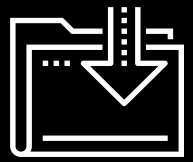
Hypothesis Testing and Statistical Tests

Data Boot Camp

Lesson 7.3



Class Objectives

By the end of today's class, you will be able to:

- Explain the difference between a null hypothesis and an alternate hypothesis.
- Apply a one-sample t-test to determine if sample and population data are significantly different.
- Apply a two-sample t-test to determine if two groups are significantly different.
- Apply ANOVA to compare the means of three or more groups.
- Perform a chi-square test to compare the distribution of categorical data.

Intro to Hypothesis Testing

Hypothesis testing can be confusing at times, mostly because you must create your null and alternative hypotheses before performing any analysis.

Hypothesis Testing

A hypothesis statement is an educated guess about something.

The hypothesis is often expressed as an `if/then` statement.

01

Hypothesis testing in statistics is a way to test the results of a survey or an experiment to determine if they are meaningful.

02

We test for two mutually exclusive outcomes:

- The null hypothesis
- The alternative hypothesis

Hypothesis Testing

Null and alternative hypotheses:

Null hypothesis (H_0)

The null hypothesis is the hypothesis that we are trying to disprove; it states that no statistical significance exists between the two variables.

Your null hypothesis assumes that your results happened by chance.

Alternative Hypothesis (H_a)

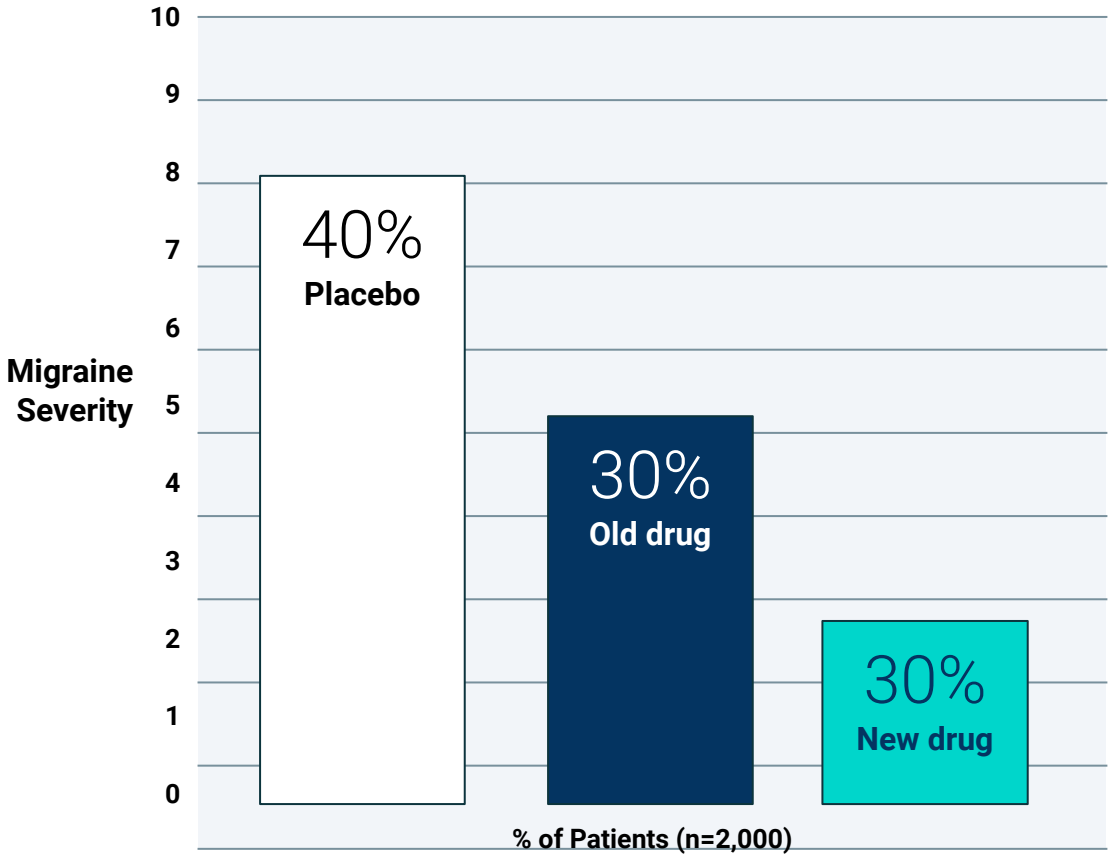
The alternative hypothesis is the opposite of the null hypothesis; it assumes that some factor influenced the results—meaning that they did not happen by chance.

Moving forward, we will refer to the alternative hypothesis as just the hypothesis.

Null and Alternative Hypotheses: Example

Null Hypothesis:
Migraine severity when taking the new drug is the same when taking the old drug.

Alternative Hypothesis:
Migraine severity when taking the new drug is statistically significantly better or worse than when taking the old drug.



Hypothesis Testing

Steps for hypothesis testing:

01 Determine the hypothesis and null hypothesis.

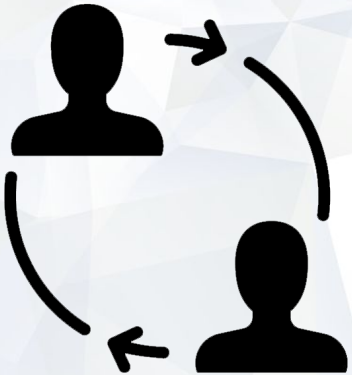
02 Identify the appropriate statistical test.

03 Determine the acceptable significance value.

04 Compute the p-value.

05 Determine if the p-value rejects the null hypothesis by comparing it to the significance value (typically, $p < 0.05$).

In an analysis, we test the hypothesis and decide whether we can reject the null hypothesis. If we can, the only option left is the alternative hypothesis!



Partner Activity: Forming a Null Hypothesis

In this activity, you and your partner will turn two provided questions into a hypothesis and a null hypothesis.

Suggested Time:

10 minutes

Partner Activity: Forming a Null Hypothesis

Instructions

Convert the following questions into a hypothesis and a null hypothesis.

- Does dark chocolate affect arterial function in healthy individuals?
- Does coffee have anti-aging properties?

Time’s Up! Let’s Review.

Does dark chocolate affect the arterial function in healthy individuals?

Hypothesis: If dark chocolate is related to arterial function in healthy individuals, consuming 30 grams (g) of dark chocolate daily for one year will result in improved arterial function.

Null Hypothesis

Consuming 30 g of dark chocolate daily for one year will result in no improvement in arterial function.

Alternative Hypothesis

Consuming 30 g of dark chocolate daily for one year will result in improvement in arterial function.

Does coffee have anti-aging properties?

Hypothesis: If coffee consumption is related to anti-aging properties, consuming 400 milligrams (mg) of coffee daily will reduce mortality from age-related diseases, such as heart disease.

Null Hypothesis

Consuming 400 mg of coffee daily will result in an equal or greater rate of heart disease.

Alternative Hypothesis

Consuming 400 mg of coffee daily will result in a reduction in the rate of heart disease.

T-Tests

A **t-test** tells you how significant the differences between groups are.

It lets you know if the differences, measured in means (averages), could have happened by chance.

Calculating a T-Test

$\overline{X_1}$	Mean of the first dataset
$\overline{X_2}$	Mean of the second dataset
s_1^2	Standard deviation of the first dataset
s_2^2	Standard deviation of the second dataset
n_1	Number of elements in the first dataset
n_2	Number of elements in the second dataset

$$t = \frac{\left(\overline{X_1} - \overline{X_2}\right)}{\sqrt{\frac{s_1^2}{n_1} - \frac{s_2^2}{n_2}}}$$

T-Tests

One-sample t-test

Determines whether the sample mean statistically differs from a known or hypothesized population mean.

Independent t-test

Determines whether there is a statistically significant difference between the means in two unrelated groups.

An independent t-test is also known as a two-sample t-test.

What type of t-test should I use?

T-Tests

There are couple of things to consider before performing t-test:

01 Do the compared groups come from a single population or two distinct populations?

02 Do you want to test the difference in a specific direction?

T-Tests

One-sample t-test

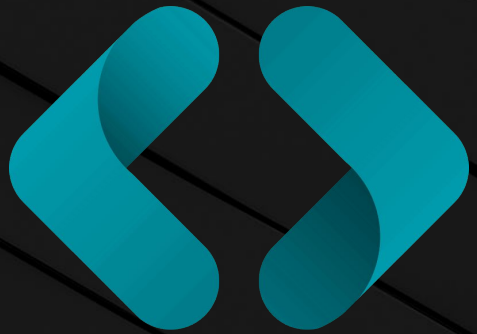
One group being compared against a standard value

Example: Comparing gasoline octane level to an octane level

Independent t-test

Groups coming from two distinct populations

Example: Different countries, different species



Time to <code>

Activity: T-Test

In this activity, you will use a t-test to compare the difference in heights of non-related individuals from two different locations.

Suggested Time:

15 minutes

Activity: T-Test

Instructions	
	Calculate the mean for each population.
	Use a t-test to determine if there is a statistically significant difference in the height of non-related individuals in San Francisco vs. Denver.
	It is up to you to determine if you should use a one-sample or an independent t-test.

Time’s Up! Let’s Review.

ANOVA

ANOVA

01

ANOVA is an extension of a t-test. With a t-test, you can test two groups to determine whether their means are statistically different.

02

An ANOVA test is a way to find out if survey or experiment results are significant.

03

It helps you figure out whether to reject the null hypothesis or to fail to reject it.

Activity: ANOVA

In this activity, you will use ANOVA to compare the differences in pain tolerance based on how many times a person exercises in a given week.

Suggested Time:

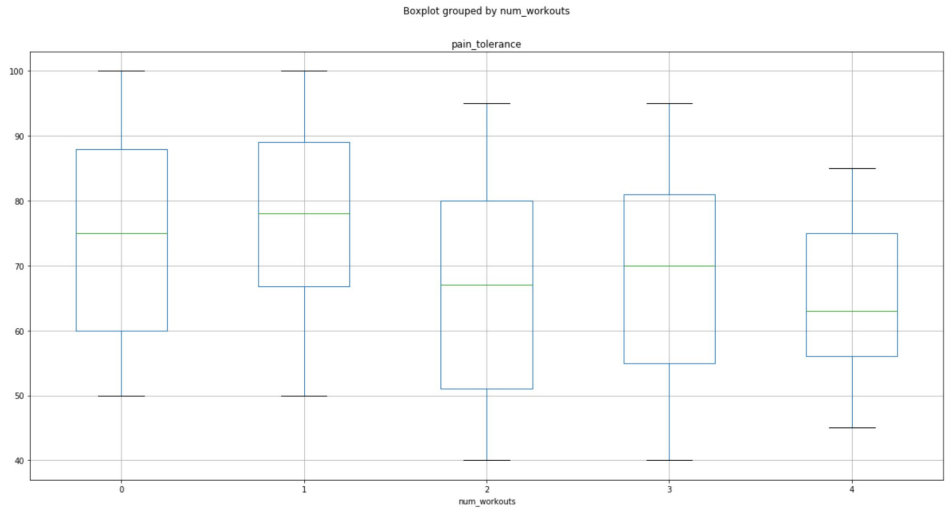
10 minutes

Activity: ANOVA

Instructions

Perform a one-way ANOVA test to determine whether there are any significant differences in number of workouts vs. pain threshold.

Create a box plot to show the distribution of pain tolerances for number of workouts.



Time’s Up! Let’s Review.

Chi Square

The Chi-Square Test

What is it used for?

- To answer the question: Is the distribution of frequencies in the dataset meaningful?
- In other words, does the data match our expectations?
- Stated another way: Do we reject the null hypothesis or fail to reject it?

The Chi-Square Test

Example: Out of 300 dinosaurs…

220 eat everything

55 eat only meat

25 eat only plants

Null hypothesis:
No statistical significance exists in the distribution of omnivores, carnivores, and herbivores. That is, this data can be explained by random distribution.

The chi-square test can help us reject or fail to reject the null hypothesis.

The Chi-Square Python Function

How is the chi-square test used?

01 Determine chi-square value.

02 Determine degree of freedom.

03 Choose a p-value.

04 Determine critical value.

05 Make a decision.

Degrees of Freedom

To determine the degree of freedom (df),
take the number of rows and subtract 1:

Omnivores	220
Carnivores	55
Herbivores	25

There are three rows, so the degree of freedom is 3 – 1 = 2

The degree of freedom is the number of figures required to fill out the table (like Sudoku).

If we have two of the numbers, we can figure out the value of the third.

P-value

The p-value is the **confidence level**, or the acceptable risk of a false positive.

p = 0.05 is widely accepted in academia but may be higher in business settings.

Importance of Findings	Significance Level	Probability of Being Wrong
Low	0.1	1 in 10
Normal	**0.05**	**5 in 100**
High	0.01	1 in 100
Very high	0.001	1 in 1,000
Extreme	0.0001	1 in 10,000

We'll say 0.05 in this example.

The Chi-Square Test Formula

A few more considerations:

- The chi-square test is used to test categorical variables; it can’t be used on continuous data.
- The categories must be mutually exclusive.
- We covered using the chi-square test formula to test goodness of fit.
- It can also be used to test independence. Feel free to explore this more on your own.

Using the Chi-Square Test In Python

Import the scipy.stats module.

```
# The statistical module used to run chi square test
import scipy.stats as stats
```

Determine the critical value.

```
# The degree of freedom is 3-1 = 2
# With a p-value of 0.05, the confidence level is 1.00-0.05 = 0.95.
critical_value = stats.chi2.ppf(q = 0.95, df = 2)
```

```
# The critical value
critical_value
```

```
5.99146454710798
```

Run the chi-square test.

```
# Run the chi square test with stats.chisquare()
stats.chisquare(df['observed'], df['expected'])
```

```
Power_divergenceResult(statistic=220.5, pvalue=1.3153258948574585e-48)
```

Activity: Chi Square

In this activity, you will perform the chi-square test, first in Python and then by hand.

Suggested Time:

10 minutes

Activity: Chi Square

Instructions	
	You are the owner of four cafés in a town of avid coffee drinkers.
	Using a chi-square goodness-of-fit test, determine whether the results suggest that customers are more likely to frequent one café over another.
	Perform the necessary calculations using Python.
	Then, perform the calculations by hand to verify your findings.
	Use a p-value of 0.05.
	Consult a chi-square table to find your critical value: https://www.medcalc.org/manual/chi-square-table.php
	On your student repository, open Stu-Cafes.ipynb.

Time’s Up! Let’s Review.

Countdown timer

40:00

(with alarm)

Questions?