Lesson 4.2

Basic Concepts of Hypothesis Testing

Norberto E. Milla, Jr.

What is a hypothesis?

- Scientific or educated guess
- It is a conjecture concerning one or more populations whose veracity can be established using information contained in the random sample
- Examples:
  - Climate change has greatly affected agricultural productivity.
  - COVID-19 vaccinations is significantly associated with deaths in people with comorbidities.
  - Approximately 9.5 percent of students suffer from depression or a depressive illness.
  - A university claims that at least 95% of their civil engineering students pass the board examination.

What is a statistical hypothesis?

- A statistical hypothesis is a claim or statement about the population *parameter*
- *Parameter* is a quantity that describes a characteristic of the units in a population.
- Examples of parameters are population mean (μ), population variance (σ^2), and population proportion (P)
- The parameter must be identified before analysis and it is dictated by the problem

Examples of statistical hypothesis

- The true mean body temperature of COVID-19 patients in COVID-19 Isolation Facilities is above $37.0^o C$.
- The parameter of interest here is μ which is the true mean body temperature of COVID-19 patients

Examples of statistical hypothesis

- The proportion of people in favor of mass vaccination is lower than 0.60.
- The parameter of interest here is P which is the true proportion of people in favor of mass vaccination.

Elements of hypothesis testing

- Null and Alternative Hypotheses
- Test statistic (and its p-value)
- Level of significance
- Decision rule
- Decision and conclusion

Null Hypothesis (H_0)

- This is the statement being tested and might be rejected based on sample evidence
- It is a statistical hypothesis which the experimenter doubts to be true
- Generally, this is a statement of **equality** or **status quo** or **no difference** or **no relationship** or **no effect**.
- Designated by the symbol H_0, read as $H - null$
- Examples:
  - $H_0 : \mu = 37^o C$
  - $H_0 : P = 0.60$

Alternative Hypothesis (H_1)

- It is the statement that must be accepted as true if the null hypothesis cannot be supported with data
- The operational statement of the theory that the experimenter believes to be true and wishes to prove
- It is sometimes referred to as the **researcher's hypothesis**
- Designated by the symbol (H_1), read as $H - 1$
- Uses the inequality symbols $>$, $<$, or $\neq$
- Examples:
  - $H_1 : \mu > 37^o C$
  - $H_1 : P < 0.60$ (Directional alternative)

Null and alternative hypotheses: more examples

- It is believed that fewer than 10% of university students do not graduate on time.
  - $H_0 : P = 0.10$
  - $H_1 : P < 0.10$ (Directional alternative)
- The average time to complete the pen-and-paper admission test is 2.35 hours. The university admissions office believed this time can be significantly changed when using the computer-based admissions test.
  - $H_0 : \mu = 2.35$
  - $H_1 : \mu \neq 2.35$ (Non-directional alternative)

Null and alternative hypotheses: more examples

- The HR Division Chief of a certain government agency wants to support the hypothesis that the true mean performance evaluation scores of the employees will improve this year as compared last year (mean = 80 points) due to several in-house trainings given to employees this year.
  - $H_0 : \mu = 80$ versus $H_1 : \mu > 80$
- A recent college graduate has been offered two similar positions, one in City A and one in City B. He is interested in knowing whether the proportion of a family's income spent on housing in city A is less than the proportion spent on housing in city B.
  - $H_0 : P_A = P_B$ versus $H_1 : P_A < P_B$

What is a test of (statistical) hypothesis?

- A **statistical test** is a method used to decide whether the data at hand sufficiently support a particular hypothesis
- It is a test on how "close" (probabilistically) the sample data is with the value being hypothesized
- The process of testing hypotheses can be compared to court trials
  - A person comes into court charged with a crime
  - A jury must decide whether the person is innocent (null hypothesis) or guilty (alternative hypothesis) based on evidence (testimonies, pictures, videos, etc.)

Types of test of hypothesis

- One-tailed (One-sided)
  - It is used to test a null hypothesis against a directional alternative hypothesis
  - $H_0 : P = 0.10$ vs $H_1 : P < 0.10$
  - $H_0 : \mu > 37^o C$ vs $H_0 : P < 0.60$
- Two-tailed (Two-sided)
  - It is used to test a null hypothesis against a non-directional alternative hypothesis
  - $H_0 : \mu = 2.35$ vs $H_1 : \mu \neq 2.35$

The test statistic

- This is a statistic whose value is calculated from sample data, which will be the basis for deciding whether to reject H_0
- This serves as the evidence that forms the basis in making a decision on H_0, whether to reject it or not
- It can be in the form of the sum, sample average, or any transformations or standardizations of these statistics
- Values of the test statistic define the decision rule or criterion to reject the H_0

Types of errors

- **Type I error**: We reject H_0 when in fact it is true
- **Type II error**: We do not reject H_0 when in fact it is not true

		True Situation	
		The null hypothesis is true.	The null hypothesis is false.
Decision	We reject the null hypothesis.	**TYPE I error *(rejecting a true null hypothesis)***	**CORRECT decision**
	We do not reject the null hypothesis.	**CORRECT decision**	**TYPE II error *(not rejecting a false null hypothesis)***

Types of errors: an example

Red Tide is a bloom of poison-producing algae- a few different species of a class of plankton called dinoflagellates. When the weather and water conditions cause these blooms, shellfish such as clams living in the area develop dangerous levels of paralysis-inducing toxin. The Bureau of Fish and Aquatic Resources (BFAR) monitors levels of the toxin in shellfish by regular sampling of shellfish along the coastline. If the mean level of toxin in clams exceeds 800 micrograms of toxin per kg of clam meat in any are, clam harvesting is banned there until the bloom is over and the levels of toxin in clams subside.

- $H_0 : \mu = 800$ versus $H_1 : \mu > 800$
- When does a Type I error occur? Type II?
- Which type has fatal consequence?

Level of significance

- The probability of committing a Type I error is the **level of significance**, denoted as α
- It is usually set to a small value such as 0.05, 0.01, or 0.1
- $\alpha = 0.05$ indicates a 5% risk of concluding that a difference (effect) exists when there is no actual difference (effect)
- Lower significance levels indicate that you require stronger evidence before you will reject the null hypothesis
- The researcher determines the significance level before conducting the study

The p-value

- The **p-value** is the probability that, if the null hypothesis were true, the test statistic will have a value as extreme or more extreme than the results obtained in the (current) sample
- It is the probability that you would obtain the effect observed in your sample, or larger, if the null hypothesis were true
- It is calculated based on sample data and under the assumption that the null hypothesis is true
- Lower p-values indicate greater evidence against the null hypothesis

The p-value

Based on a survey conducted three years ago, the mean age of graduate students at a university is 31 years with a standard deviation of 2 years. Recently, a random sample of 15 graduate students is taken and the sample mean is 32 years with sample standard deviation of three years.

- $H_0 : \mu = 31$ vs $H_1 : \mu > 31$
- Assuming normal distribution,

$$\begin{aligned} p - value &= P(\overline{Y} > 32 | \mu = 31) \\ &\approx 0.0264 \end{aligned}$$

- The above probability was obtained using the R command **pnorm(32,31,2/sqrt(15),lower.tail=F)**
- This means that if the null hypothesis were true, that is $\mu = 31$, there is a only a 2.64% chance that we observe a sample mean more extreme than 32.

The Decision Rule

- We reject the null hypothesis if the sample data provides sufficient evidence against it
- We do not reject the null hypothesis if the sample data does not provide sufficient evidence against it
- **Rejection Region** approach: Reject H_0 if the value of the test statistic falls in the rejection region
- **p-value** approach: Reject H_0 if the p-value is less than the significance level. That is, reject H_0 if $p - value \leq \alpha$

Steps in hypothesis testing

1. State the null and alternative hypotheses
2. Specify the level of significance (α)
3. Collect data and compute the test statistic and its p-value. (Can use JASP, Excel or R)
4. Formulate a decision rule either using the critical value from statistical tables or using the p-value approach
5. Make a decision.
6. Draw conclusion.