



## Module 9: Nonparametric Tests

The Applied Research Center

# Module 9 Overview

---

- ▶ Nonparametric Tests
- ▶ Parametric vs. Nonparametric Tests
- ▶ Restrictions of Nonparametric Tests
- ▶ One-Sample Chi-Square Test
- ▶ Chi-Square Test of Independence
- ▶ Other Nonparametric Tests

# What Are Nonparametric Tests?

---

- ▶ Nonparametric tests require few, if any assumptions about the shapes of the underlying population distributions
- ▶ For this reason, they are often used in place of parametric tests if or when one feels that the assumptions of the parametric test have been too grossly violated (e.g., if the distributions are too severely skewed).

# Parametric or Nonparametric tests?

---

- ▶ If all assumptions are met, use Parametric techniques
- ▶ Use Nonparametric techniques
  - ▶ When the dependent variable is either nominal or ordinal
  - ▶ If the distribution of the dependent variable is skewed
  - ▶ When the assumptions are not met, specifically:
    - ▶ Normality
    - ▶ Homogeneity of variance

# Restrictions

---

- ▶ Nonparametric tests do have at least two major disadvantages in comparison to parametric tests:
- ▶ First, nonparametric tests are less powerful. Why? Because parametric tests use more of the information available in a set of numbers.
- ▶ Parametric tests make use of information consistent with interval or ratio scale (or continuous) measurement, whereas nonparametric tests typically make use of nominal or ordinal (or categorical) information only.

## Restrictions (cont' d)

---

- ▶ Second, parametric tests are much more flexible, and allow you to test a greater range of hypotheses. For example, ANOVA designs allow you to test for interactions between variables in a way that is not possible with nonparametric alternatives.
- ▶ There are nonparametric techniques to test for certain kinds of interactions under certain circumstances, but these are much more limited than the corresponding parametric techniques.

# Chi-Square Test

---

- ▶ Used to test variables that have nominal data
- ▶ Examples of nominal data
  - ▶ Gender
  - ▶ Political affiliation
  - ▶ Course delivery method
- ▶ Three main types:
  - ▶ Goodness-of-fit (or one-sample) test
  - ▶ Test of independence (or association)
  - ▶ Independent-samples test

# Chi-Square Goodness-of-Fit Test

---

- ❑ Evaluates whether the proportions of individuals who fall into categories are equal to hypothesized values
- ❑ The variable can have two or more categories
- ❑ The categories can have quantitative (one category reflects a higher value than another; e.g., Likert scale responses of Agree and Disagree) or qualitative grouping (e.g., course delivery method)
- ❑ Note: The chi-square test does not recognize any quantitative distinction among categories; it simply assesses whether the proportions associated with the categories are significantly different from the hypothesized values



# Chi-Square Goodness-of-Fit Test (cont' d)

---

- ▶ When conducting a chi-square test in SPSS, you must first specify the values for the hypothesized proportions
- ▶ The proportions can be
  - ▶ Equal (proportion =  $1 / \text{number of categories}$ )
  - ▶ Not equal (determined based on previous research)

# Chi-Square Goodness-of-Fit Test (cont' d)

---

- Next, you must formulate a null hypothesis (the expected frequencies = hypothesized proportions)
- Then, specify the expected frequencies associated with the categories.
  - ▣ The expected frequency =  
hypothesized proportion x total sample size.
  - ▣ If hypothesized proportions are equal, the expected frequencies will be equal
  - ▣ If hypothesized proportions are unequal, the expected frequencies will be unequal

# Contingency Tables

---

- ▶ Helps arrange the data in a study for the purpose of showing how each group split itself up on the variable
- ▶ Make it easier to understand the null hypothesis and why the data led to the rejection or not of the null.

# Goodness-of-Fit Example

---

- A professor was interested in whether students prefer to take classes face to face or online. A poll was taken with 100 students to determine their preference for learning.
- Step 1: State the null and alternate hypotheses
  - $H_0$  : The number of students who prefer to take classes face to face is equal to the number of students who prefer to take classes online
  - $H_a$  : The number of students who prefer to take classes face to face is not equal to the number of students who prefer to take classes online

# Goodness-of-Fit Example (cont' d)

---

## ▶ Step 2: Create the contingency table

	Observed Frequencies	Hypothesized Proportions	Expected Frequencies
Face to Face	39	.50	50
Online	61	.50	50

## ▶ Step 3: Input the data into SPSS

- ▶ Create a variable labeled DeliveryMethod
- ▶ Assign values: 1 = Face to Face, 2 = Online
- ▶ Enter the data (in one column): 39, 1's and 61, 2's

# Goodness-of-Fit Example (cont' d)

---

## ▶ Step 4: Run the Analysis

- ▶ Analyze → Nonparametric → Legacy Dialogs → Chi-square
- ▶ Move the variable (DeliveryMethod) to the “Test variable list” box
- ▶ Under “Expected Values,” select “All categories equal”
- ▶ Click “Ok”

# Goodness-of-Fit Example (cont' d)

---

**DeliveryModel**

	Observed N	Expected N	Residual
Facetoface	39	50.0	-11.0
Online	61	50.0	11.0
Total	100		

**Test Statistics**

	DeliveryModel
Chi-Square	4.840 <sup>a</sup>
df	1
Asymp. Sig.	.028



# Goodness-of-Fit Example (cont' d)

---

- ▶ Step 4: Make a decision regarding the null
  - ▶  $\chi^2 = 4.84$
  - ▶  $df = 1$  (Number of categories – 1)
  - ▶  $p = .028$
- ▶ What is our decision regarding the null?



# Goodness-of-Fit Example (cont' d)

---

- ▶ Using the level of significance = .05, do we reject or fail to reject the null?
  - ▶ If  $p < .05$ , we reject the null
  - ▶ if  $p > .05$ , we fail to reject the null
- ▶ According to SPSS,  $p = .028$
- ▶  $.028 < .05$ , therefore, we reject the null!

# Goodness-of-Fit Example (cont' d)

---

- ▶ Step 5: Write up your results.
  - ▶ The null hypothesis stated that the number of students who prefer to take classes face to face is equal to the number of students who prefer to take classes online. A chi-square goodness-of-fit test revealed that significantly more students prefer to take courses online (61%) compared to face to face (39%,  $X^2 = 4.84$ ,  $df = 1$ ,  $p = .028$ ). Consequently, the null hypothesis was rejected.

# Chi-Square Test of Independence

---

- ▶ A test to see whether a relationship (or association) exists between two nominal variables
- ▶ Similar to a bivariate correlation

# Test of Independence Example

---

- ▶ A professor is interested in whether there is a relationship between gender and preference for learning statistics face to face or online. A poll was taken with 100 students (50 males and 50 females) to determine their preference for learning.
- ▶ Step 1: State the null and alternate hypotheses
  - ▶  $H_0$  : There is no relationship between gender and preference for learning statistics face to face or online
  - ▶  $H_a$  : There is a relationship between gender and preference for learning statistics face to face or online

# Test of Independence Example (cont' d)

---

- ▶ Step 2: Create the contingency table

	Face to face	Online
Males	20	30
Females	30	20

- ▶ Step 3: Input the data into SPSS

- ▶ Create 2 variables: Gender and DeliveryMethod
- ▶ Assign values
- ▶ Enter the data:
  - ▶ For Gender: 50, 1's and 50, 2's
  - ▶ For DeliveryMethod: 20 males and 30 females (face to face), 30 males and 20 females (online)

# Test of Independence Example (cont' d)

---

## ▶ Step 4: Run the Analysis

- ▶ Analyze → Descriptives → Crosstabs
- ▶ Move Gender to the “Row” box and DeliveryMethod to the “Column” box
- ▶ Select “Statistics” then select “Chi square” and “Phi and Cramer’s V” boxes, Click “Continue”
- ▶ Select “Cells” then under “Counts” select “Observed” and “Expected,” Click “Continue”
- ▶ Select “Display clustered bar charts” at the bottom left.
- ▶ Click “Ok”

# Test of Independence Example (cont' d)

## Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender * DeliveryMethod	100	100.0%	0	.0%	100	100.0%

## Gender \* DeliveryMethod Crosstabulation

			DeliveryMethod		Total
			Face to face	Online	
Gender	Male	Count	20	30	50
		Expected Count	25.0	25.0	50.0
	Female	Count	30	20	50
		Expected Count	25.0	25.0	50.0
Total		Count	50	50	100
		Expected Count	50.0	50.0	100.0



# Test of Independence Example (cont' d)

**Chi-Square Tests**

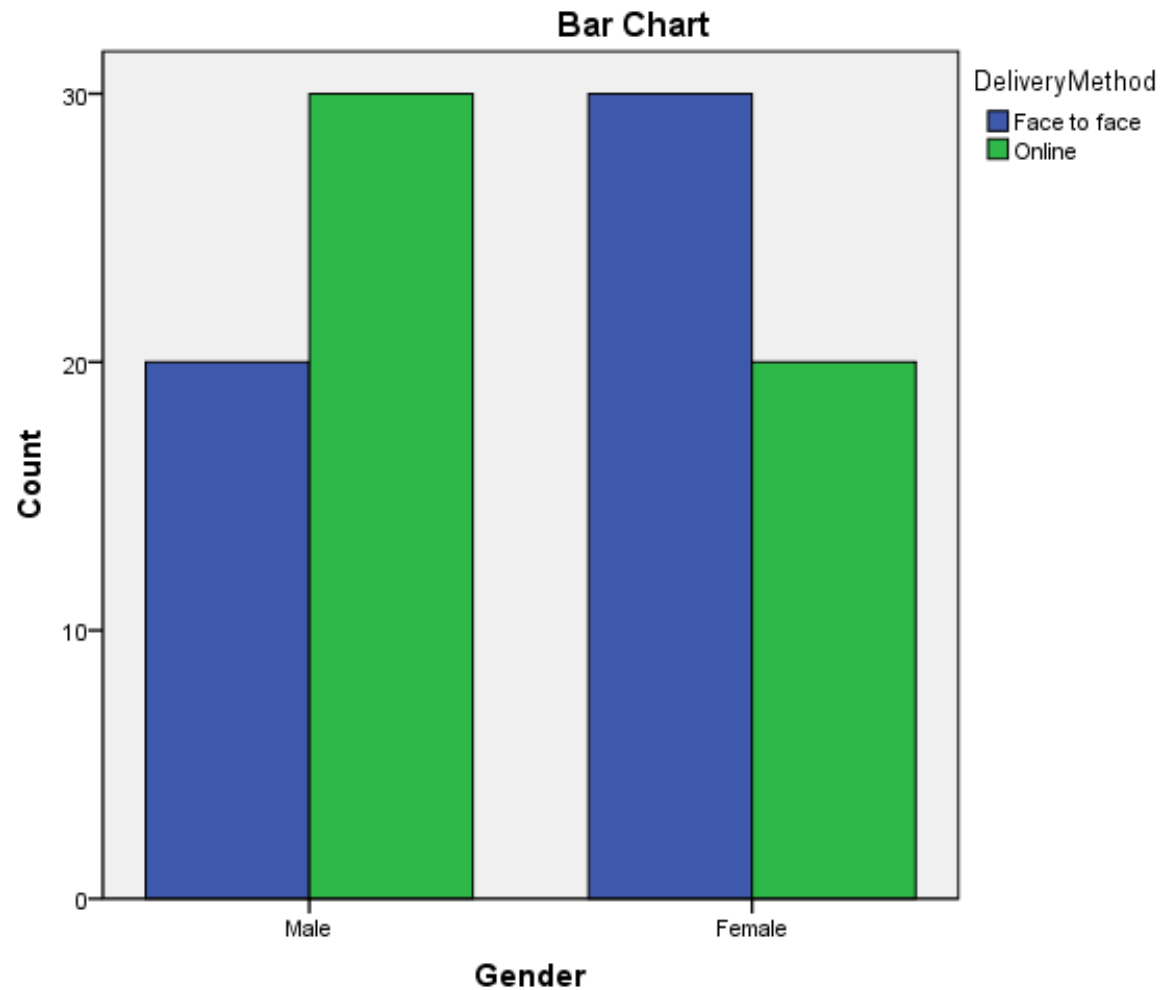
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.000 <sup>a</sup>	1	.046		
Continuity Correction <sup>b</sup>	3.240	1	.072		
Likelihood Ratio	4.027	1	.045		
Fisher's Exact Test				.071	.036
Linear-by-Linear Association	3.960	1	.047		
N of Valid Cases	100				

**Symmetric Measures**

		Value	Approx. Sig.
Nominal by Nominal	Phi	-.200	.046
	Cramer's V	.200	.046
N of Valid Cases		100	



# Test of Independence Example (cont' d)



# Test of Independence Example (cont' d)

---

- ▶ Step 4: Make a decision regarding the null
  - ▶  $X^2 = 4.00$
  - ▶  $df = 1$ 
    - ▶ (Number of categories in 1<sup>st</sup> variable – 1)(Number of categories in 2<sup>nd</sup> variable – 1)
  - ▶  $p = .046$
- ▶ What is our decision regarding the null?

# Test of Independence Example (cont' d)

---

- ▶ Using the level of significance = .05, do we reject or fail to reject the null?
  - ▶ If  $p < .05$ , we reject the null
  - ▶ if  $p > .05$ , we fail to reject the null
- ▶ According to SPSS,  $p = .046$
- ▶  $.046 < .05$ , therefore, we reject the null!

# Test of Independence Example (cont' d)

---

- ▶ **Step 5: Write up your results.**
  - ▶ The null hypothesis stated that there is no relationship between gender and preference for learning statistics face to face or online. A chi-square test of independence revealed a significant relationship between gender and preference for learning statistics,  $\chi^2 (1, N = 100) = 4.00; p = .046$ . Consequently, the null hypothesis was rejected.

# Nonparametric Tests for Nominal Data

---

- ▶ The sign test
- ▶ The binomial test
- ▶ Fischer's Exact Test
- ▶ The chi-square test
- ▶ McNemar's test
- ▶ Cochran's test

# Nonparametric Tests for Ordinal Data

---

- ▶ The Mann-Whitney  $U$  test
  - ▶ Analogous to independent samples  $t$  test
- ▶ Wilcoxon signed-ranks test
  - ▶ Analogous to paired samples  $t$  test
- ▶ The Kruskal-Wallis one-way analysis of ranks
  - ▶ Analogous to a one-way ANOVA
- ▶ The Friedman's ANOVA
  - ▶ Analogous to RM ANOVA
- ▶ Spearman's rho
  - ▶ Analogous to Pearson correlation

# Module 9 Summary

---

- ▶ Nonparametric Tests
- ▶ Parametric vs. Nonparametric Tests
- ▶ Restrictions of Nonparametric Tests
- ▶ One-Sample Chi-Square Test
- ▶ Chi-Square Test of Independence
- ▶ Other Nonparametric Tests

# Review Activity

---

- ▶ Please complete the review activities at the end of the module.
- ▶ You can complete the review activities and module as many times as you like.
- ▶ Upon completion of the review activities you will have an opportunity to print your Certificate of Completion, along with instructions for submitting your certificate to the ARC.



# Statistics PD Course Modules

---

- ▶ Module 1: Introduction to Statistics
- ▶ Module 2: Introduction to SPSS
- ▶ Module 3: Descriptive Statistics
- ▶ Module 4: Inferential Statistics
- ▶ Module 5: Correlation
- ▶ Module 6: *t*-Tests
- ▶ Module 7: ANOVAs
- ▶ Module 8: Linear Regression
- ▶ Module 9: Nonparametric Procedures