Introduction to UX Research

Lesson 11: Statistical Analysis

The Erdős Institute

Basic Hypothesis Testing

Hypothesis testing involves investigating how an independent variable impacts a dependent variable

- Does store location affect sales rates?
- Does price affect brand perception?
- Does customer age affect loyalty behaviors?

Basic Hypothesis Testing

We need to formulate our theory into two versions of reality that are mutually exclusive.

H_0 = “null hypothesis”

H_a = “alternative hypothesis”

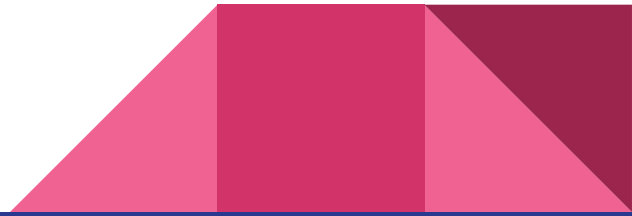
Basic Hypothesis Testing

Hypothesis

You are scouting a location for a new coffee shop. You want to test the idea that coffee shops within 1 mile of campus have higher Saturday morning sales than shops more than 1 mile from campus.

H_0: sales for shops <= 1 mile are ***the same as*** sales for shops > 1 mile

H_a: sales for shops <= 1 mile are ***greater than*** sales for shops > 1 mile



Basic Hypothesis Testing

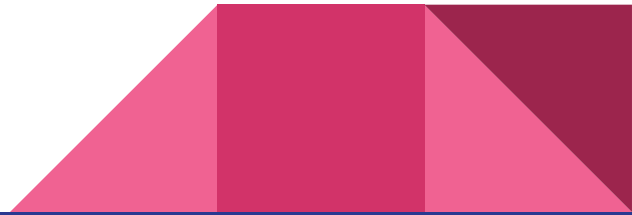
To test our hypothesis, we want to *disconfirm the null*.

H_0 = "null hypothesis"

H_a = "alternative hypothesis"

We always start by assuming the null hypothesis is true.

- If we **do not have enough evidence** to disconfirm the null, we say the null adequately describes reality.
- If we **do have enough evidence** to disconfirm the null, we must accept the alternative.



Basic Hypothesis Testing

Hypothesis

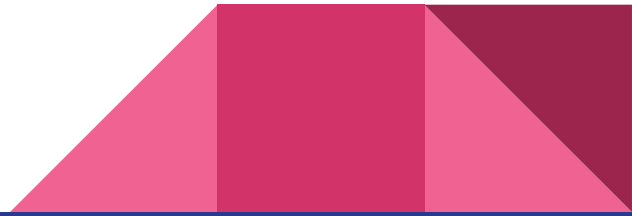
Store A = 1 mile from campus

Store B = 2 miles from campus

H_0: Shop A sales = Shop B sales

H_a: Shop A sales > Shop B sales

We assume that reality matches the null – there's no difference in sales based on shop distance from campus. We need enough evidence to convince us that the null hypothesis is not true.



Basic Hypothesis Testing

We collect data on Store A and B's Saturday morning coffee sales for one year.

Store A – 1 mile from campus

Week	Sales
1	514.9
2	515.0
3	514.5
4	513.8
5	513.1
6	514.2
7	517.0
8	517.3
9	514.8
10	513.5

Mean: 514.83
SD: 1.88

Store B – 2 miles from campus

Week	Sales
1	508.7
2	512.9
3	509.7
4	512.2
5	509.5
6	509.1
7	504.7
8	507.2
9	511.2
10	511.5

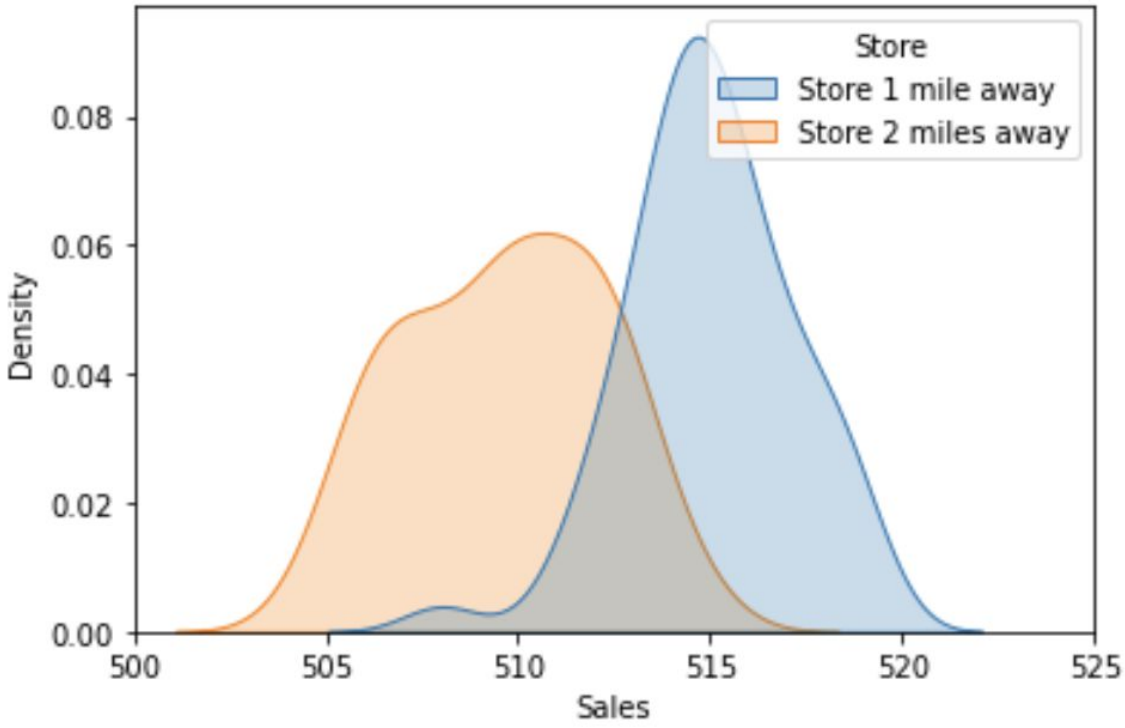
Mean: 510.02
SD: 2.71

Basic Hypothesis Testing

Now, we need to look at the consistency between the sales result from our data and the sales data under the null hypothesis.

In other words, do we have enough evidence to disprove the null hypothesis – that the sales from Store A and Store B are the same?

Basic Hypothesis Testing



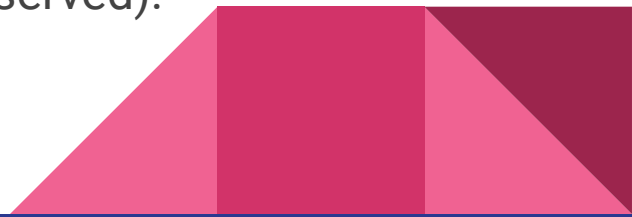
Basic Hypothesis Testing

P values

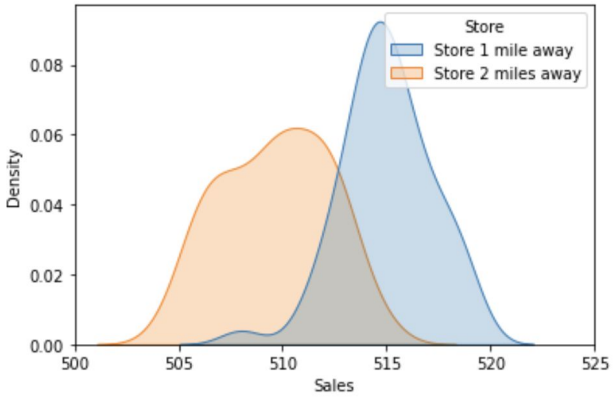
- Assumes the null hypothesis is true
- The probability that the obtained result, or a result even more discrepant from the null hypothesis than the obtained result

Typically, we reject the null hypothesis if the p-value is < 0.05.

This means that, if the null hypothesis is true, there is a 5% chance of obtaining your result (or a value more extreme than the one you observed).



Basic Hypothesis Testing



T-test

- A statistical test to compare the means of two groups
- A p-value will be associated with the t-test

For our data

A (one-tailed) t-test comparing the Saturday coffee sales of Store A and Store B

- t = 11.403
- *p* = 3.26 e-20 – we can reject the null hypothesis!

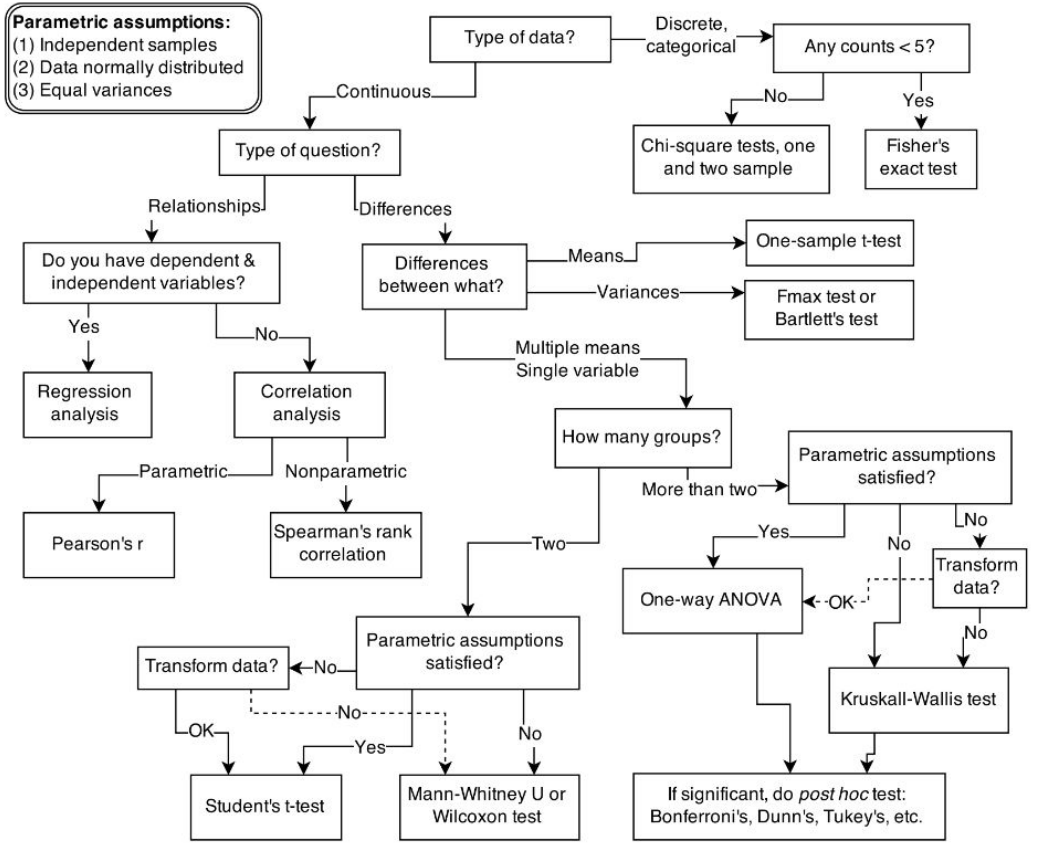
On average, a store 1 mile from campus sold more coffee (mean = $514, SD = 1.88) than a store 2 miles from campus (mean = $510, SD = 2.71), t(51) = 11.403, *p < 0.001* (one-tailed).

Basic Hypothesis Testing

Type of dependent variable	Type of independent variable						
	Ordinal/categorical				Normal/interval (ordinal)	More than 1	None
	Two groups		More groups				
	Paired	Unpaired	Paired	Unpaired			
2 categories	McNemar Test, Sign-Test	Fisher Test, Chi-squared-Test	Cochran's Q-Test	Fisher Test, Chi-squared-test	(Conditional) Logistic Regression	Logistic Regression	Chi-squared-Test
Nominal	Bowker Test	Fisher Test, Chi-squared-Test		Fisher Test, Chi-squared-test	Multinomial logistic regression	Multinomial logistic regression	Binomial Test
Ordinal	Wilcoxon Test, Sign-Test	Wilcoxon-Mann-Whitney Test	Friedman-Test	Kruskal-Wallis Test	Spearman-rank-test	Ordered logit	Median Test
Interval	Wilcoxon Test, Sign-Test	Wilcoxon-Mann-Whitney Test	Friedman-Test	Kruskal-Wallis Test	Spearman-rank test	Multivariate linear model	Median Test
Normal	t-Test (for paired)	t-Test (for unpaired)	Linear Model (ANOVA)	Linear Model (ANOVA)	Pearson-Correlation-test	Multivariate Linear Model	t-Test
Censored Interval	Log-Rank Test		Survival Analysis, Cox proportional hazards regression				
None	Clustering, factor analysis, PCA, canonical correlation						

https://www.r-bloggers.com/2019/03/overview-of-statistical-tests/

Basic Hypothesis Testing



https://xcelab.net/rm/statistical-rethinking