

# Hypothesis Testing

Project 1

**Hypothesis Testing** is a way  
to test a theory or a question.

# Example data questions

1. Does this new medication work?
2. Is the cost of living higher in this city?
3. Does this group score higher than another group?

A large part of Statistics and the Scientific Process is to **identify a Hypothesis** and then try **to determine if the observations or effects could be due to chance.**

# Key Points about Hypothesis Testing

1. Hypothesis Testing is applying statistical methods to determine if something happened purely by chance.
2. Scientists and researchers often form a Hypothesis for their observations along with a Null Hypothesis. The Null Hypothesis assumes that observations or effects are simply by chance.
3. **The goal of hypothesis testing is to reject the Null Hypothesis** through statistical tests.
4. **Rejecting the Null Hypothesis presumes that the Hypothesis was true.**
5. A Hypothesis is a **measurable** and **testable statement** about something that you expect will happen.

# Hypothesis as If/Then

Hypothesis is often expressed as If/Then

"If San Diego, CA is warmer than Austin, TX in July, then the average daily temperature will be higher."

# Null Hypothesis

The Null Hypothesis is typically stated that NO differences exist between the variables or groups of interest.

"If San Diego, CA is not warmer than Austin, TX in July, then there will be no difference in the average daily temperatures."

# p-value

Is the probability of rejecting the null hypothesis when it is true (false positive).

For example, a significance level of 0.05 indicates a 5% risk of concluding that a difference exists when there is no actual difference.



# Rejecting the Null Hypothesis is never absolute

Instead, statisticians calculate the **probability of observing the event. This is called the P value.**

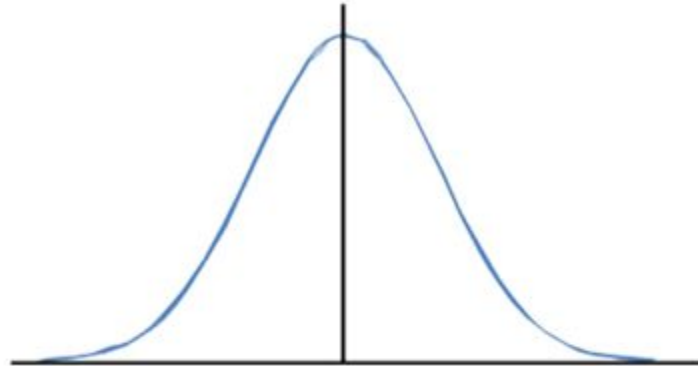
The P value is then compared to a fixed significance level to determine if the Null Hypothesis can be rejected. A smaller P value indicates stronger evidence against the Null Hypothesis.

# Steps for Hypothesis Testing

1. Determine the Hypothesis and Null Hypothesis.
2. Identify the appropriate statistical test.
3. Determine the acceptable significance value.
4. Compute the P-value.
5. Determine if the P-value rejects the Null Hypothesis by comparing it to the significance value (Typically  $< 0.05$ ).

# Central Limit Theorem

The Central Limit Theorem states that the sampling distribution of the sample means approaches a normal distribution as the sample size gets larger — no matter what the shape of the population distribution. This fact holds especially true for sample sizes over 30. All this is saying is that as you take more samples, especially large ones, your graph of the sample means will look more like a normal distribution.



**Figure 1** Normal distribution curve.

# Assumption of Equal Variance

What Is the Assumption of Equal Variance? ... Statistical tests, such as analysis of variance (ANOVA), assume that although different samples can come from populations with different means, they have the same variance. Equal variance is when the variances are approximately the same across the samples.

# ANOVA (Analysis of Variance)

- Analysis of Variance (ANOVA) is a test to compare the means of multiple groups.
- ANOVA assumes the Null Hypothesis that there is no difference between groups.
- Any mean that is significantly different from the rest will result in a low p-value.

**ANOVA Example: Does each drink below affect reaction time differently?**



$H_0$ : The mean for all three groups is the same.

So... which test is right for our project?

# T-Test

Allows you to answer the question...

- Are these 2 groups statistically different from each other?
- **Null Hypothesis:** There is no statistical difference between the means of the 2 groups.
- When you reject the null hypothesis with a t-test, you are saying that the means are statistically different.
- The difference is meaningful!



# ANOVA (Analysis of Variance)

Is the equivalent of running multiple t-tests between more than 2 groups.

- Will need to run additional analysis to determine which individual level comparisons are significant.

# Chi-Square

Allows you to test whether there is a relationship between two variables.

BUT, it does not tell you the direction or the size of the relationship.

- **Null Hypothesis:** There is no relationship between the two variables.
- When you reject the null hypothesis, you are saying that there is a relationship between the 2 variables.



**Hey girl**

**You must be  $p > .05$ ,  
because I fail to reject you.**

quickmeme.com