



onrad

School of Entrepreneurship and Business

Agenda

01

Recap

02

Comparing
Two
Populations

03

Hands-on
Data
Analysis



Hypothesis Testing

- The general goal of a hypothesis test is to rule out chance (sampling error) as a plausible explanation for the results from a research study.
- Is also called *significance testing*
- Tests a claim about a parameter using evidence (data in a sample)

Key Terms

- A. Null and alternative hypotheses
- B. Test statistic
- C. P-value and interpretation

P-Value

- P -value answer the question: What is the probability of the observed test statistic ... **when H_0 is true?**
- Thus, smaller and smaller P -values provide stronger and stronger evidence against H_0
- Small P -value \Rightarrow strong evidence

Conventions*

$P > 0.10 \Rightarrow$ non-significant evidence against H_0

$0.05 < P \leq 0.10 \Rightarrow$ marginally significant evidence

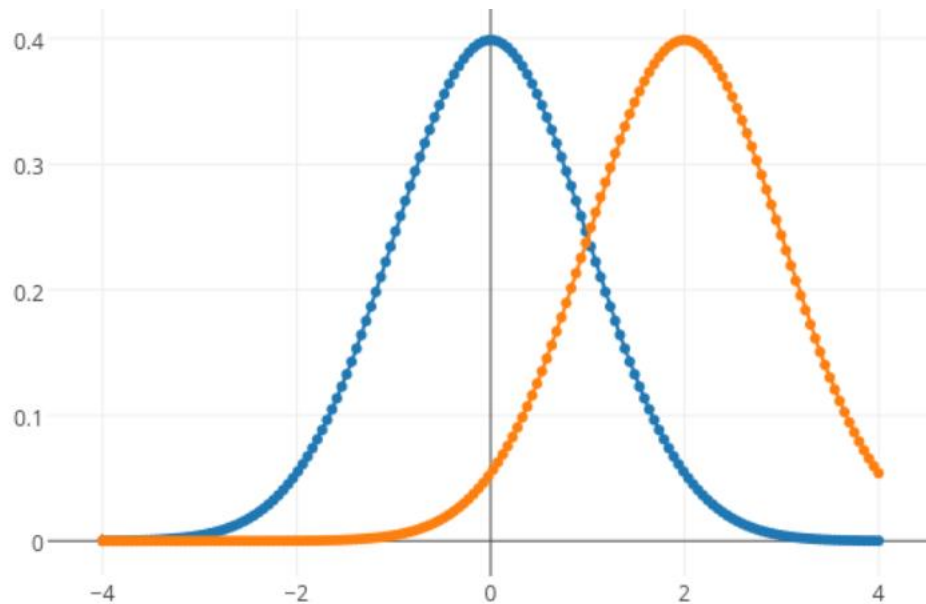
$0.01 < P \leq 0.05 \Rightarrow$ significant evidence against H_0

$P \leq 0.01 \Rightarrow$ highly significant evidence against H_0

Comparing Two Populations



Comparing Means of Two Populations



Inferences About the Difference Between Two Population Means: σ_1 and σ_2 Known

Inferences About the Difference Between Two Population Means: σ_1 and σ_2 Unknown

Estimating the Difference Between Two Population Means

Let μ_1 equal the mean of population 1 and μ_2 equal the mean of population 2.

The difference between the two-population means is $\mu_1 - \mu_2$.

To estimate $\mu_1 - \mu_2$, we will select a simple random sample of size n_1 from population 1 and a simple random sample of size n_2 from population 2.

Let \bar{x}_1 equal the mean of sample 1 and \bar{x}_2 equal the mean of sample 2.

The point estimator of the difference between the means of the populations 1 and 2 is $\bar{x}_1 - \bar{x}_2$

Sampling Distribution of $\bar{x}_1 - \bar{x}_2$

Expected Value

$$E(\bar{x}_1 - \bar{x}_2) = \mu_1 - \mu_2$$

Standard Error

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

where: σ_1 = standard deviation of population 1
 σ_2 = standard deviation of population 2
 n_1 = sample size from population 1
 n_2 = sample size from population 2

STAY CLOSE
SCREEN

Interval Estimation of $\mu_1 - \mu_2$: σ_1 and σ_2 Known

Interval Estimate

$$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

where:

$1 - \alpha$ is the confidence coefficient

So, if we need a 95% CI, then $\alpha = 0.05$

And for 99% CI, $\alpha = 0.01$

Hypothesis Tests About $\mu_1 - \mu_2$: σ_1 and σ_2 Known

Hypotheses

$H_0: \mu_1 - \mu_2 \geq D_0$ $H_a: \mu_1 - \mu_2 < D_0$	$H_0: \mu_1 - \mu_2 \leq D_0$ $H_a: \mu_1 - \mu_2 > D_0$	$H_0: \mu_1 - \mu_2 = D_0$ $H_a: \mu_1 - \mu_2 \neq D_0$
-------------------------------------------------------------	-------------------------------------------------------------	-------------------------------------------------------------

Left-tailed

Right-tailed

Two-tailed

Test Statistic

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

PSD1

PSD2

Example



Titleist is a manufacturer of golf equipment and has developed a new golf ball that has been designed to provide “extra distance.”

In a test of driving distance using a mechanical driving device, a sample of Titleist balls was compared with a sample of golf balls made by Callaway, a competitor. The sample statistics are below:



	Sample #1 <u>Titleist</u>	Sample #2 <u>Callaway</u>
Sample Size	120 balls	80 balls
Sample Mean	275 yards	258 yards

Based on data from previous driving distance tests, the two population standard deviations are known with $\sigma_1 = 15$ yards and $\sigma_2 = 20$ yards.

Hypothesis Tests About $\mu_1 - \mu_2$: σ_1 and σ_2 Known

■ p -Value Approach

Develop the hypotheses.

$$H_0: \mu_1 - \mu_2 \leq 0$$

$$H_a: \mu_1 - \mu_2 > 0$$

Right-tailed test

where:

μ_1 = mean distance for the population of Titleist golf balls

μ_2 = mean distance for the population of Callaway golf balls



Compute the value of the test statistic.

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$z = \frac{(235 - 218) - 0}{\sqrt{\frac{(15)^2}{120} + \frac{(20)^2}{80}}} = \frac{17}{2.62} = 6.49$$



Hypothesis Tests About $\mu_1 - \mu_2$: σ_1 and σ_2 Known

p -Value Approach

Compute the p -value.

For $z = 6.49$
the p -value $< .0001$.

Determine whether to reject H_0 .

Conventions*

- $P > 0.10 \Rightarrow$ non-significant evidence against H_0
- $0.05 < P \leq 0.10 \Rightarrow$ marginally significant evidence
- $0.01 < P \leq 0.05 \Rightarrow$ significant evidence against H_0
- $P \leq 0.01 \Rightarrow$ highly significant evidence against H_0

Because p -value $\leq .01$, we reject H_0 .

Social Science Statistics

$$H_0: \mu_1 - \mu_2 \leq 0$$

$$H_a: \mu_1 - \mu_2 > 0$$

Right-tailed test

Titleist
#1 ball in golf.®



Test Statistic When Population Standard Deviations are not Known

Interval Estimation of $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown

When σ_1 and σ_2 are unknown, we will:

- use the sample standard deviations s_1 and s_2 as estimates of σ_1 and σ_2 and replace $z_{\alpha/2}$ with $t_{\alpha/2}$.

Interval Estimate

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Where the degrees of freedom for $t_{\alpha/2}$ are:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1} \left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1} \left(\frac{s_2^2}{n_2}\right)^2}$$

Difference Between Two Population Means:

σ_1 and σ_2 Unknown



Example: Amazon

Are female customers more likely to recommend Amazon to their friends and family than male customers?

2 Brand X Customer Strategy Quant Survey Instrument – Survey #1

S1	Are you...	GO TO
	Male	S2
	Female	

Handwritten red annotations: a checkmark and a squiggle.

Likelihood Recommend						
Q13 How likely are you to recommend the following retailer to friends and family members, on a scale from 0-10, where 0=Not at all likely to recommend, and 10= Will definitely recommend.						
PROGRAMMER: ALLOW ONLY ONE ANSWER PER BRAND. SHOW EACH RETAILER ON SEPARATE SCREENS.						
	Brand X	JC Penney Target	Kohl's	Nordstrom	Amazon	TJ Maxx
0	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Hypothesis Tests About $\mu_{\text{Females (1)}} - \mu_{\text{Males (2)}}$ $\sigma_{\text{Females (1)}}$ and $\sigma_{\text{Males (2)}}$ Unknown

Hypotheses

$$H_0: \mu_1 - \mu_2 \geq D_0$$

$$H_a: \mu_1 - \mu_2 < D_0$$

Left-tailed

$$H_0: \mu_1 - \mu_2 \leq D_0$$

$$H_a: \mu_1 - \mu_2 > D_0$$

Right-tailed

$$H_0: \mu_1 - \mu_2 = D_0$$

$$H_a: \mu_1 - \mu_2 \neq D_0$$

Two-tailed

Test Statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Which of the above are we using?

Example: Amazon

Are female customers more likely to recommend Amazon to their friends and family than male customers?

SPSS®

	Name	Scale	Measure	Decimals	Label	Values	Missing	Columns	Align
1	respid	N	0	0	respid	None	None	8	Right
2	status	S	0	0	status	None	None	20	Left
3	interview_start	D	0	0		None	None	8	Right
4	interview_end	D	0	0		None	None	8	Right
5	htotaltimein...	N	0	0		None	None	8	Right
6	userid	S	0	0		None	None	50	Left
7	subsid	S	0	0	SubsID	None	None	50	Left
8	s1	N	0	0	Are you...	{1, Male}...	None	8	Right
9	os2	S	0	0	Please enter yo...	None	None	5	Left
10	dregion	N	0	0	dRegion	{1, NORTH...	None	8	Right
11	dregion1	N	0	0	dRegion1	{1, Northeas...	None	8	Right
12	s3	N	0	0	Please enter yo...	None	None	8	Right
13	hage	N	0	0	PLEASE CODE...	{1, 1-17}...	None	8	Right
14	s4_1	N	0	0	No, not of Hisp...	None	None	8	Right
15	s4_2	N	0	0	Yes, Mexican, ...	None	None	8	Right
16	s4_3	N	0	0	Yes, Puerto Ric...	None	None	8	Right
17	s4_4	N	0	0	Yes, Cuban (Ar...	None	None	8	Right
18	s4_5	Numeric	1	0	Yes, Central / S...	None	None	8	Right
19	s4_6	Numeric	1	0	Yes, other Hisp...	None	None	8	Right
20	s5_1	Numeric	1	0	White / Caucasi...	None	None	8	Right
21	s5_2	Numeric	1	0	Black / African ...	None	None	8	Right

- Power Analysis >
- Reports >
- Descriptive Statistics >
- Bayesian Statistics >
- Tables >
- Compare Means >
- General Linear Model >
- Generalized Linear Models >
- Mixed Models >
- Correlate >
- Regression >
- Loglinear >
- Neural Networks >
- Classify >
- Dimension Reduction >
- Scale >
- Nonparametric Tests >
- Forecasting >
- Survival >
- Multiple Response >
- Missing Value Analysis... >
- Multiple Imputation >
- Complex Samples >
- Simulation... >
- Quality Control >
- Spatial and Temporal Modeling... >
- Direct Marketing >

- Means...
- One-Sample T Test...
- Independent-Samples T Test...
- Summary Independent-Samples T Test
- Paired-Samples T Test...
- One-Way ANOVA...

357614 SPSS.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

	Type	Width	Decimals	Label	Values
1	Numeric	8	0	respid	None
2	String	8	0	status	None
3	start Date	10	0	interview_start	None
4	end Date	10	0		None
5	in... Numeric	17	0		None
6	String	1	0		None
7	String	18	0		None
8	Numeric	1	0		{1, Male}...
9	String	5	0		None
10	Numeric	1	0		{1, NORTH...
11	Numeric	1	0		{1, Northeas...
12	Numeric	2	0		None
13	Numeric	2	0	...NG AGE BREAKS:	{1, 1-17}...
14	Numeric	1	0	No, not of Hispanic, Latino or Spanish origin (Are you of Hispanic, Latin...	None
15	Numeric	1	0	Yes, Mexican, Mexican American or Chicano (Are you of Hispanic, Latin...	None
16	Numeric	1	0	Yes, Puerto Rican (Are you of Hispanic, Latino or Spanish origin?)	None
17	Numeric	1	0	Yes, Cuban (Are you of Hispanic, Latino or Spanish origin?)	None
18	Numeric	1	0	Yes, Central / South American (Are you of Hispanic, Latino or Spanish ...	None
19	Numeric	1	0	Yes, other Hispanic, Latino or Spanish origin (Are you of Hispanic, Lati...	None
20	Numeric	1	0	White / Caucasian (Which of the following describes your racial or eth...	None
21	Numeric	1	0	Black / African American (Which of the following describes your racial o...	None

Independent-Samples T Test

Test Variable(s): Q13_1 (How likely are you to ...)

Grouping Variable: s1(2 1)

Options... Bootstrap...

OK Paste Reset Cancel Help

Independent-Samples T Test: Options

Confidence Interval Percentage: 95 %

Missing Values

Exclude cases analysis by analysis

Exclude cases listwise

Continue Cancel Help

SPSS Output

T-Test

Group Statistics

Are you...	N	Mean	Std. Deviation	Std. Error Mean
Female	1140	8.76	1.676	.050
Male	468	8.74	1.457	.067



$$H_0: \mu_1 - \mu_2 \leq D_0$$

$$H_a: \mu_1 - \mu_2 > D_0$$

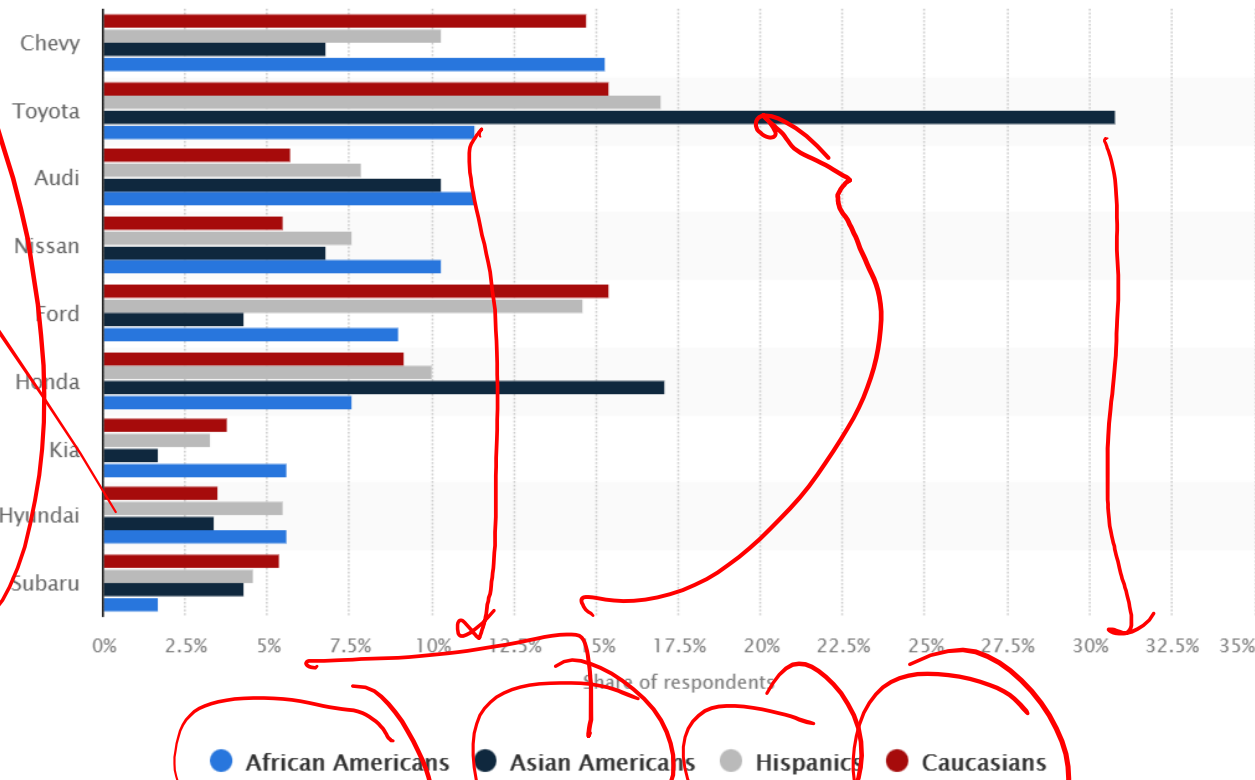
Right-tailed

One tailed Sig. = $0.736/2 = 0.368$

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Q13_1 (How likely are you to recommend the following retailer to friends and family members, on a scale from 0-10, where 0=Not at all likely to recommend, and 10= Will definitely recommend. Amazon)	Equal variances assumed	3.991	.046	.337	1606	.736	.030	.089	-.144	.204
	Equal variances not assumed			.357	992.175	.721	.030	.084	-.134	.194

Comparing Proportions of Two Populations



Number of respondents

2,484*

Age group

18 years and older

Special properties

Respondents aged 18-74

Method of interview

Online survey

Supplementary notes

* African Americans, n=301
 Asian Americans, n=117
 Hispanics, n=329
 Caucasians, n=1,737.

Estimating the Difference Between Two Population Proportions

Let p_1 equal the proportion in population 1 and p_2 equal the proportion in population 2.

The difference between the two-population proportions is $p_1 - p_2$.

To estimate $p_1 - p_2$, we will select a simple random sample of size n_1 from population 1 and a simple random sample of size n_2 from population 2.

The point estimator of the difference between the proportions of the populations 1 and 2 is $\bar{p}_1 - \bar{p}_2$

Sampling Distribution of $\bar{p}_1 - \bar{p}_2$

Expected Value

$$E(\bar{p}_1 - \bar{p}_2) = p_1 - p_2$$

Standard Error

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

where: n_1 = size of sample taken from population 1
 n_2 = size of sample taken from population 2

Sampling Distribution of $\bar{p}_1 - \bar{p}_2$

If the sample sizes are large, the sampling distribution of $\bar{p}_1 - \bar{p}_2$ can be approximated by a normal probability distribution.

The sample sizes are sufficiently large if all of these conditions are met:

$$n_1 p_1 \geq 5$$

$$n_1(1 - p_1) \geq 5$$

$$n_2 p_2 \geq 5$$

$$n_2(1 - p_2) \geq 5$$

Interval Estimation of $p_1 - p_2$

Interval Estimate

$$\bar{p}_1 - \bar{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}$$

where:

$1 - \alpha$ is the confidence coefficient

Hypothesis Tests about $p_1 - p_2$

Hypotheses

Let's stick with the Two-tailed test (i.e. $p_1 = p_2$)

$$H_0: p_1 - p_2 \geq 0$$

$$H_a: p_1 - p_2 < 0$$

Left-tailed

$$H_0: p_1 - p_2 \leq 0$$

$$H_a: p_1 - p_2 > 0$$

Right-tailed

$$H_0: p_1 - p_2 = 0$$

$$H_a: p_1 - p_2 \neq 0$$

Two-tailed

Hypothesis Tests about $p_1 - p_2$

Pooled Estimate of Standard Error of $\bar{p}_1 - \bar{p}_2$

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\bar{p}(1 - \bar{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

where:

$$\bar{p} = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2}$$

Hypothesis Tests about $p_1 - p_2$

Test Statistic

$$z = \frac{(\bar{p}_1 - \bar{p}_2)}{\sqrt{\bar{p}(1 - \bar{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

where:

$$\bar{p} = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2}$$

Difference Between Two Population Proportions:

Example: Amazon vs. Nordstrom

There is a difference between the proportion of Amazon customers who are likely to purchase within the next month as compared to the proportion of Nordstrom customers?

NORDSTROM



Likelihood to Purchase						
Q12 For the following retailers, when do you anticipate your next purchase at that store might be? PROGRAMMER: ALLOW ONLY ONE ANSWER PER BRAND.						
	Brand X	JC PenneyTarget	Kohl's	Nordstrom	Amazon	TJ Maxx
Within the next month	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Within next 1-3 months	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
With next 3-6 months						
Within next 6-9 months	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Within 9 months – 1 year	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Longer than 1 year	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I'm not sure when I will shop there again						

Hypothesis Tests about $p_{Amazon} - p_{Nordstrom}$

Hypothesis

$$H_0: p_1 - p_2 = 0$$

$$H_a: p_1 - p_2 \neq 0$$

amazon

NORDSTROM

Test Statistic

$$z = \frac{(\bar{p}_1 - \bar{p}_2)}{\sqrt{\bar{p}(1 - \bar{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

where:

$$\bar{p} = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2}$$



SPSS doesn't directly do this test, so we seek help elsewhere 😊

Social Science Statistics

	Name				Label	
689	d3d_7	N			target (Do you own a credit or charge card from any of these department ...	N
690	d3d_8	N			CPenney (Do you own a credit or charge card from any of these depart...	N
691	d3d_9	N			ohl's (Do you own a credit or charge card from any of these department ...	N
692	d3d_10	N			one of these (Do you own a credit or charge card from any of these dep...	N
693	d4	N			which of the following statements best describes your current employe...	{1
694	d4a	N			which of the following statements best describes you?	{1
695	responseid	N		0	responseid	N
696	q12x1	N		0	Brand X	{1
697	q12x2	N		0	JC Penney	{1
698	q12x11	N		0	Kohl's	{1
699	q12x3	N		0	Nordstrom	{1
700	q12x23	N		0	Amazon	{1
701	q12x15	N		0	TJ Maxx	{1
702	q13_1_1	N		0	Q13_1 (How likely are you to recommend the following retailer to friends ...	{0
703	q13_1_2	N		0	Q13_1 (How likely are you to recommend the following retailer to friends ...	{0
704	q13_1_11	N		0	Q13_1 (How likely are you to recommend the following retailer to friends ...	{0
705	q13_1_3	N		0	Q13_1 (How likely are you to recommend the following retailer to friends ...	{0
706	q13_1_23	Numeric	2	0	Q13_1 (How likely are you to recommend the following retailer to friends ...	{0
707	q13_1_15	Numeric	2	0	Q13_1 (How likely are you to recommend the following retailer to friends ...	{0
708	q19a_24	Numeric	1	0	Vince Camuto (How familiar are you with the following brands, for women'...	{1
709	q19anew_1	Numeric	1	0	Clinique (How familiar are you with the following brands, for skincare and ...	{1

- Power Analysis >
- Reports >
- Descriptive Statistics >
- Bayesian Statistics >
- Tables >
- Compare Means >
- General Linear Model >
- Generalized Linear Models >
- Mixed Models >
- Correlate >
- Regression >
- Loglinear >
- Neural Networks >
- Classify >
- Dimension Reduction >
- Scale >
- Nonparametric Tests >
- Forecasting >
- Survival >
- Multiple Response >
- Missing Value Analysis... >
- Multiple Imputation >
- Complex Samples >
- Simulation... >
- Quality Control >
- Spatial and Temporal Modeling... >
- Direct Marketing >

- Frequencies...
- Descriptives...
- Explore...
- Crosstabs...
- TURF Analysis
- Ratio...
- P-P Plots...
- Q-Q Plots...

	Name	Type	Width	Decimals	Label
689	d3d_7	Numeric	1	0	Target (Do you own a credit or charge card from any of these department ...
690	d3d_8	Numeric	1	0	JCPenney (Do you own a credit or charge card from any of these depart...
691	d3d_9	Numeric	1	0	Kohl's (Do you own a credit or charge card from any of these department ...
692	d3d_10	Numeric	1	0	None of these (Do you own a credit or charge card from any of these dep...
693	d4	Numeric	1	0	Which of the following statements best describes your current employe...
694	d4a	Numeric	1	0	Which of the following statements best describes you?

Frequencies

Variable(s):
 Nordstrom [q12x3]
 Amazon [q12x23]

Display frequency tables
 Create APA style tables

OK Paste Reset Cancel Help

Statistics...
 Charts...
 Format...
 Style...
 Bootstrap...

695	responseid	Numeric			
696	q12x1	Numeric			
697	q12x2	Numeric			
698	q12x11	Numeric			
699	q12x3	Numeric			
700	q12x23	Numeric			
701	q12x15	Numeric			
702	q13_1_1	Numeric			How likely are you to recommend the following retailer to friends ...
703	q13_1_2	Numeric	2	0	Q13_1 (How likely are you to recommend the following retailer to friends ...
704	q13_1_11	Numeric	2	0	Q13_1 (How likely are you to recommend the following retailer to friends ...
705	q13_1_3	Numeric	2	0	Q13_1 (How likely are you to recommend the following retailer to friends ...
706	q13_1_23	Numeric	2	0	Q13_1 (How likely are you to recommend the following retailer to friends ...
707	q13_1_15	Numeric	2	0	Q13_1 (How likely are you to recommend the following retailer to friends ...
708	q19a_24	Numeric	1	0	Vince Camuto (How familiar are you with the following brands, for women'...
709	q19anew_1	Numeric	1	0	Clinique (How familiar are you with the following brands, for skincare and ...

Output2 [Document2] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output

- Log
- Frequencies
 - Title
 - Notes
 - Statistics
- Frequency Table
 - Title
 - Nordstrom
 - Amazon

Statistics

		Nordstrom	Amazon
N	Valid	1590	1608
	Missing	2741	2723

Frequency Table

Nordstrom

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Within the next month	262	6.0	16.5	16.5
	Within next 1-3 months	350	8.1	22.0	38.5
	With next 3-6 months	300	6.9	18.9	57.4
	Within next 6-9 months	147	3.4	9.2	66.6
	Within 9 months – 1 year	135	3.1	8.5	75.1
	Longer than 1 year	105	2.4	6.6	81.7
	I'm not sure when I will shop there again	291	6.7	18.3	100.0
Total		1590	36.7	100.0	
Missing	System	2741	63.3		
Total		4331	100.0		

Amazon

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Within the next month	1068	24.7	66.4	66.4
	Within next 1-3 months	336	7.8	20.9	87.3
	With next 3-6 months	108	2.5	6.7	94.0
	Within next 6-9 months	26	.6	1.6	95.6
	Within 9 months – 1 year	17	.4	1.1	96.7
	Longer than 1 year	6	.1	.4	97.1
	I'm not sure when I will shop there again	47	1.1	2.9	100.0
Total		1608	37.1	100.0	
Missing	System	2723	62.9		
Total		4331	100.0		

NORDSTROM

amazon

IBM SPSS Statistics Processor is ready Unicode:ON

Z Score Calculator for 2 Population Proportions

The z-score test for two population proportions is used when you want to know whether two populations or groups (e.g., liberals and conservatives) differ significantly on some single (categorical) characteristic - for example, whether they watch *South Park*.

To use the calculator, just input the proportions (or absolute numbers) for your two samples in the textboxes below, together with the size of each sample. Then press the "Calculate Z" button.

Sample 1 Proportion (or total number)

Sample 1 Size (N_1)

Sample 2 Proportion (or total number)

Sample 2 Size (N_2)

Significance Level:

- 0.01
- 0.05
- 0.10

One-tailed or two-tailed hypothesis?:

- One-tailed
- Two-tailed

No calculation performed yet.

NORDSTROM

Z Score Calculator for 2 Population Proportions

Success!

You'll find the values for z and p below. Blue means your result is significant, red means it's not.

Sample 1 Proportion (or total number)

Sample 1 Size (N_1)

Sample 2 Proportion (or total number)

Sample 2 Size (N_2)

Significance Level:

- 0.01
- 0.05
- 0.10

One-tailed or two-tailed hypothesis?:

- One-tailed
- Two-tailed

The value of z is 28.6262. The value of p is $< .00001$. The result is significant at $p < .05$.

NORDSTROM

Hypothesis

$$H_0: p_1 - p_2 = 0$$
$$H_a: p_1 - p_2 \neq 0$$



Group Work

- Since we have 77 registered students, 19 groups will be created with students being randomly assigned
- Please submit your work no later than 10:00pm EST. There is no drop box, work needs to be submitted via email to harvir.bansal@uwaterloo.ca
- Please ensure that you clearly outline the steps taken in your analysis as well as the results.
- I'll make rounds in each "room" to answers any questions that you might have
- As I have said multiple times, data analysis is not a spectator sport so PLAY and have fun

Hands-on Data Analysis

- **All questions below refer to the brand tracking questionnaire and data made available for this class on LEARN**
 1. Are there any stores where males and females differ in their commitment to the brand? Explain via analysis
 2. Are Kohl's customers equally likely to purchase a product within the next 9 months as compared to TJ Maxx's customers? Explain via analysis
 3. As compared to Amazon's NPS score which retailer must work hardest to compete? Explain via analysis