. You plan to use the data to refute the claim that the method of teaching she proposes is better.**

Notice the Null Hypothesis ALWAYS has equality associated with it.

Null Hypothesis: $H_0 : \mu \geq 10$ points

Alternative Hypothesis: $H_a : \mu < 10$ points

One-sided alternative hypothesis.

  - **A study involving men with alcoholic blackouts is done to determine if abuse patterns have changed. A previous study reported an average of 15.6 years since a first blackout with a standard deviation of 11.8 years. A second study involving 100 men is conducted, yielding an average of 12.2 years and a standard deviation of 9.2 years. It is claimed that the average number of years has changed between blackouts. Is there evidence to support this claim?**
    - **(Information reported in the *American Journal of Drug and Alcohol Abuse*, 1985, p.298)**

The researchers only wanted to see if the number of years had "changed." They weren't looking for a direction of change.

Null Hypothesis: $H_0 : \mu = 15.6$ years

Alternative Hypothesis: $H_a : \mu \neq 15.6$ years

Two-sided alternative hypothesis.

HYPOTHESIS TESTING FOR THE MEAN CONT'D

- **Test Statistics (TS):**

$$z = \frac{\bar{y} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

- **Decision Rule: Given** $\alpha = P(\text{Type−I error})$

  – $H_a: \mu > \mu_0$ **Reject** H_0 **in favor of** H_a **if** $z > z_\alpha$

  – $H_a: \mu < \mu_0$ **Reject** H_0 **in favor of** H_a **if** $z < -z_\alpha$

  – $H_a: \mu \neq \mu_0$ **Reject** H_0 **in favor of** H_a **if** $|z| > z_{\alpha/2}$

- **Remark: Note that the test statistics is based on CLT (Central Limit Theorem), and** σ **is assumed to be known.**

HYPOTHESIS TESTING FOR THE MEAN CONT'D

- **Assumptions:**
  1. σ is known
  2. $n \geq 30$ or the sample is drawn from a normal population.

- **Example:**
  Let $\{y_1, y_2, \ldots, y_{100}\}$ be sample of blood pressures of 100 patients with a certain disease. We want to investigate that the population of patients with this disease have high blood pressure. Suppose that the mean normal blood pressure is 120. Assume that $\sigma = 5.0$.

- **Sample Information:** $\bar{y} = 121.5$

- **Do we have sufficient evidence to conclude that this population has high blood pressure?** $\alpha = 0.05$.

EXAMPLE CONT'D

- $H_0: \mu = 120$ **(Null Hypothesis)**
- $H_a: \mu > 120$ **(Research Hypothesis)**

- **Test Statistics (TS):** $z = \frac{\bar{y} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{121.5 - 120}{\frac{5.0}{\sqrt{100}}} = 3.0$

- **Decision Rule: Reject H_0 in favor of H_a if $z > z_\alpha$.**

  $\alpha = 0.05$. **Thus** $z_\alpha = 1.64$.

  Reject H_0 in favor of H_a if $z > 1.64$.

- **Conclusion: Is $z > 1.64$? Yes, because $z = 3.0$. Thus we reject H_0 in favor of H_a, and conclude that the population of patients with the disease have high blood pressure.**

BOOK EXAMPLE 5.9

- **Research question: Do anti-bacterial soap work?**

- **An experiment was done on 35 petri dishes which were first cultured with E.Coli bacteria, and then a solution of the antibacterial soap was added. After 24-hour incubation period, bacterial counts were recorded on each of 35 dishes.**

- **Sample Information:** $\bar{y} = 31.2, \quad s = 8.4$

- **If the mean number of bacterial counts under ordinary soap is known to be 33, does the data provide sufficient evidence to conclude that the anti-bacterial soap is effective?**
  Use $\alpha = 0.05$.

BOOK EXAMPLE 5.9 CONT'D

- $H_0: \mu = 33$
- $H_a: \mu < 33$

- **T.S.** $z = \frac{\bar{y} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$
  - **Assumption.** $n \geq 30$, **but the population st.dev.** σ **is unknown. Assume** $\sigma \approx s = 8.4$.

$$z = \frac{\bar{y} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{31.2 - 33}{8.4/\sqrt{35}} = -1.27$$

- **Decision Rule: Reject** H_0 **in favor of** H_a **if** $z < -z_\alpha$.

  $\alpha = 0.05$. **Thus** $z_\alpha = 1.64$.

  Reject H_0 **in favor of** H_a **if** $z < -1.64$.

BOOK EXAMPLE 5.9 CONT'D

- **Conclusion: Is $z < -1.64$? No. So, we cannot reject H_0 in favor of H_a. Can we conclude based on the data that the antibacterial soap is not effective?**
- **NO**

- **Why did we not say that the antibacterial soap is not effective? Which means why we did not accept H_0.**
- **Answer: For accepting H_0, we need to look at**

$$\beta = P(\text{Falsely accept } H_0)$$

- **As we mentioned earlier that β can be very large (in fact as much as $1 - \alpha = 0.95$). So, by accepting H_0, we might be committing a large error.**

COMPUTING $\boldsymbol{\beta}$ (TYPE-II ERROR)

- $\beta =$ **P(Falsely accept H_0) = P(Accept H_0 if H_a is True)**

- **Formula:**
- **One-Tail Test:** $$\beta = \mathrm{P}\left(\mathrm{Z} \leq \mathrm{z}_{\alpha} - \frac{|\mu_0 - \mu_a|}{\sigma/\sqrt{n}}\right)$$
- **Two-Tail Test:** $$\beta = \mathrm{P}\left(\mathrm{Z} \leq \mathrm{z}_{\alpha/2} - \frac{|\mu_0 - \mu_a|}{\sigma/\sqrt{n}}\right)$$

- **Note that, to compute β, you also need μ_a of the True H_a.**

BOOK EXAMPLE 5.10 (CONT. EXAMPLE 5.9)

- **Note that, in example 5.9, we tested**
- $H_0 : \mu = 33 \quad vs. \quad H_a : \mu < 33$

  $n = 35, \;\; \bar{y} = 31.2, \;\; \sigma = 8{,}4, \;\; \alpha = 0.05$

- **Based on this,** $z = \frac{\bar{y} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = -1.27$. **We fail to reject** H_0**.**
- **Should we accept** H_0 **(antibacterial soap is same as the ordinary one)?**
- **If** $\mu_a \leq 28$**, what is the probability of falsely accept** H_0**?**

BOOK EXAMPLE 5.10 CONT'D

- **If $\mu_a \leq 28$, what is the probability of falsely accept H_0?**

$$\beta(\mu_a) = P\left(Z \leq z_\alpha - \frac{|\mu_0 - \mu_a|}{\sigma/\sqrt{n}}\right)$$

- $z_\alpha = 1.64, \mu_0 = 33, \mu_a = 28, \sigma = 8.4, n = 35$

$$\beta(28) = P\left(Z \leq 1.64 - \frac{|33 - 28|}{8.4/\sqrt{35}}\right) = P(Z \leq -1.88) = 0.0301$$

- **Note that if $\mu_a < 28$, then** $\beta(\mu_a) < 0.0301$

- **Since this error is low, it would be reasonable to accept H_0, and declare that the antibacterial soap is not effective.**

ANOTHER EXAMPLE

- **The milk price of a gallon of** 2% **milk is normally distributed with standard deviation of** $\$0.10$**. Last week the milk price was** $\$2.78$**. We want to determine if this week the price is different. Based on sample of** 25**,** $\bar{y} = 2.80.\ \alpha = 0.05.$

- $H_0: \mu = 2.78 \quad vs.$
- $H_a: \mu_a \neq 2.78$

- **TS:** $z = \frac{\bar{y} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{2.80 - 2.78}{\frac{0.10}{\sqrt{25}}} = 1.00$

- **Decision Rule: Reject** H_0 **in favor of** H_a **if** $|z| > z_{\alpha/2} = 1.96$

- **Conclusion: Is** $|z| > 1.96?$ **No. Fail to reject** H_0**.**
  We cannot conclude that milk price has changed.

ANOTHER EXAMPLE CONT'D

- **If $|\mu_a - 2.78| \geq 0.05$, can we conclude that price has not changed significantly?**

$$\beta(\mu_a) = P(Z \leq z_{\alpha/2} - \frac{|\mu_0 - \mu_a|}{\frac{\sigma}{\sqrt{n}}})$$

- $\beta(\mu_a) = P\left(Z \leq 1.96 - \frac{0.05}{\frac{0.10}{\sqrt{25}}}\right) = P(Z \leq -0.54)$

  $= 0.2946$

- **Since this is a large error, we cannot conclude that the price is same as the price last week.**

- **The problem here is that the sample size $n = 25$ is too small.**

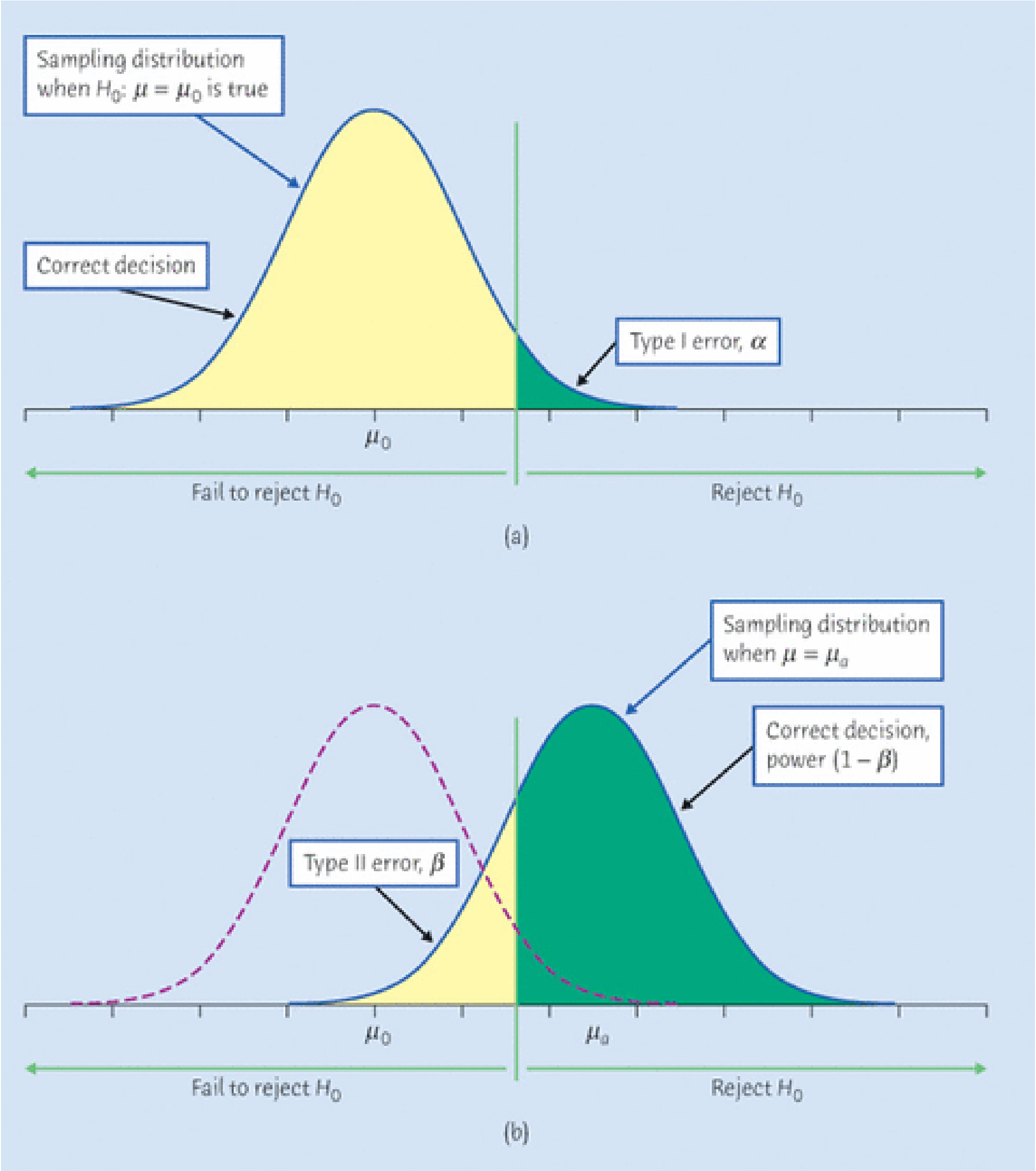
POWER ANALYSIS

- **Power analysis is a tool of statistics, usually researchers employ to make sure that there is enough sample size to make discovery if in fact there is a discovery. In a statistical term,**

- $\text{Power} = P(\text{correct Discovery})$
  $= P(\text{Reject } H_0 \text{ if } H_a \text{ is True})$

- **You plan to conduct an expensive study when you want to prove that a treatment is effective.**
  - H_0: Treatment is not effective
  - H_a: Treatment is effective **(Research Hypothesis)**

- **[Power Analysis Applet]**



POWER ANALYSIS CONT'D

- **What will increase the Power?**
  - **Smaller** σ
  - **Further away** μ_a **from** μ_0
  - **Increasing sample size** n

- **You want to have sufficient sample so that you can correctly discover that the treatment is effective.**
- $H_0: \mu = \mu_0$ $\quad$ $H_a: \mu > \mu_0$
  or $H_a: \mu < \mu_0$
  or $H_a: \mu \neq \mu_0$
- **TS** $\quad z = \frac{\bar{y} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$
- **What is the sufficient sample size to correctly discover** H_a **at a given** α **with power** $P(\mu_a)$.

POWER ANALYSIS CONT'D

- **Formula for finding sample size:**
- **One sided (one-tail) test: ($H_a: \mu < \mu_0$ or $H_a: \mu > \mu_0$)**

$$n = \sigma^2 \frac{\left(z_\alpha + z_\beta\right)^2}{\Delta^2}$$

- **Two sided (two-tail) test: ($H_a: \mu \neq \mu_0$)**

$$n = \sigma^2 \frac{\left(z_{\alpha/2} + z_\beta\right)^2}{\Delta^2}$$

- **where**

$\beta = 1 - \text{Power}$ **and** $\Delta = |\mu_a - \mu_0|$

BOOK EXAMPLE 5.11

- A cereal company sell boxes of cereal with the labeled weight of 16 oz.
- The production is based on the mean weight of 16.37 oz.
- Why?
- Answer: So that only small portion of boxes have weight less than 16 oz.
- They suspect that due to some production defect the weight filled in the boxes have mean less than 16.37

BOOK EXAMPLE 5.11 CONT'D

- **Test the Hypothesis**
- $H_0: \mu = 16.37$ **vs** $H_a: \mu < 16.37$
- **At** $\alpha = 0.05.$ **Assume** $\sigma = 0.225$**.**
- **How many boxes should be sampled in order to correctly discover that mean is less than** 16.37 **with the power of** 0.99 **if in fact the true mean is** $\mu \leq 16.27?$
- **Formula:** $n = \frac{\sigma^2 (z_\alpha + z_\beta)^2}{\Delta^2}$
- $\beta = 1 - \text{Power} = 0.01.$ **Thus** $z_\beta = 2.33.$
- $z_\alpha = 1.64.$ $\Delta = |16.27 - 16.37| = 0.10$
- **Thus** $n = \frac{0.225^2 (1.64 + 2.33)^2}{0.10^2} = 79.99.$ $n = 80.$