CARL F. SIEBERT
AND DARCY CLAY SIEBERT

Data Analysis with Small Samples and Non-Normal Data

Nonparametrics and Other Strategies

POCKET GUIDES TO
SOCIAL WORK RESEARCH METHODS

Data Analysis with Small Samples and Non-Normal Data

POCKET GUIDES TO
SOCIAL WORK RESEARCH METHODS

Series Editor
Tony Tripodi, DSW
Professor Emeritus, Ohio State University

Determining Sample Size
Balancing Power, Precision, and Practicality
Patrick Dattalo

Preparing Research Articles
Bruce A. Thyer

Systematic Reviews and Meta-Analysis
Julia H. Littell, Jacqueline Corcoran, and Vijayan Pillai

Historical Research
Elizabeth Ann Danto

Confirmatory Factor Analysis
Donna Harrington

Randomized Controlled Trials
Design and Implementation for Community-Based Psychosocial Interventions
Phyllis Solomon, Mary M. Cavanaugh, and Jeffrey Draine

Needs Assessment
David Royse, Michele Staton-Tindall, Karen Badger, and J. Matthew Webster

Multiple Regression with Discrete Dependent Variables
John G. Orme and Terri Combs-Orme

Developing Cross-Cultural Measurement
Thanh V. Tran

Intervention Research
Developing Social Programs
Mark W. Fraser, Jack M. Richman, Maeda J. Galinsky, and Steven H. Day

Developing and Validating Rapid Assessment Instruments
Neil Abell, David W. Springer, and Akihito Kamata

Clinical Data-Mining
Integrating Practice and Research
Irwin Epstein

Strategies to Approximate Random Sampling and Assignment
Patrick Dattalo

Analyzing Single System Design Data
William R. Nugent

Survival Analysis
Shenyang Guo

The Dissertation
From Beginning to End
Peter Lyons and Howard J. Doueck

Cross-Cultural Research
Jorge Delva, Paula Allen-Meares, and Sandra L. Momper

Secondary Data Analysis
Thomas P. Vartanian

Narrative Inquiry
Kathleen Wells

Structural Equation Modeling
Natasha K. Bowen and Shenyang Guo

Finding and Evaluating Evidence
Systematic Reviews and Evidence-Based Practice
Denise E. Bronson and Tamara S. Davis

Policy Creation and Evaluation
Understanding Welfare Reform in the United States
Richard Hoefer

Grounded Theory
Julianne S. Oktay

Systematic Synthesis of Qualitative Research
Michael Saini and Aron Shlonsky

Quasi-Experimental Research Designs
Bruce A. Thyer

Conducting Research in Juvenile and Criminal Justice Settings
Michael G. Vaughn, Carrie Pettus-Davis, and Jeffrey J. Shook

Qualitative Methods for Practice Research
Jeffrey Longhofer, Jerry Floersch, and Janet Hoy

Analysis of Multiple Dependent Variables
Patrick Dattalo

Culturally Competent Research
Using Ethnography as a Meta-Framework
Mo Yee Lee and Amy Zaharlick

Using Complexity Theory for Research and Program Evaluation
Michael Wolf-Branigin

Basic Statistics in Multivariate Analysis
Karen A. Randolph and Laura L. Myers

Research with Diverse Groups:
Diversity and Research-Design and Measurement Equivalence
Antoinette Y. Farmer and G. Lawrence Farmer

Conducting Substance Use Research
Audrey L. Begun and Thomas K. Gregoire

A Social Justice
Approach to Survey Design and Analysis
Llewellyn J. Cornelius and Donna Harrington

Participatory Action Research
Hal A. Lawson, James Caringi, Loretta Pyles, Janine Jurkowski, and Christine Bolzak

Developing, Selecting, and Using Measures
David F. Gillespie and Brian E. Perron

Mixed Methods Research
Daphne C. Watkins and Deborah Gioia

Content Analysis
James W. Drisko and Tina Maschi

Group Work Research
Charles D. Garvin, Richard M. Tolman, and Mark J. Macgowan

Developing Cross-Cultural Measurement in Social Work Research and Evaluation, Second Edition
Thanh V. Tran, Tam H. Nguyen, and Keith T. Chan

Data Analysis With Small Samples and Non-Normal Data
Nonparametrics and Other Strategies
Carl F. Siebert and Darcy Clay Siebert

CARL F. SIEBERT
DARCY CLAY SIEBERT

Data Analysis with Small Samples and Non-Normal Data

Nonparametrics and Other Strategies

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and certain other countries.

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America.

Library of Congress Cataloging-in-Publication Data
Names: Siebert, Carl F., author. | Siebert, Darcy Clay, author.
Title: Data analysis with small samples and non-normal data :
nonparametrics and other strategies / Carl F. Siebert, Darcy Clay Siebert.
Description: New York : Oxford University Press, [2018] |
Series: Pocket guides to social work research methods |
Includes bibliographical references and index.
Identifiers: LCCN 2017010962 (print) | LCCN 2017014700 (ebook) |
ISBN 9780199391509 (updf) | ISBN 9780199391516 (epub) |
ISBN 9780199391493 (alk. paper)
Subjects: LCSH: Nonparametric statistics.
Classification: LCC QA278.8 (ebook) | LCC QA278.8 .S52 2017 (print) |
DDC 519.5/4—dc23
LC record available at https://lccn.loc.gov/2017010962

1 3 5 7 9 8 6 4 2

Printed by WebCom, Inc., Canada

We dedicate this book to our children and grandchildren,
who are constant sources of joy and inspiration.

Contents

Acknowledgments

We both thank Dr. Tony Tripodi for his encouragement and patience. Thanks, too, to Dr. Myles Hollander for sparking Carl's interest in non-parametric analysis and for all the lessons learned in his courses. We especially wish to acknowledge Dr. Carl Henley for his early support of Darcy's abilities, and for teaching her that statistics are her friends, not her enemies.

I, Carl, also want to thank my co-author, wife, and partner for allowing me the privilege of spending my life with her, experiencing unending love, and being constantly inspired by her intelligence and dedication to helping others.

Data Analysis with Small Samples and Non-Normal Data

1

Introduction to Nonparametrics

OVERVIEW

Research training in the social sciences must cover a wide range of methodological and statistical coursework in a few years of doctoral study. Frequently, doctoral students take two or three statistics courses to learn basic analytic techniques, and students who take additional courses often do so to learn advanced inferential techniques, such as Hierarchical Linear Modeling (HLM) and Structural Equation Modeling (SEM). These techniques can be very useful when one is working with large data sets and normally distributed data.

However, researchers regularly find themselves working with small data sets, and this is especially commonplace for intervention research studies, which are the hallmark of social science research. Similarly, researchers may find themselves with data sets that simply do not meet the assumptions required for the parametric techniques they were trained to use. Yet very little, if any, time in their research training is devoted to the analytic techniques that would allow them to properly analyze these kinds of data—nonparametric statistics. This gap in education leaves individuals without the necessary tools to conduct suitable analyses of their data, or more importantly, the lack of exposure may prevent researchers from recognizing that nonparametric tools are available to them. The researchers' primary option is to hope for robustness when they utilize parametric methods, and this has the potential to guide the person to the wrong conclusions based on the inappropriate use of these statistical procedures.

This book is written in clear, understandable language for nonstatisticians who want to conduct statistical analyses on a small sample and/or variables that are not normally distributed. The nonparametric procedures covered herein describe how researchers can analyze data that otherwise might fall by the wayside. For example, successful grant proposals are frequently supported by believable pilot data, and these pilots often consist of small samples. Funding may hinge on researchers' ability to show significant findings, and using parametric methods for data that do not meet the assumptions or have insufficient power can result in nonsignificant findings when significant findings are actually present. Conversely, the researcher may find significant results that are an artifact of using the inappropriate analytic technique.

In this book, we assume that you, the reader, want to understand nonparametric statistics conceptually, and then learn the steps to complete a nonparametric statistical test. We intend the book to be a reference source more than a textbook, so we have formatted the information in ways to make it easy to understand. For example, you will find sections in this book where we have used bullet points rather than standard text in paragraphs. This is to make the information easy to find and simple to read.

We also do not cover the mathematical formulas, nor do we walk you through the actual derivation of how to obtain the critical values for hypothesis testing. In addition, we do not cover how to conduct a nonparametric test by hand, as we assume you have access to a PC and have a basic understanding of IBM® SPSS® Statistics software (SPSS) software. We have chosen SPSS software because it offers a point-and-click process for nonparametric procedures that is reasonably easy to follow. Thus, this book does not offer R or SAS users the code to conduct nonparametric tests; however, we include the SPSS syntax in Appendix A. Accordingly, this book is written for nonstatisticians, yet it might prove useful for applied statisticians as well, because we supply links to additional resources that will meet the needs of those interested in the mathematical foundations of nonparametrics.

Another important topic that is not integrated into the chapters is the difference between statistical significance (i.e., results from a statistical test) and practical importance. Statistical tests are mathematical procedures that provide a mathematical test of a hypothesis. Finding significance from a statistical test is not the same as determining how important the statistical finding might be for a particular research question. Your data analyses may indicate statistical significance, when in the real world, the difference is practically inconsequential. We strongly urge anyone conducting statistical analyses, especially those making decisions from statistical findings, to review fully the topic of practical importance related to statistics. (For more information on statistical significance versus practical importance, see Hojat & Xu, 2004.)

Finally, we do not introduce bootstrapping as an option in our chapters. Bootstrapping is another option for researchers dealing with small data sets, and in simple terms, it draws inferences about the population under investigation by resampling the data set from which a sample of the population came. Since the introduction of the bootstrap (Efron, 1979), procedures have been improved and its use has been demonstrated to be a viable alternative when dealing with challenging data (Davidson & Hinkley, 1998; Janssen & Pauls, 2003; Wu, 1986). Because bootstrapping is an advanced technique, we leave the discussion of it and how to use it to others.

ORGANIZATION

This book is organized so you can quickly locate the nonparametric technique that is optimal for your data, and then follow the clear, step-by-step instructions for using SPSS to conduct your analysis. Each chapter includes a scenario typically found in social science research to contextualize the analytic techniques, including details about the research design, population, sample, groups within the sample, number of cases/observations/data points, time (if relevant), variables, level of measurement, description of the variables' distribution, and a series of research questions and hypotheses. Next, using the scenario as the example, relevant research questions and hypotheses guide the discussion of the nonparametric method(s) that can be used to answer them.

In this way, you can easily match your particular questions to the appropriate technique, and then follow the example to its conclusion and interpretation.

Beginning with Chapter 2, each chapter starts with a list of chapter objectives, a scenario that serves as the backdrop for the chapter's research questions, and a table of the nonparametric methods covered in the chapter. Chapter 2 covers nonparametric methods using single variables or groups. Chapter 3 presents procedures for comparing two or more independent groups on a particular characteristic, and Chapter 4 describes techniques for comparing two or more related groups. The fifth chapter describes methods for predicting a particular value or level based on a set of independent variables, when the data do not meet the assumptions for least-squares regressions.

The primary software used throughout the book is SPSS, but because SPSS currently does not include nonparametric regression techniques in its software, we utilize an add-in to Excel—XLSTAT—for the steps to conduct the predictive analyses.

Additional useful information is supplied in the Appendices (e.g., SPSS syntax, missingness) and from SPSS Help (e.g., how to prepare your data for analysis). We also suggest additional resources (i.e., web links, articles, books) that will provide detailed foundational information and formulas. On this book's companion website you will find all the data sets used in the book's analyses. After downloading the data sets, you can follow the steps for each nonparametric method using the example data to reinforce your learning and understanding of the process.

You may be tempted to go straight to the data analysis chapters at this point, but if you do, you will not necessarily understand the foundation for what you are doing. Consequently, you may err in describing your procedures to others, or in presenting the conclusions and interpretations of your analyses. We urge you to review carefully all the information in this chapter before proceeding.

WHAT ARE NONPARAMETRIC STATISTICS?

For some, nonparametric statistics or tests can be difficult to understand at first, because most "Introduction to Statistics" courses are based on parametric procedures. Nonparametric statistics are a set of tests that are based on rankings, and they are often called distribution-free statistics, meaning that nonparametric techniques "rely on no or few assumptions about the shape or parameters of the population distribution from which the sample was drawn" (Hoskin, n.d., p. 2). In addition, nonparametric procedures can be used with nominal and ordinal data, as well as the interval and ratio variables required for many parametric procedures.

KEY TERMS

Several key terms used are used throughout the book. These terms are as follows:

- Hypothesis — a testable proposition that guides your study. The null hypothesis indicates no statistically significant difference between two groups or measured phenomena; that is, they are drawn from the same sample. Thus, any difference in the outcome variable is due to sampling error, and the independent variable won't make a difference. The alternative (research) hypothesis suggests that the null hypothesis is incorrect, and that a relationship does exist. The alternative hypothesis can be either nondirectional or directional. In a nondirectional hypothesis, a difference exists, but the direction of the relationship (e.g., higher or lower) is unknown. The directional alternative hypothesis indicates the direction of the hypothesis (e.g., mean scores in X are greater than mean scores in Y), and it has more power to detect the direction of the relationship in which you are interested.
- Levels of measurement — a classifying scheme describing the kind of information ascribed to the values representing a variable.
- Parametric assumptions — guidelines that the data representing a variable(s) must meet before being used in a parametric statistical procedure.
- Robustness — a property of a statistical technique in which assumption violation will not adversely influence the test results (i.e., the test's actual Type I error rate is close to the .05 guideline).

- *P*-value — the probability of getting results *at least as extreme* as the ones you observed, given that the null hypothesis (the hypothesis of no effect/no difference) is correct.
- Independence — when one respondent does not influence how another person responds to a question.
- Sample size — the number of subjects or records in a data set.
- Effect size — a quantitative measure of the magnitude of a difference or relationship (e.g., the correlation between two variables, a mean difference, or the influence of an intervention) that is not influenced by sample size. Typical effect sizes for social sciences (see Cohen, 1988, 1992) are: .1 = small effect, .3 = moderate effect, .5 = large effect.
- Significance — in this book, the term for statistical significance, meaning that the *p*-value is less than the a priori alpha level chosen for the study. Practical significance, or the importance of a finding in the real world, is labeled as such.
- Response coverage — the degree to which respondents provide data for all possible response options for an item or variable.
- Outlier — a response that is extreme in value, either low or high, from the majority of other responses.
- Ties — when two responders or observations provide the same value.

WHEN SHOULD NONPARAMETRIC PROCEDURES BE UTILIZED?

Nonparametric methods are the methods of choice when samples are small and/or do not meet the assumptions for parametric methods. These parametric assumptions are exceptionally important, as they are the glue holding the mathematical processes together. If your data do not meet one or more of the assumptions and you use the parametric procedure anyway (yes, the statistical software will provide results for a parametric procedure even if the assumptions are not met), you may very well come to incorrect conclusions. In addition, when samples are small, the ability to verify that your data meet parametric assumptions becomes difficult and sometimes impossible.

Although not all parametric procedures have the same set of assumptions, the primary parametric assumptions are as follows:

1. The population data are approximately normally distributed.
2. The population variances are equal (i.e., homogeneity of variance among groups is present).
3. The sample is adequately large.
4. A linear relationship exists between or among the variables.
5. Homoscedasticity is present in regression models.
6. Residuals (error) follow a normal distribution.

Other methodological assumptions are fundamental when performing any kind of statistical analysis. These assumptions are as follows:

7. The data are sampled randomly from a defined population.
8. The data consist of independent observations (apart from paired data).
9. The levels of measurement match the variable design.

Let us examine these assumptions, one at a time, to understand, conceptually, the circumstances in which a violation in an assumption might require the use of a nonparametric strategy.

PARAMETRIC ASSUMPTIONS

Assumption 1. The Population Data Are Approximately Normally Distributed

That the data are approximately normally distributed may be the most common parametric assumption (Hoskins, n.d., p. 4). The normality assumption is important whenever a scale variable (i.e., continuous variable) is used in a parametric statistics test. Most parametric tests involve a scale variable that is assumed to represent data from a symmetric bell-shaped curve distribution. Small departures from the symmetrical bell curve often have limited negative consequences for an analysis, but larger departures from normality can make the results wholly incorrect and/or uninterpretable. Furthermore, although some parametric tests like ANOVA purport to be robust against mild violations of normality, this does not mean that you can ignore the normality assumption, particularly with small samples. In addition, you are left to answer the question as to what, exactly, constitutes a mild departure from normality.

All scale variables must be tested for normality prior to including them in a parametric test. This is especially true when conducting multivariable analyses in which multiple scale variables are included in a single statistical model (e.g., multivariate regression). Each variable must demonstrate its association with a smooth, continuous, and bell-shaped distribution. However, when the sample size is small, the gaps between values (i.e., the data points on the curve) make the test of normality more likely to fail. This is particularly true when testing a Likert-type response variable that is being used as a scale variable rather than an ordinal variable, and even more so when the response options number five or fewer. Thus, although many social scientists tend to use Likert-scaled response options as if they represent a continuous variable, one must be sure to take the size of the sample into consideration. For example, it is easier to justify the use of a Likert-type variable if the sample size is 1,000 and the response options range from 1 to 7 or 1 to 9 than if the sample size is 50 and the Likert response options are 1 to 4. Testing to determine if a variable can be treated as a scale variable is included in upcoming chapters.

You have two primary options when your variable fails the normality test. Your first option is to use a transformation strategy to normalize your variable so that it will pass a normality test (e.g., log or square transformation). Not all variables can be transformed into a normal distribution (e.g., data with zeroes and negative values cannot be log transformed), and when your transformation attempts fail, you cannot simply use them anyway, hoping that your test will be robust against the normality violation. When a non-normal variable does respond well to transformation, you may use it in a parametric procedure, but then you cannot interpret your findings as if the variable were not transformed. You are faced with making sense of the transformed variable in the analysis and in your findings, and this can be complicated, if not impossible. (For more information on variable transformation, see McDonald, 2014, and Stat Trek, n.d.) Your second choice, when faced with a non-normal variable, is simply to use a nonparametric test or the aforementioned bootstrapping.

Another challenge to verifying normality is when your data have outliers. Because a scale variable is assumed to come from a bell-shaped normal curve, this includes the assumption that every value along the curve is a possible value for your variable (i.e., it is a reasonable representative of your sample). Outliers are data points that fall away from the mean value and that are multiple standard deviations away from the center of the normal curve. Therefore, variables with data that include outliers have a more difficult time passing normality tests. When your variable fails a normality test due to potential outliers, you must investigate the validity of including the case with the outlier. If you have irrefutable evidence that an outlier is not a legitimate member of your sample (e.g., it was a data entry error), you can choose to remove it or correct the error. If it does represent a legitimate data point, although extreme, it must be included in your analysis, even though it contributes to non-normality. In this situation, you should not run a parametric test anyway, hoping that the outlier is not influential. Instead, use a nonparametric procedure. (For more information on outliers, see Aguinis, Gottfredson, & Joo, 2013; Barnett & Lewis, 1978; Taylor, n.d.)

Finally, we must mention the use of Monte Carlo simulations. You may find in the literature a reference suggesting that someone has proved that it is reasonable to ignore assumption violations because they (or someone else, often referencing Glass, 1972) conducted a Monte Carlo study — basically research conducted with many dummy data sets. The dummy data sets are created by drawing values for variables that represent hypothetical distributions. With hundreds of data sets that are slightly different from one another, researchers can explore how sensitive a procedure is to a violation in an assumption. The challenge is to fully understand how close your data situation is to the dummy data sets studied and therefore the importance of the assumption to your analysis. In other words, a Monte Carlo simulation is basically the creation of many data sets using automated procedures to represent a particular data model. The data sets allow a researcher to conduct a risk or other similar analysis to identify generalizable results related to the data model. However, with small data sets, the limited information a small data set provides restricts the ability to identify the actual model the data represent. Because parametric statistics are constructed with assumptions and because each data set is unique, assuming your data set is one that conforms to these earlier simulation studies may place your findings at risk if you choose to ignore parametric assumptions.

Assumption 2. The Population Variances Are Equal (i.e., Homogeneity of Variance Among Groups is Present), in that the Scores in Different Groups Have Similar Variances

Variance is how far a group of numbers are spread. No variance means the data points are equal, a small variance means the data points are close together around the mean, and a large variance means they are spread out a good deal from one another. The equal variances assumption is important for parametric tests that investigate differences between or among groups within your targeted population. A commonly used example to explain equal variances is the testing of height differences between men and women. When examining differences between men and women, think of the distribution of height for men and women as separate groups. The variance of men's height is expected to be similar to the variance of women's height, thereby meeting the assumption of equal variances among groups.

A few parametric tests are appropriate for situations when the equal variances assumption is not met, but the equal variance assumption is not the only assumption for these tests (e.g., normality, sample size). Therefore, when conducting parametric tests to look for group differences, it is important to know if variances are different among groups. When your group variances are different, you cannot simply move ahead as if they meet the assumptions. Instead, you have two options. First, use a parametric test that allows for unequal variances while meeting all the other assumptions, or second, use a nonparametric test that does not assume equal or unequal variances. (For more information on testing for equal variances, see Elliott, 2015.)

Assumption 3. The Sample Is Adequately Large

Terms in the statistical literature like "adequately large" or "small" are essentially relative terms. For example, researchers working with large, secondary data sets may believe that 500 cases constitute a small sample, but many intervention researchers would find 500 cases to be an extraordinarily large sample. In practical applications, however, the conceptualization of sample size is often simply related to the statistical method being utilized. If the sample in use does not meet the rules for the statistical method a researcher wishes to use — for example, the accepted practice of 10 to 15 cases per variable in the case of an ordinary least squares (OLS) regression — then the sample becomes "too small."

In addition, as the sample size decreases, the likelihood of verifying that the assumptions for conducting parametric procedures are met also decreases. Remember that parametric procedures require data that are approximately normally distributed, and most texts recommend a minimum sample of 100 cases to investigate normality. Thus, you might infer that samples of fewer than 100 cases cannot adequately verify the normal distribution assumption, and nonparametric procedures must be used. Because most nonparametric texts suggest that samples of 30 cases or fewer must always utilize nonparametric procedures (Hoskins, n.d.; Mathews, 2010; Ryan, 2013), samples between 30 and 100 cases are unclear, and you must examine the variables in these samples carefully before proceeding.

Small samples are also problematic for parametric analyses because sample size is linked to the statistical power needed to make inferences about a population from your sample. Simply said, a small sample makes it more likely that you will not find a significant effect in your analysis, particularly if the effect size (e.g., of an intervention) is small. When faced with a small sample, you have two options. First, you can always use nonparametric procedures with small samples. Second, you can utilize a parametric test along with its parallel nonparametric test and include both in any report and in publication, explaining the strengths and limitations of the findings for each.

Assumption 4. A Linear Relationship Exists Between or Among the Variables

Linearity is an important assumption for parametric tests like correlation and regression. For example, obtaining a value for a correlation between two variables is accurate only if their relationship can be represented by a linear combination of parameters. In simple regression, two parameters called regression coefficients explain the relationship between a dependent variable (i.e., the outcome variable) and an independent variable (i.e., predictor variable). The two parameters (i.e., intercept and slope) can be used to graph a straight line, with the slope of the line being the correlation value representing the relationship between the dependent and independent variables. Of course, placing a line that best fits a scatter plot for two variables is always possible, but as the scatter plot reveals a departure from a straight line, the correlation value may no longer represent the relationship between the variables. The assumption of a linear relationship dictates that a straight line can represent the relationship.

Simple and multiple regression are parametric models that assume a linear relationship, and the models are used to predict values of a dependent variable based on values from a single independent variable or multiple independent variables. To make a prediction, the relationship between the independent and dependent variables must be clear. In other words, the independent variables in a statistical model like multiple regression must suggest only one value for a dependent variable. Having the linear relationship between the independent and dependent variables makes prediction possible.

In simple regression (i.e., regression with only one independent variable), the relationship between the dependent and independent variables can be seen by graphing a scatter plot of the values. However, in multiple regression (i.e., regression with multiple independent variables), verifying the linear relationship is a bit more complicated. In addition, adding higher power regression coefficients can model curvilinear relationships between the dependent and independent variables. (For more information on testing the linear relationship assumption, see StatSoft, 2000.)

Conducting a simple or multiple regression when the assumptions are not met has the potential for a misinterpretation of the relationship between the independent and dependent variable(s), or worse, you could identify a predicted value for the dependent variable that is truly unfounded. You have several options in this situation. One option is to attempt to transform one or a few of your variables to more closely resemble a normal distribution, even though the transformation does make explaining the variable relationships much more difficult. Another option is to use nonparametric or nonlinear regression, discussed in Chapter 5.

Assumption 5. Homoscedasticity is Present in Regression Models

Equivalent to the assumption about equal variances, an assumption of homoscedasticity is required when conducting parametric regression.

In simple terms, you are assuming that the data points for the relationship between the dependent and independent variable have similar variability along the entire range of the prediction line. For example, if you were to examine a scatter plot of the prediction line and data points, homoscedasticity is present if the points along the prediction line show a similar pattern along the entire range of the prediction line. A violation of this assumption (i.e., heteroscedasticity) results in some cases having more pull than others, as demonstrated by scatter plots that can sometimes look like a cone or a fan. Heteroscedasticity can result in invalid prediction values or inaccurate estimates of variable relationships. When faced with heteroscedasticity, you have two options. First, you can look for variables that are causing the heteroscedasticity and examine the cases that appear to be problematic to see if you can legitimately replace them. However, it is not appropriate to go “case shopping,” because any case that represents a real respondent cannot be excluded without cause. The other option is to use a nonparametric approach, such as nonparametric or nonlinear regression, as discussed in Chapter 5. (For more information on testing the homoscedasticity assumption, see Gaskin, 2015.)

Assumption 6. Residuals (Errors) Follow a Normal Distribution

When conducting regression, there is rarely, if ever, a time when the dependent and independent variables suggest a prediction line that matches perfectly with the data. The differences between the estimated prediction line and the data used to establish the line are called residuals or errors. These residual values help explain how well your prediction line matches the data. Larger residual values suggest less support for prediction values, while smaller values suggest more support for prediction values (i.e., poor prediction model versus better prediction model, respectively). Given this, the distribution of the residual values is assumed to follow a normal distribution. If they do not, the potential to make incorrect predictions or to draw incorrect conclusions about variable relationships is high. The non-normal residuals indicate that your model is not constructed appropriately for predicting the dependent variable. When faced with a non-normal residuals distribution, your options are to find the variable(s) that are causing the issue and determine if you can legitimately exclude them from the model, or to use a nonparametric process, such as nonparametric or nonlinear regression. (For more information on testing the normal residuals distribution assumption, see NIST/SEMATECH, 2012.)

OTHER ASSUMPTIONS

The assumptions considered next apply to both parametric and nonparametric statistical techniques, and are the product of a well-designed study. Violating any of the assumptions below weakens or possibly invalidates any conclusions you draw from your statistical analyses.

Assumption 7. The Data Are Sampled Randomly from a Defined Population — A Probability Sample

If you are planning a research study that will select respondents randomly from a population, you should spend a good deal of time identifying your target population and learning about all the potential differences and similarities within the population. You must establish clear inclusion and exclusion criteria so that your population is clearly defined. Next, you will probably decide to collect data from a sample of the population rather than the entire population, for one of any number of reasons — e.g., you do not have access or enough time or money to collect data from all the population members. Thus, the believability of your study and the conclusions you draw about the population will be based on your sample. Importantly, your sample must reflect the same differences and similarities you identified in the population. To accomplish this, you must utilize a probability sampling method that gives every person in your target population an equal opportunity to be in your sample. This reassures the consumers of your research that your sample is representative of the population.

You won't always be able to utilize a probability sampling strategy for your research, and sometimes your secondary data will not come from a random sample. In these situations, you must clearly understand and then explain that the analysis and findings are from a convenience sample (i.e., a nonrandom sample from a population). Convenience samples are commonplace in social science research, and especially when the population of interest is a clinical population. Nevertheless, incorrectly treating a convenience sample as a probability sample is extremely risky, as it undermines your ability to draw accurate conclusions from your data. Using a nonprobability sample must be justified by the research question and has to be explained carefully in any reports or publications as a potential limitation.

Assumption 8. The Data Consist of Independent Observations (Apart from Paired Data, but for Which the Pairs Must Be Independent)

Independent observations involve the assumption that the information that one respondent provides does not influence the information provided by any other respondent. In other words, one response from an individual does not have anything to do with responses from other individuals, except for what is being measured. Unfortunately, trying to anticipate all the possible threats to collecting independent observations is very challenging, but if this challenge is ignored, it can add damaging bias to your study. If unrecognized, it can have a potentially devastating effect on the accuracy of your findings. The independent observation assumption is related to your ability to understand fully the potential influences on the data you collect. The examples are endless and range from not being able to conduct an unbiased probability sample to the natural occurrence of subjects' falling into previously unrecognized groups. The most important step in verifying independent observations is to examine closely how the data were collected, both procedurally and theoretically.

If your observations are not independent and you can identify the dependence issue, one option is to use multilevel modeling, which provides a way to model a dependence among your observations (e.g., students within classrooms within schools). Unfortunately, most of these models require a large data set. Ignoring a recognized dependence problem is not an option in either parametric or nonparametric strategies.

Assumption 9. The Levels of Measurement Match Variable Design

Measurement level is an important issue for all statistical procedures (see Box 1.1, "Levels of Measurement"). This assumption requires you to identify the appropriate level of measurement accurately for each variable used in a statistical analysis. In social sciences, ratio data are

Box 1.1 Levels of Measurement

Nominal — Qualitative information (i.e., names or labels) that is often classified into dichotomous typologies (e.g., heads or tails, yes or no) or more than two (e.g., race, religion). The categories do not have an implied order. We can determine if cases are equal or different on the variable, but nothing else. Mode is the only central tendency measure.

Ordinal — Qualitative information that is rank ordered, indicating the direction of difference, but that does not have equal distances between the data points (e.g., the distance between strongly disagree, disagree, agree, and strongly agree). We can determine if cases are equal or not equal, and greater or less than on the variable only. Median and mode are measures of central tendency.

Interval — Information that encompasses data points with equal distances between them, so we can determine amount of difference, but without an arbitrary zero (e.g., people's scores on a summary measure). We can determine if cases are equal or not, greater or less than, and we can add or subtract. Mode, median, and arithmetic mean are measures of central tendency.

Ratio — Information that includes a meaningful, nonarbitrary zero and that is often used in age or counting (e.g., 2 is twice as many as 4; 50 years is five times older than 10 years). We can determine if cases are equal or not, greater or less than, and we can add, subtract, multiply, and divide. Mode, median, arithmetic mean, and geometric mean are measures of central tendency.

Measurement Levels in SPSS

Categorical

- Nominal — see under "Measurement Levels from Textbooks and Research Courses"
- Ordinal — see under "Measurement Levels from Textbooks and Research Courses"
- Single items with Likert-type response option scales with fewer than 6 points

Scale — *Any variables that can be treated as continuous, such as*

- Single items with Likert-type response option scales with 6 points or longer
- Multiple items in a measure, summed for the scale score
- Interval or ratio — see under "Measurement Levels from Textbooks and Research Courses"

uncommon, because they require a meaningful, nonarbitrary possible value of zero. However, even interval data are infrequent, as the data points must have equal distances between them. For scale variables, the equal-interval assumption is extremely important. (The term *equal interval* refers to the implied space between whole-number values of your variable.) For example, if the theoretical space between a value of 8 and 10 is not the same as the space between 20 and 22, you do not have a variable with equal intervals. Not having equal intervals means that your variable is not really a scale variable; rather, it is actually an ordinal variable.

Nevertheless, social science researchers often treat variables with Likert-like measures as if they were interval or scale measures, assuming that the ordered responses (e.g., never, sometimes, frequently, always) are equally separated. In other words, the space between "never" and "sometimes" is the same as the space between "sometimes" and "frequently." This can be problematic, because the researcher who designed the question may believe that the spaces between the response options are the same, but the respondents may believe otherwise. For a question to be reliable, each respondent must see the response options in the same way. In other words, if some of the respondents perceive equal spaces between the response options and other respondents do not, the study may not be capturing the information in the way that the researcher intended. The researcher may not recognize that the respondents are approaching the questions differently, leaving the potential for the researcher to report invalid findings.

Response coverage is another conceptual issue of importance that is related to level of measurement and to the normality assumption. Response coverage refers to whether you have enough responses for each response option in a question. For example, assume that 100 people respond to a 4-point Likert-type variable with response options of *strongly disagree, disagree, agree,* and *strongly agree.* If 37 people choose *strongly disagree* and 57 people choose *strongly agree*, you are left with only six responses total in the *disagree* and *agree* options. The information you have for *agree* and *disagree* does not support the use of the 4-point Likert-type response options scale, and you should not use the responses as a continuous variable. Instead, you may need to transform the responses into a dichotomous variable (i.e., variable with only two categories). Although you may need to choose a different statistical test, the validity of the transformed variable's contribution is greatly increased. These decisions can be challenging and you should not base them exclusively on mathematical guidelines. A solid conceptual understanding of what the variable represents is an important consideration in these situations.

In summary, researchers faced with measures that do not meet the requirements for parametric techniques should not use them anyway and simply hope that their results are meaningful. It is easy to do, because statistical software packages like SPSS assume that the variables are measured and defined accurately. Frankly, some journals will publish articles based on parametric analyses using inappropriate measurement levels for the variables, as the practice has become commonplace. Yet this practice diminishes the believability of the findings and weakens the reputation of social science research. Fortunately, nonparametric techniques offer a viable alternative analytic strategy. When researchers find themselves with nominal or ordinal data, including Likert-like measures with few response options, and they are inclined to utilize parametric techniques, a solid strategy would be to follow up the parametric analysis with the parallel nonparametric analysis to determine if meaningful differences exist between the two. Including both findings will serve to strengthen the science in a meaningful way.

ADVANTAGES AND DISADVANTAGES OF NONPARAMETRIC PROCEDURES

Nonparametric analytic techniques offer a number of important advantages.

- They have few, if any, assumptions.
- These techniques are useful for dealing with samples containing outliers, especially extreme outliers that cannot be logically removed, because nonparametric techniques reduce excessive influence of outliers.
- They reduce the effect of heterogeneity of variance.
- Some of the nonparametric tests provide an exact p-value, which is the probability a null hypothesis is true given the data used in the test.
- They can be used when (a) variable(s) violate the normality assumption, (b) transforming the variable(s) does not solve the

normality requirement, or (c) transforming the variable makes its interpretation unintelligible.
- They are useful for analysis of ordinal data (e.g., 5-point Likert-type responses).

So why should a researcher not always use nonparametrics techniques? Nonparametric analytic techniques are accompanied by several disadvantages.

- Most importantly, when data are normally distributed, the parallel parametric technique is more powerful. The parametric technique has a higher probability that the procedure will reveal significant findings — e.g., that two variables are associated with one another when they are actually associated. However, having more power justifies the use of parametric statistics only when all the assumptions for the parametric technique are met.
- Nonparametric procedures provide no estimates of variance and many times do not provide confidence intervals.
- Because nonparametric procedures frequently use rankings rather than the actual data, statistical results can sometimes be more difficult to interpret.
- Tied values can be problematic if too many of them exist in the data set, and software packages are not always straightforward in describing how they deal with ties. An example of a tie is when two respondents provide the same value or response to a scale-type question.
- Calculating power is less straightforward, requiring Monte Carlo simulation methods (Mumby, 2002). Monte Carlo simulations will allow the testing of different combinations of sample sizes, group member comparisons (if applicable), effect size, etc., all a priori to estimate the sensitivity of a future test to reject a null hypothesis given different data scenarios.

MISCONCEPTIONS ABOUT NONPARAMETRIC TESTS

Misconceptions are commonplace about the conditions in which nonparametric statistics should be applied and about the characteristics that set them apart from parametric procedures. Over the years, many people have disseminated the misconceptions, while others have refuted them (e.g., Hunter, 1993; Sawilowsky, 2005). However, the repudiations do not seem to have made much of an impact in correcting people's beliefs. We attribute this, in large part, to the lack of inclusion of nonparametric training in statistical courses. The confusion about these misconceptions makes it difficult to avoid making errors when choosing your statistical strategies. The list of misconceptions is quite long, but a few that seem to come up more frequently than others are as follows:

- *Parametric procedures are always more powerful than nonparametric procedures.* This is simply not true, especially when dealing with non-normal data.
- *Nonparametric statistics should be avoided because article reviewers do not understand these procedures.* Not true. Reviewers are selected by their expertise and ability to review article content. If you want your article to be accepted, using an appropriate procedure for your data will improve your odds greatly. Just be sure to justify why you chose the nonparametric test for your particular research question or hypothesis.
- *Nonparametric statistics should not be used on large data sets.* Untrue. The mathematical procedures for nonparametric statistics are just as valid for large data sets as for small data sets, but they are used more frequently with small data sets and data with non-normal distributions.
- *Unlike nonparametric tests, parametric tests are robust.* Again, not true. Robustness implies that violation of statistical assumptions is not problematic. In fact, neither nonparametric nor parametric tests are immune to violations of assumptions, which is why it is important to test assumptions before accepting any test's findings.

A FEW WORDS OF CAUTION

1. You should not get the mistaken impression that because nonparametric analytic techniques are available, it is always OK to conduct research with small samples. We urge you to

conduct an a priori power analysis when planning a study, so that all statistical techniques can be available for analyzing the data, and so you will be able to detect small or moderate effect sizes. Remember, parametric procedures used on large samples with normally distributed data that meet all the statistical assumptions are more powerful than nonparametric strategies. Nevertheless, nonparametric strategies are extremely important if data from your large sample do not meet the assumptions of parametric tests, particularly the normality assumption. Studies of human behavior, for example, often result in highly skewed or kurtotic data (e.g., your outcome variable has clear limits of detection or is censored, or your variables are ranked or ordinal level). In addition, you may run into situations you do not anticipate that result in small data sets (e.g., attrition from your study sample), so it is useful to have nonparametric strategies to use in these situations.

2. It can be easy to interpret p-values incorrectly, so it is important to cover this topic a bit more extensively here rather than to integrate it into the chapters. As defined, p-values are the probability of getting results *at least as extreme* as the results you observed, assuming that the null hypothesis (the hypothesis of no effect/no difference) is correct. In other words, p-values evaluate how well the sample data support the notion that the null hypothesis is true. Thus, a small p-value (typically less than .05) indicates that the sample provides sufficient evidence to reject the null hypothesis for the whole population. For example, assume that you conduct an intervention that found a p-value of .03. The proper interpretation would be that *if you assume the intervention had no effect, you'd obtain the observed difference (or more) in 3% of the cases, due to random sampling error.*

Nevertheless, the literature contains many misinterpretations of the p-value. It is *not* the probability of making an error by rejecting a true null hypothesis — a Type I error in traditional statistical language. Furthermore, p-values cannot indicate the likelihood that the null hypothesis is true or false. A small p-value can specify that your data are improbable if, indeed, the null were true, but it cannot indicate if the null is false, or if the null is true but the sample was unusual. Thus, you *cannot* say *if the null hypothesis is rejected, you have a 3% likelihood of making an error.*

The American Statistical Association arrived at six statements about *p*-values, as follows:

1. The *p*-value can indicate how incompatible the data are with a specified statistical model.
2. The *p*-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
3. Scientific conclusions and business or policy decisions should not be based only on whether a *p*-value passes a specific threshold.
4. Proper inference requires full reporting and transparency.
5. A *p*-value, or statistical significance, does not measure the size of an effect or the importance of an effect.
6. By itself, a *p*-value does not provide a good measure of evidence regarding a model or hypothesis. (Wasserman & Lazar, 2016)

It is also important that you choose a level of significance (alpha) for your threshold before you begin your analyses. Social scientists typically choose .05 as their threshold, but in nonparametric analyses it is sometimes reasonable to choose .10 instead (M. Hollander, personal communication, 2003), despite its tendency to provide an inflated experimentwise Type I error rate. This information is integrated into the chapters of this book.

CONVENTIONS UTILIZED IN THIS BOOK

- Each chapter begins with a research scenario, followed by research questions and hypotheses to illustrate each nonparametric technique.
- Levels of measurement are described in SPSS terms. Thus, *categorical* includes nominal and ordinal measurement, which includes dichotomous variables. Another term used to describe categorical variables is *discrete*, which represents nominal, ordinal, and count data. In addition, by our recommendation

and for use in this book, *categorical* also includes Likert-type response option variables with fewer than 6 points. *Scale* includes any variable that can be treated as continuous. A continuous variable is one that can take on any value within a particular range of values related to what the variable represents. In this book, scale variables include multiple items in a measure that are summed for the scale score, interval and ratio levels of measurement, and single items with Likert-type response option scales of 6 or more points.

- Notes are strategically placed for clarity and additional explanation.
- Variable names are capitalized (e.g., Stress, Burnout).
- Response options for an item/question are italicized (*low, medium, high*).
- SPSS labels (i.e., words or phrases in the SPSS window) are boldface (**Choose Tests**, **Fields**). When you see boldface words in SPSS, you can click on the words to make something happen (e.g., if you click on **Variable View,** you will find the list of variables).
- Window names are underlined (e.g., <u>One-Sample Nonparametric Tests</u>).
- "Select =>" indicates that you should click to change the appearance of a screen or open/close a window.
- "Click" indicates that you are choosing an option on the screen or highlighting an item on the screen.
-  is the icon representing a move arrow. When in SPSS, clicked variables can be moved in and out of a variable list.
- In SPSS, the icon for nominal level measures is 
- In SPSS, the icon for ordinal level measures is 
- In SPSS, the icon for scale level measures is 
- We use the language and symbols that SPSS uses.
- **Sig.** = significance level.
- **Adj. Sig**. = adjusted significance.
- Success = SPSS's term for differentiating the values of a dichotomous variable (e.g., if you have a 1 and a 0 as the values of a dichotomous variable, you must choose one of them to represent *success*). *Note:* The default for SPSS is to choose whatever value it encounters first as the *success* value. If you run

a success-type test on a categorical variable with three or more values, SPSS will choose the first value as *success* and the rest as *failure*. This becomes clearer as you work your way through the relevant chapters.

- We report the number of decimal places that are included in the SPSS output (typically three), rather than rounding values.

FIRST STEPS

If you plan to follow the steps in the chapters to analyze your own data, you will need to get your data into SPSS. Procedures for getting your data into SPSS are found in SPSS's help feature (click on the Help menu option). Once your data are in SPSS, the next step is to verify that all your variables are assigned to the appropriate level of measurement. Figure 1.1 displays the data view for the example data used in Chapter 2. Note that each variable is assigned one of the three levels of measurement values (e.g., nominal, ordinal, or scale).

Once your data are in SPSS, you recognize SPSS notation, and you are familiar with the key terms, you are ready to move on to Chapter 2.

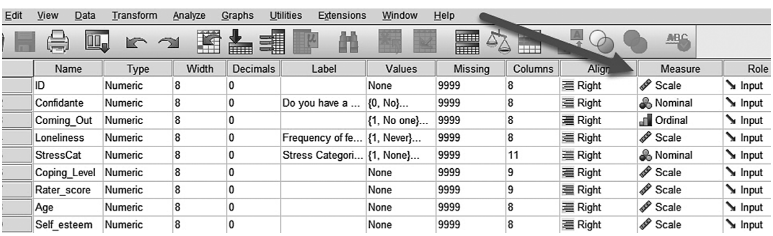


Figure 1.1. Data view within SPSS.

SUGGESTIONS FOR FURTHER READING

Manfred te Grotenhuis, Manfred, & Matthijssen, Anneke (2015). *Basic SPSS Tutorial*. Thousand Oaks, CA: Sage Publications.

Huff, D., & Geis, I. (1993). *How to Lie with Statistics*. NY: W. W. Norton & Company.

Pett, M. A. (2015). *Nonparametric Statistics for Health Care Research: Statistics for Small Samples and Unusual Distributions.* Thousand Oaks, Ca: Sage Publications.

Salkind, N. J. (2015). *Excel Statistics: A Quick Guide.* Thousand Oaks, CA: Sage Publications.

2

Analyzing Single Variables and Single Groups

CHAPTER OBJECTIVES

This chapter focuses on techniques for assessing single variables and single groups, organized by the variable's level of measurement. The following topics are covered:

- Examining variables for the quality of information they provide
- Dichotomous nominal variables — binomial test
- A nominal variable with three or more categories — chi-square test
- An ordinal variable with three or more categories — chi-square test
- A Likert-type variable treated as continuous — Kolmogorov-Smirnov test
- A summed scale variable — Wilcoxon signed-rank test

- A scale variable needing testing for randomness — runs test
- A single group, analyzing the association between two or more scale variables — Spearman's rho and Kendall's tau-b coefficients

INTRODUCTION

Chapter 1 describes the importance of understanding the data and variables prior to and while using them in statistical tests. This begins by first identifying the measurement level of the variables in the data set. The measurement level dictates which nonparametric procedures can be used to answer research questions. The second step is to review the quality of the information provided by each variable. For scale-type variables, the quality of information is related to capturing the true variability or distribution of the data. In addition, variables that represent concepts or ideas that have very little variability are not as informative and sometimes contribute less to a statistical model. For example, a variable that has very little variability (i.e., small variance) does not offer as much information about the characteristic the variable is attempting to capture as a variable that has more variability. In other words, in most cases, the more recorded variability, the more information available to answer research questions. This is not to say that variables with large variances are always preferred over variables with smaller variances, just that scale (i.e., continuous) variables with minimal variances do not contribute as much to the analysis. Therefore, before using any scale variable, you should review the underlying concept you are measuring to investigate if the observed variance has theoretical support, and if there is enough variance to warrant its inclusion in an analysis.

For categorical and dichotomous variables, the quality of information is related to coverage in the variable categories. The more coverage, the more information you have for the analyses. For example, if using a dichotomous variable for ethnicity (i.e., a nominal categorical variable) and the data set contains 30 cases (i.e., a small data set) with 27 Hispanic/Latino(a) people and three non-Hispanic/Latino(a) people, you have much more information for Hispanic people than for non-Hispanic people. In this situation, you must decide if you have enough information to consider differences between the two ethnicities. Similarly, consider a categorical variable that represents a 4-point Likert-type scale (i.e., an ordinal categorical variable) with response options *strongly disagree, disagree, agree,* and *strongly agree*. In this situation, if, for example, only two respondents selected *agree* and the other respondents were equally distributed among the three other response options, the variable offers very little to no information on people who *agree*. Therefore, one option is to review carefully the construct of the variable and to decide if collapsing the responses into *disagree* and *agree* has theoretical support. After creating the new collapsed variable, you should run the analysis twice — first on the noncollapsed variable and then on the collapsed variable — to identify how the results differ between the two versions of the variable. The difference between the two analyses will provide insight into how to proceed.

Software packages that facilitate the analysis of variables do not take into consideration the quality of information captured by a variable. Therefore, the responsibility falls to you, the researcher, to determine when the information provided by a variable is too limited to use in a statistical test. When using categorical variables (e.g., analyzing a 2 × 2 table that represents two dichotomous variables), one frequently used guideline (dating from 1952) dictates that a categorical variable should have a minimum of five respondents representing each category before using the variable in an analysis (Cochran, 1952). On face value, this makes sense because any category that has fewer than five responses does not have much representation within the variable. In the example above involving ethnicity, three non-Hispanic people do not provide enough information to effectively use the variable Ethnicity in a statistical test. Categorical variables with a small number of respondents for a category are especially problematic for parametric procedures involving group comparisons (e.g., *t*-tests, ANOVA). The lack of information limits the ability to verify that the data meet the assumptions for parametric procedures. However, the lack of abundant information for non-Hispanic people is not as problematic for nonparametric procedures, because nonparametric procedures use ranks and order to test hypotheses. Therefore, the likelihood of a Type I error (reject a true null hypothesis) is much lower when conducting nonparametric tests. This is not to say that using nonparametric procedures releases a researcher from fully understanding the information quality within each variable. A thorough review of the variable characteristics is still the first and most important step in analyzing data.

You have two major reasons for looking at variables individually. The first reason is understanding the quality of information captured by variables in a data set. The second reason relates to investigating research questions that can be answered with a single variable. Multiple nonparametric procedures provide the opportunity to run statistical tests on a single variable.

For most people learning how to use statistics, it is often helpful to provide a framework or context for the discussion. This is especially true when the discussion is your initial introduction to new statistical procedures. Therefore, each chapter begins with a scenario describing a hypothetical research line of study that can be explored with research questions.

Box 2.1 Research Scenario for Chapter 2

Dr. Chonody has just completed a study investigating the coming out process for people who are gender and sexually diverse. She conducted extensive, semi-structured interviews with 29 LGBTQ youth who were recruited through websites and newspaper advertisements. Dr. Chonody used grounded theory and narrative approaches, audio taping the interviews. The tapes were transcribed and then analyzed using a combination of grounded theory and cross-case procedures. She ensured the quality of the data through peer debriefing, independent coding into key categories, and member checking. After analyzing the qualitative data, Dr. Chonody asked the respondents to fill out a paper-and-pencil questionnaire that contained several validated measures that could confirm or extend her findings. In combination, the qualitative and quantitative data provide improved validity and a more complete picture of the respondents. Dr. Chonody has hired you as her statistical consultant. Her research questions follow.

Six research questions guide the presentation of nonparametric procedures related to the above scenario. Each research question describes the reasoning for using nonparametric procedures. Table 2.1 lists the nonparametric procedures presented in Chapter 2, the types of variables for each procedure, and the names of the variables found in the database for Chapter 2, which is available online.

Table 2.1. Summary of Analyses and Variables in Chapter 2

Research Question #	*Nonparametric Test*	*Variable Measure*	*Variable Name in DB*
1	Binomial	Dichotomous or scale	Confidante
2	Chi-square	Nominal or ordinal	Coming_out
3	Kolmogorov-Smirnov	Scale	Loneliness
4	Wilcoxon signed-rank	Scale	Coping_level
5	Runs	Dichotomous, nominal, ordinal, or scale	Confidante and Rater_score
6	Spearman's rho and Kendall's tau-b coefficients	Scale	Age, Self_esteem, and Coming_Out

Box 2.2 Research Question #1

After analyzing her qualitative data, Dr. Chonody believed that only 60% of the youths had a confidante with whom to discuss the coming out process. However, she did not ask that question explicitly in her interviews — she simply inferred this from her qualitative data. Dr. Chonody examined this question explicitly by asking it on her follow-up questionnaire. After she collects the data from the follow-up questionnaire, Dr. Chonody asks you to test her inferred percentage.

Research Question #1

What proportion of youths have a confidante?

BINOMIAL TEST

The binomial test is used for testing the statistical significance of possible differences between an expected proportion and an observed proportion. The binomial test is called an exact test because it provides the exact probability that the data represent the stated expected proportion. A comparable parametric test is a *t*-test that uses *z* scores to test a hypothesis of an expected proportion, but it is used rarely for proportion testing outside of introductory statistics courses. The assumptions associated with the binomial test are that the variable used in the test is dichotomous — two categories representing *success* and *failure* — or can represent a dichotomous situation by identifying what values represent *success* and what values represent *failure* with at least five values for each of the two categories. The strengths of the binomial test are that it provides the exact probability and, unlike the *t*-test, it does not assume any associated distribution. However, like many nonparametric procedures, the binomial test has less power to find statistical significance compared to the other tests with an assumed distribution (if the data meet the distribution assumption).

Other examples of binomial tests are as follows:

1. You recently opened a new health clinic, and you know the percentage of people with STDs in your city, so you want to test if the proportion of people coming to your new clinic is similar to the city percentage or significantly different.
2. A treatment center's administrators believe that 90% of their patients are satisfied with their care, and you suspect that the administrators are overestimating the satisfaction level.

Prior to conducting any hypothesis test, you must examine the information provided by the variables. For Research Question #1 (see Box 2.2), the variable is dichotomous, representing two nominal categories — having a confidante or not having a confidante. For dichotomous variables, the easiest way to examine the variable information is to run a frequency of the responses.

SPSS Process

To begin your analysis

- □ Select **Analyze** => **Descriptive Statistics** => **Frequencies** to open the Frequencies window.
- □ Click to select the variable Confidante in the list of variables on the right (i.e., Do you have a confidante?).
- □ Click the move arrow to move the variable name over to **Variable(s):** area.

Because the interest is only in frequencies, leave all the options for the buttons on the right in their default position.

☐ Make sure the **Display frequency tables** option at the bottom of the Frequencies window is selected.
☐ Select => **OK** at the bottom of the window.

SPSS Output

Two tables will appear in the IBM SPSS Statistical Viewer window (i.e., analysis output window).

- The first table shows the number of valid and missing cases found in the data set. Use this table to verify that the table matches the number of cases intended to be part of the analysis.
- The second table shows the frequency at which the 28 youths report having a confidante (see Table 2.2 — Frequency Table for Research Question #1).

Table 2.2. Frequency Table for Research Question #1, Do you have a confidante?

		Frequency	*Percent*	*Valid Percent*	*Cumulative Percent*
Valid	No	8	27.6	27.6	27.6
	Yes	21	72.4	72.4	100.0
	Total	29	100.0	100.0	

Table 2.2 shows that 21 out of the 29, or 72.4% of the youths, have a confidante. Therefore, do the data collected in the follow-up support Dr. Chonody's inference that 60% of participants had a confidante? At first glance, it appears that they do not (60% is too low), but a rigorous researcher will test further to be sure.

You might be tempted to utilize a parametric test like a *t*-test using *z* scores, despite a sample size of fewer than 30. If you run a one-sample *t*-test in SPSS (i.e., **Analyze** => **Compare Means** => **One Sample T-test**. . .), the 95% confidence interval for the percentage is .55 to .90 (55% to 90%). This would support Dr. Chonody's inference that 60% of the youths have confidantes. However, because the sample has fewer than the sample-size guideline, you should use a nonparametric procedure called a binomial test, which is well suited for this analysis because of the ranking approach used in testing hypotheses.

SPSS Process

To begin this analysis,

- ☐ Select => **Analyze** => **Nonparametric Tests** => **One Sample**.

Once the One-Sample Nonparametric Tests window opens, SPSS needs to know which variable to analyze.

- ☐ Select => **Fields** tab.
- ☐ Click to select the **Use custom field assignments** option.
- ☐ Click the variable you want to examine by highlighting Do you have a confidante?
- ☐ Click the move arrow to move the variable over to the **Test Fields:** area.

Note: Depending on how you selected the last nonparametric test options, some or all the available variables may show up in the **Test Fields:** area on the right side of the window. If this is the case, simply select all the variables except the Confidante variable in the **Test Fields:** area and then use the move arrow to move them to the **Fields:** area on the left side of the window.

- ☐ Select => **Settings** tab to select test options and which nonparametric test to run.

The **Settings** tab in the One-Sample Nonparametric Tests window is where you can select which nonparametric test to run, where you can set different test options, and where you set how you want to handle missingness. Selecting the different items listed in the **Select an item:** area gives you access to these choices.

- ☐ Select => **Test Options** in the **Select an item:** area if you wish to change the significance level from the default (0.05) for a test, or if you wish to change the confidence interval range from the default (95.0%). Here you can also change the missing data default (exclude cases test-by-test).
- ☐ Select => **Choose Tests** from the **Select an item:** area to see two options for choosing tests.

The first option (**Automatically choose the test based on the data**) allows SPSS to decide automatically which tests are appropriate based on the recorded SPSS variable definition (i.e., defined level of measurement for each variable). Because it is important to follow the analysis process closely to understand fully how the analysis is performed, click to select the **Customize tests** option to gain access to the test selection boxes.

☐ Click the first option in the list (i.e., binomial test). Figure 2.1 shows the One-Sample Nonparametric Tests window with the test selected that matches Research Question #1.

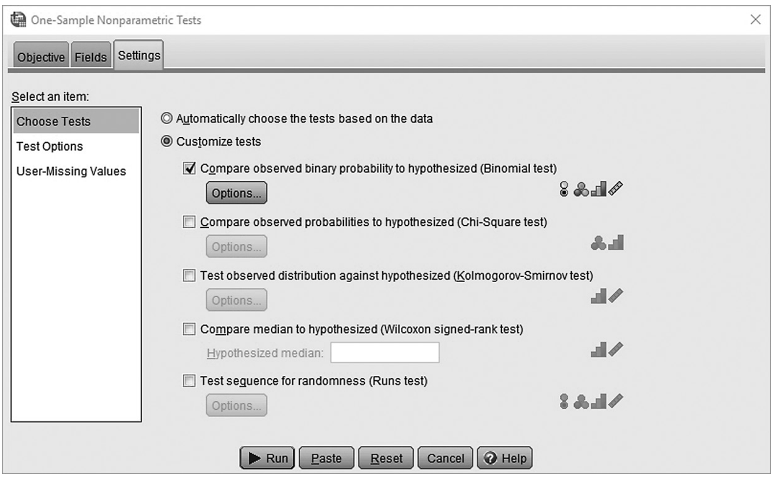


Figure 2.1. One-Sample Nonparametric Tests window.

Shown in Figure 2.1 is the list of different tests available for analyzing a single variable. The icons shown on the right side of the screen represent the tests that are appropriate for each measurement level. For Research Question #1, the binomial test is appropriate because it compares a hypothesized proportion of responses to the observed proportion of responses in the variable categories. For a dichotomous variable, the default is that half of the respondents selected one category while the other half selected the other category (i.e., successes are as equally likely to occur as fails, or 50% success and 50% fails). However, Research Question #1 proposes a 60% rate for *yes* responses, so

- ☐ To change the default, select => **Options**. . . to open the Binomial Options window.

The Binomial Options window provides an opportunity to change the expected or hypothesized proportion, request confidence intervals, identify which category value to use for the hypothesized proportion, and define how to group low vs. high values if analyzing a scale variable.

Binomial Options Decisions

1. **Hypothesized Proportion** — You can identify any hypothesized proportion value between 0 and 1 to test against your data. The SPSS default hypothesized proportion value is 0.5.
2. **Confidence Interval** — SPSS provides the option of three different calculation methods to estimate a confidence interval for likely proportions based on the data. You may have an occasion in which obtaining a confidence interval would be helpful (e.g., wanting to identify a plausible percentage range of youths with confidantes). When you do, select all three options and compare their results.
3. **Define Success for Categorical Fields** — SPSS uses the terms *success* and *failure* to identify values in binomial tests. If a value for *success* is not provided, SPSS will use the first value it finds in the variable as *success*. Managing what values are designated as *success* and *failure* will help in explaining your results.
4. **Define Success for Continuous Fields** — The binomial test is very flexible, and SPSS provides a way to conduct the test using a scale variable. This option decision allows you to identify a range of values that represent *success* and values that represent *failure*.

For Research Question #1, 0.60 is used as the hypothesized proportion and 1 as the value for *success* (i.e., Yes — youth has a confidante).

- ☐ Click in the **Hypothesized proportion:** area and change the value from .5 to .6.
- ☐ Click all three options under **Confidence Interval** (i.e., Clopper-Pearson, Jeffreys, and Likelihood ratio) so you can compare the

results from the three different ways of calculating confidence intervals.

- ☐ Click the **Specify success values** option for **Define Success for Categorical Fields.**
- ☐ Click in the **Success Values:** white area and enter a 1 (to indicate 1 as the *success* value).
- ☐ After you review your choices, select => **OK** to close the window.
- ☐ Select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

A variety of information shows up on the output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., active data set)
- Hypothesis Test Summary (i.e., link to Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click the Hypothesis Test Summary information (i.e., point the mouse within the Hypothesis Test Summary table and double-click).

The Model Viewer window will open to display the detailed information (see Figure 2.2).

The Model Viewer window has access to more information than what is shown in Figure 2.2.

- At the bottom of the window are dropdown lists on the left for **View** and **Field Filter**, and on the right for **Test**, **Field(s)**, and **View**.
- Because Research Question #1 involves only one variable (i.e., you moved only one variable to **Test Fields:**), some of the dropdowns do not offer multiple options. Changing the dropdown selection for **View:** on the left from **Hypothesis**

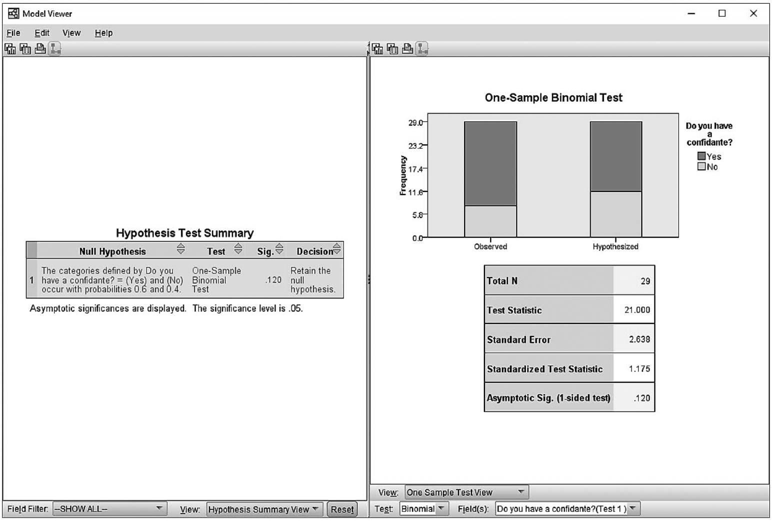


Figure 2.2. Model Viewer window with detailed test results.

> **Summary View** to **Confidence Interval Summary View** displays the lower and upper values for the three selected confidence interval calculation methods.

Findings

When you set the **View** in the Model Viewer window to **Hypothesis Summary View,** it displays the results of the hypothesized 60%. The **Sig.** value is .120, which fails to reject the hypothesis that the proportions of with and without confidantes are .60 and .40, respectively. Whenever you run a statistical test, you want to verify that your settings in SPSS were as you intended. You can verify your settings by closely reviewing the wording in the tables, values of the confidence intervals, the look of the histograms, and the information below the histograms, etc. You can find all this information by selecting the different dropdown options in the Model Viewer.

Next, select the dropdown for **View:** on the left side of the window and select **Confidence Interval Summary View**. The values under the Lower and Upper columns show the boundary values for the 95% confidence interval for the percent *success* identified in this data set. The calculation methods are slightly different among the three, which is why the values are slightly different.

In statistical terms, you can say a binomial test of the null hypothesis that the percent of youths with confidantes is equal to 60% fails to be rejected at an alpha level of .05 with a *p*-value of .120. Another explanation from the results is that if the youths responding to the follow-up survey are typical for gay youths, then there is a 12% chance that the proportion of youths who have confidantes is 60%. A Clopper-Pearson calculation provides a 95% confidence interval for proportion rate of 52.8% to 87.3%. In this analysis, selecting 95% for the confidence interval provides a range of possible percentages of youths with a confidante in which we are 95% confident that the true percentage falls. If 90% was selected, the range would be a bit wider, but the confidence that the true percentage falls within the range would be smaller. Note that this finding differs from that of the (inappropriate) parametric *t*-test mentioned earlier that found lower and upper bounds of 55% to 90%.

Practically, the findings suggest that, for this study, the number of gay youths with a confidante is a bit higher than the number without a confidante. For example, if you had a group of 100 gay youths who were similar to the youths in Dr. Chonody's study, the number of youths with a confidante would be somewhere between 53 and 87.

Box 2.3 Research Question #2

In her qualitative data, Dr. Chonody did not see a general pattern in the degree to which her participants had come out, but she wanted to test this with a survey question asking participants the number of people to whom they had disclosed their sexual orientation. After reviewing this variable's distribution, Dr. Chonody found that the responses fell into four main categories — coming out to no one, to one person, to a few people, and to everyone. However, she did not have enough responses in the two middle categories (i.e., fewer than five responses in "to one person" and "to a few people" categories), so she collapsed those responses into one category. She now has a variable with three categories of responses — coming out to no one, to one or a few people, and to everyone. Dr. Chonody expects no differences in these three categories, but she wants you to test this statistically.

Research Question #2

Are the proportions equivalent among youths coming out to a single confidante, to multiple confidantes, or not coming out at all?

CHI-SQUARE TEST

Different kinds of chi-square tests address different data analysis situations (e.g., 2 × 2 contingency table analysis, Pearson chi-square, likelihood-ratio chi-square, etc.). Basically, the chi-square test is suited for testing possible proportion differences between an expected proportion and an observed proportion when more than two combinations of categories are possible. When proportions among categories are examined, this is similar to investigating whether the numbers within each category are similar. The kind of variable used in a chi-square test in this situation is either a nominal or an ordinal variable (i.e., categorical variable representing two or more responses — e.g., *none, sometimes,* or *often* for each respondent). The assumptions associated with the chi-square test are that the variables used in the test are categorical and that responses are independent. In addition, each category should be represented with at least five responses to allow for appropriate information coverage for the respondent categories. The strengths of the chi-square test are that the results are easy to interpret and you can apply the test to all kinds of categorical variables. One limitation is that the chi-square test is much more sensitive to sample sizes and often requires a larger sample to find significant proportional differences; however, if the sample is too large, the test finds significance too easily. Lastly, no alternative parametric procedure is possible without manipulating the data into something testable with a *t*-test.

Other uses for the chi-square test are as follows:

1. You are examining the data from a pilot study, and you want to investigate the proportions of people subscribing to different religions (i.e., the expected vs. observed).
2. In your small sample of gay youths, you wish to explore your suspicion that only 10% have come out to no one and the other 90% are evenly distributed between coming out to one or a few and coming out to everyone.

The binomial test investigates dichotomous variables, but the chi-square test examines if the response proportions are equal across three or more categories, which you need to answer Research Question #2. The SPSS steps, however, are similar. For Research Question #2, the variable Coming_Out is operationalized with three categories (1 = *no one*, 2 = *one or a few*, and 3 = *everyone*).

SPSS Process

To begin your analysis,

- ☐ Select => **Analyze** => **Nonparametric Tests** => **One Sample** to open the One-Sample Nonparametric Tests window.

When the One-Sample Nonparametric Tests window opens, SPSS needs to know which variables to use in the analysis and which analysis routines to run.

- ☐ Select => **Fields.**
- ☐ Click to select **Use custom field assignments** option.
- ☐ Click to select Coming_Out from the list of variables in the **Fields:** area.
- ☐ Click the move arrow to move Coming_Out over to **Test Fields:** area.

Note: If other variables are listed in the **Test Fields:** area, select them and move them back to the **Fields:** area on the left side of the window.

Now that SPSS knows what variables to use,

- ☐ Select => **Settings** tab to see the list of available nonparametric tests.
- ☐ Select => **Choose Tests** in the **Select an item:** area.
- ☐ Click to select **Customize tests** and then click to select the checkbox for the **Chi-Square test.**

In Figure 2.1, the variable measure icons indicate the variables with the appropriate level of measurement. For the chi-cquare test, the levels of measurement are nominal and ordinal, and the icons change to color when you check the **Chi-Square Test** checkbox.

The options for the chi-square test are different from the binomial test. The only time a change to the default options is necessary is when you are testing the category proportions for something other than being equal across all categories. Because Research Question #2 is testing for equal proportions, the chi-square test options can remain as the default (i.e., all categories have equal probability).

- □ After carefully reviewing your choices, select => **Run** to conduct the analysis

See Appendix A for complete SPSS syntax.

SPSS Output

To see the details of the test, double-click the **Hypothesis Test Summary** information (i.e., point the mouse within the **Hypothesis Test Summary** table and double-click). The Model Viewer window will open to display the detailed information.

Findings

The null hypothesis for the chi-square test for Coming_Out is that equal proportions of responses exist for the three categories. When the Model Viewer window first opens, it presents a histogram on the right and the test results on the left. The **Sig.** value for Coming_Out indicates that the test fails to reject the null hypothesis (i.e., **Sig.** value = .185 is larger than the default alpha level of .05), thereby suggesting that there is not enough information to identify a proportional difference among youths' Coming_Out categories. Another explanation is that the proportions are similar enough to not be identified statistically as significantly different. In statistical terms, you may say that the chi-square test for a null hypothesis of equal proportion of youth Coming_Out categories fails to be rejected at $\alpha = .05$, with $p = .185$.

However, when you examine the histograms for Coming_Out among gay youth categories (i.e., *no one, one or a few*, and *everyone*), they seem to show possible differences despite not being significantly different, statistically. More than likely, if the sample size were larger and followed similar proportions, there would be enough power to identify differences. This is one important reason always to include as many respondents in your study as possible, to increase the likelihood of finding statistical differences. In this situation, you do not have enough information from the variables to identify statistical significance; that is, the statistics are keeping you from drawing a potentially wrong conclusion based on viewing the histograms only.

Box 2.4 Research Question #3

Dr. Chonody believed, after reviewing the qualitative data, that her respondents' feelings of loneliness may have important relationships with other issues in their lives. To investigate this possibility, one of the questions on her follow-up survey asked respondents how frequently they feel lonely. The response options were never, sometimes, often, and all the time. Dr. Chonody wants to use the responses to this question as a continuous variable for further analysis, but she isn't sure if this ordinal-level variable sufficiently approximates a normal distribution, so she asks you to test this for her.

Research Question #3

Should a categorical Likert-type variable with four response options representing levels of loneliness be treated as a continuous variable representing a normal distribution in a statistical analysis?

KOLMOGOROV-SMIRNOV TEST

The Kolmogorov-Smirnov test examines how well a distribution from a scale variable matches a probability distribution. In SPSS's menu structure for this nonparametric procedure, you can test the variable against four different types of distribution — normal, uniform, exponential, and Poisson. The normal distribution is most often selected because of the applications of the normal distribution to parametric assumptions. On the parametric side of SPSS's menu structure, the option to use the Kolmogorov-Smirnov test exists only for the normal distribution (**Normality plots with tests** option within the Explore: Plots window). You can consider the Kolmogorov-Smirnov test a type of goodness-of-fit test, because it examines how well the data conform to the selected distribution. For information on distributions, see http://pages.stern.nyu.edu/~adamodar/New_Home_Page/StatFile/statdistns.htm. (Be sure not to include the concluding period in the web link.) The kind of variable used in a Kolmogorov-Smirnov test is a scale variable, a continuous variable representing responses along a range of potential values — e.g., Age, Burnout level, etc. If you want to examine a Likert-like variable to see if it follows a normal distribution, you will need to redefine its level of measurement to scale before attempting to perform the test.

Other applications for the Kolmogorov-Smirnov test are as follows:

1. From your pilot study, you wish to analyze a scale/continuous variable to determine if it is represented by any of four distributions (i.e., normal, uniform, exponential, or Poisson).
2. You have conducted a pilot for a larger survey, and you wish to test if the ages of respondents follow a normal distribution with a mean of *x* and a standard deviation of *y*.

Use great caution when deciding to treat a Likert-like categorical variable as continuous. A continuous variable has the characteristic that all values are possible between each response point and within a specific range supported by the variable's conceptual design. Altering the theoretical construct of a categorical variable should not be taken lightly and must be preceded by a close review to verify that the change does not disrupt how the variable represents the initial concept.

One important component of a close review includes the examination of the potential distributions of the categorical variable. The Kolmogorov-Smirnov test offers a convenient way to look at a variable's distribution. However, the Kolmogorov-Smirnov test will work only with scale variables. Therefore, for Research Question #3, you must change the definition for the variable How frequently do you feel lonely? temporarily from ordinal to scale within SPSS's variable view area.

SPSS Process

To change the level of measurement for a variable

- □ Select => **Variable View** tab at the bottom of the IBM SPSS Statistics Data Editor window.
- □ Find the **Measure** column and the row for Loneliness, and then click on the Ordinal icon to show the dropdown list.
- □ Select Scale. (*Note:* If Scale is already selected, then you have no need to make a measure change.)

To begin the analysis,

- ☐ Select=> **Analyze** => **Nonparametric Tests** => **One Sample**.
- ☐ Select => **Fields** tab.
- ☐ Click to select **Use custom field assignments** option.
- ☐ Click to highlight the Loneliness variable and then click the move arrow to move it to **Test Fields:** area.

Note: Another method for moving variables from one side of the window to the other side is to point your mouse at the variable name and double-click. SPSS interprets the double-click as a message to move the variable to opposite area.

Now that SPSS knows what variable to use,

- ☐ Select => **Settings** tab to see the list of available nonparametric tests.
- ☐ Select => **Choose Tests** to highlight it in the **Select an item:** area.
- ☐ Click to select the **Customize tests** option and then click to select the checkbox for the **Kolmogorov-Smirnov test** (uncheck other **Customize tests** options if checked).
- ☐ Select => **Options**. . . to open the Kolmogorov-Smirnov Test Options window. The options allow you to test your variable against four different hypothesized distributions.
- ☐ Click to select all four distributions (i.e., **Normal, Uniform, Exponential**, and **Poisson**) to compare all four potential distributions rather than just testing for normality.
- ☐ For Research Question #3, use the default (i.e., the **Use sample data/Sample mean**) for each of the four distribution mean and parameters options.

Note: The decision to use the default is based on the details in Research Question #3, but in other situations the default may not be appropriate. For example, if the desire was to investigate if a Likert-like variable can be from a normal distribution with a mean of 5 and a standard deviation of 1, then changing the **Distribution Parameters** for the **Normal** option is accomplished by clicking on **Custom** and entering the **Mean** and **Std Dev.** values.

- ☐ Select => **OK** at the bottom to close the Kolmogorov-Smirnov Test Options window.

- □ After reviewing your selections, select => **Run** at the bottom of the window to begin the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

Shortly after selecting **Run,** SPSS's output window will display and show the **Hypothesis Test Summary** results for the analysis. Because all four distributions were selected to test against the distribution for the Loneliness variable, four results are presented in the summary.

- To see the details of the test, double-click the **Hypothesis Test Summary** information (i.e., point the mouse within the **Hypothesis Test Summary** table and double-click).
- The <u>Model Viewer</u> window will open to display the detailed information.
- When the <u>Model Viewer</u> window opens, the first test for the normal distribution is highlighted in the list of tests shown on the left side of the window, and the corresponding test information is displayed on the right. Some of the corresponding test information is a histogram with an overlay of a normal distribution using the sample data to establish parameter values.

Findings

According to the Kolmogorov-Smirnov test, there is sufficient information to conclude that the data for the Loneliness variable do not come from a normal distribution. A review of the graph clearly shows a highly right-skewed distribution for Loneliness. Recall, one of the challenges of using nonparametric procedures is lower power in rejecting the null hypothesis, but in this case, it does not seem to be even close with a **Sig.** value less than .001. Even so, the decision to use an ordinal variable as continuous (i.e., scale) should not be based on a single statistical test alone.

Clicking on the other test summaries on the left side of the window (i.e., clicking within the **Hypothesis Test Summary** table) shows the corresponding information for each test. A review of the histograms and the hypothesized distribution shows the level of conformity of the Loneliness data to the uniform, Poisson, and exponential distributions. The Kolmogorov-Smirnov test found that the Loneliness data did not reflect uniform or exponential data (i.e., **Sig.** values less than or equal to .001), but did suggest a distribution that followed a Poisson distribution (i.e., **Sig.** = .836). In more statistical terms, the test found that there is less than a 5% chance that the Loneliness data follow a normal distribution if the null hypothesis is true.

For Research Question #3, although the Kolmogorov-Smirnov test did reject the normal distribution, a review of the histograms for Loneliness provides enough information that you should not use it as a continuous variable with an assumption of normality. However, if you had an opportunity to use Loneliness in a statistical model in which it would represent a Poisson distribution, it would be acceptable.

Practically, you should take great care when choosing to use an ordinal variable with Likert-like response options as continuous. If you have any doubts about the ordinal variable's ability to represent a normally distributed set of values, then do not use the ordinal variable in a test that assumes normality.

Box 2.5 Research Question #4

Based on her interviews, Dr. Chonody believes that her respondents have an average ability to cope with stressful events (e.g., coming out). To test this hypothesis, she investigates her respondents' scores on a validated measure of coping with stressful events. An average score would be 15 out of 30 possible, and Dr. Chonody would like you to test the likelihood that a score of 15 represents her respondents' median score.

Research Question #4

Is the median score 15 for these respondents' ability to cope with stressful events?

WILCOXON SIGNED-RANK TEST

The Wilcoxon signed-rank test is used for many different tests, but in SPSS under the One-Sample Nonparametric Tests, the Wilcoxon signed-rank test examines the probability that a specified median value is possible for a target population given the available data from a single continuous variable. Therefore, the Wilcoxon signed-rank test examines the likelihood that a specific value represents a median value for the data.

When examining small samples, the use of a median value is often preferred over the use of a mean value because a single response can change the mean score dramatically in a small sample (i.e., an extreme score can overly influence a mean calculation). When outliers are present, using a median value is preferred because this removes the outliers' excessive influence on the analysis. However, if no outliers are present, the mean is the preferred choice.

The parametric *t*-test is similar to the Wilcoxon signed-rank test, except that the *t*-test uses a specified mean value rather than a median value. The kind of variable used in a Wilcoxon signed-rank test is a scale/continuous variable representing responses along a range of potential values — e.g., Age, Burnout level — based on a summed scale score from multiple items in the measure, etc. The Wilcoxon signed-rank test is the test to replace the parametric *t*-test when you cannot verify the normality assumption. However, like many nonparametric procedures, the Wilcoxon signed-rank test has less power to find statistical significance than the *t*-test when the data meet the normality assumption for the *t*-test.

Other uses for the Wilcoxon signed-rank test are as follows:

1. You are conducting a small study of a new weight-loss intervention, and you want to determine if the average weight loss is equal to or greater than 5 pounds.
2. Before starting a new therapy group, you wish to learn if the median level of self-esteem for your future clients is similar to, or different from, that of people in general.

For Research Question #4, the continuous variable from the data set is Coping_Level. The Coping_Level variable in the data set represents a scale score from a validated measure that examines coping levels with stressful events. Twenty-nine youths completed the validated measure, and their responses were used to set the values for Coping_Level.

SPSS Process

To begin your analysis

- □ Select => **Analyze** => **Nonparametric Tests** => **One Sample**.
- □ Select => **Fields** tab.

- ☐ Click to select **Use custom field assignments** option.
- ☐ Click to highlight the Coping_Level variable in the list of variables in the **Field:** area.
- ☐ Click the move arrow to move it over to **Test Fields:** area.

Now that SPSS knows what variable to use,

- ☐ Select => **Settings** tab to see the list of available nonparametric tests.
- ☐ Select => **Choose Tests** in the **Select an item:** area.
- ☐ Click to select the **Customize tests** option to gain access to the list of tests.
- ☐ Click the checkbox for the **Wilcoxon signed-rank** test.
- ☐ Click in the **Hypothesized median:** area to place focus on the white space, and enter the number 15.

Unlike the other test examples, the Wilcoxon signed-rank test does not have any options — it requires only the identification of the hypothesized median value. The level of measurement icon to the right indicates that this test is only for interval or scale variables. Therefore, if the variable is nominal, testing for a hypothesized median value is not appropriate.

- ☐ After carefully reviewing your choices, select => **Run** to begin the analysis

See Appendix A for complete SPSS syntax.

SPSS Output

To see the details of the test, double-click the **Hypothesis Test Summary** information (i.e., point the mouse within the **Hypothesis Test Summary** table and double-click). The Model Viewer window will open to display the detailed information.

Findings

The histogram displayed on the right side of the Model Viewer window shows the relative distribution of the Coping_Level variable, the hypothetical median used in the Wilcoxon signed-rank test, and the observed median. A review of this histogram does not show any significant difference between the hypothetical and observed median, which indicates that no information supports the belief that the youths as a group have a different expected level of coping with stressful situations (i.e., **Sig.** = .537, which is larger than the $\alpha = .05$ threshold).

In statistical terms, the Wilcoxon signed-rank test fails to reject the null hypothesis that the median value for Coping_Level is 15 ($p = .537$) at $\alpha = .05$.

Practically, the only way to know for sure that a set of values does not have a specific median value is to have enough evidence to be certain to think otherwise. A statistical test is one method that contributes to identifying what "enough evidence" means. In this situation, the data show that a median value of 15 is possible.

Box 2.6 Research Question #5

Dr. Chonody had four experts in this substantive area review her semi-structured interview questions before she began her study, asking them to rate each question on a 1 to 10 scale for how well it addressed an important issue. When she collated the experts' ratings, she noticed what she thought was a pattern in one of the rater's scores. To test this possibility, she wants you to conduct a runs test on the rater's sum scale scores. Of course, Dr. Chonody placed the questions randomly in a list prior to expert rating to minimize the possibility of question quality order.

Research Question #5

Are the values in the experts' ratings showing a pattern (i.e., not random)?

RUNS TEST

The runs test, also called the Wald-Wolfowitz test, provides an opportunity to see if the values in a variable are random. The runs test in SPSS is constructed around a dichotomous variable, but SPSS can examine other levels of measurement variables for randomness by identifying two groups within the data. Randomness associated with testing for runs is the hypothesis that the order in which the two values in a dichotomous variable appear does not show any kind of pattern (i.e., the two values in the dichotomous variable appear at random). In statistics, many different procedures can test for randomness in all different kinds of variables, because knowing when a sequence of numbers is random or dependent is extremely important when analyzing data. However, in SPSS the runs test is the only procedure that directly tests for randomness in a variable.

Other examples of a Runs test are as follows:

1. A supervisor hopes that her caseworkers' client satisfaction surveys vary in a random way. However, she suspects that they are more negative at the end of the day, or that they are more positive based on the specific work group.
2. The administrator of a small organization wants to know if the number of people who do not show up for their appointments is greater toward the end of the month.

For Research Question #5, the variable from the dataset is Rater_score.

Part 1 of 2

Because an investigation for randomness may be necessary with any level of variable (i.e., nominal, ordinal, or scale), a few additional steps are sometimes required before actually running the tests. For example, because SPSS defines Rater_score as a scale variable, you must first identify a cut point — the value used to demarcate the two sides on which the variable values fall. The cut point, sometimes called a threshold, will be used to identify how many times the variable values switch from above to below (or vice versa) the cut point value, which is the method used by the Runs test to determine if the values are not necessarily random.

One way to identify the cut point is to run a descriptive analysis on Rater_score.

- ☐ Select => **Analyze** => **Descriptive Statistics** => **Descriptive** to open the Descriptives window.
- ☐ Click to highlight the Rater_score variable in the list of variables on the left side of the window.
- ☐ Click the move arrow to move it over to the **Variable(s):** area.
- ☐ Select => **OK** button to run the descriptive analysis.

The analysis reveals that the mean = 16.45 and the standard deviation = 9.482, so you must next decide if 16.45 is a practically reasonable threshold for this measure before using it for your analysis. Using the mean value, as you did in this procedure, is a statistical option for determining a cut point value. However, other options are available to you. For example, you can utilize a cut point commonly found in the literature (e.g., a cut point of 16 for the CESD depression scale). Another option is to review the variable's distribution to see if you can identify visually a cut point value. When you have identified your cut point/threshold, you may now conduct the Runs test.

SPSS Process

After you identify your cut point, continue your analysis:

- ☐ Select=> **Analyze** => **Nonparametric Tests** => **One Sample**.
- ☐ Select => **Fields** tab to display the list of available variables.
- ☐ Click on **Use custom field assignments** option.
- ☐ Click to highlight Rater_score in the **Fields:** area.
- ☐ Click the move arrow to move Rater_score to the **Test Fields:** area.

Now that SPSS knows what variable to use,

- ☐ Select => **Settings** tab to see the list of available nonparametric tests.
- ☐ Select => **Choose Tests** to highlight it in the **Select an item:** area.
- ☐ Click to select **Customize tests** option and then click the checkbox for the **Runs test** (uncheck other **Customize tests** options if checked).
- ☐ Select => **Options**. . . to open the Runs Test Options window (see Figure 2.3 — Runs Test Window).

When the Runs Test Options window opens, you will notice that the default for scale variables is to use the **Sample median** option. Because the descriptive information was obtained for Rater_score and the mean value was identified so it could be used in the test, you will need to make a change.

- ☐ Click to select **Sample mean** option in the **Define Cut Pont for Continuous Fields** area.

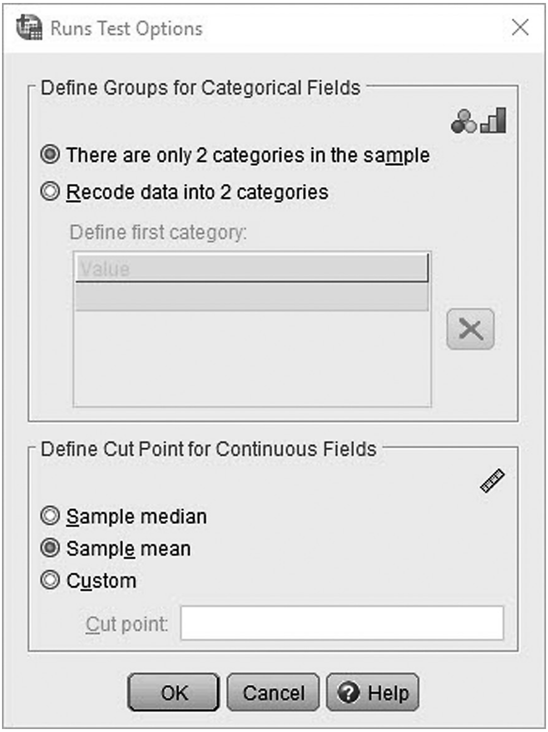


Figure 2.3. Runs Test Options window.

Note: If a cut point other than the mean value is more appropriate for the situation being tested, selecting the **Custom** option provides access to the **Cut point** area, where a specific value can be entered.

- ☐ Select => **OK** button at the bottom to close the Runs Test Options window.
- ☐ After carefully reviewing your choices, select => **Run** to begin the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

Shortly after you select => **Run,** SPSS's output window will display and show the **Hypothesis Test Summary** results for the analysis. To see the details of the test, double-click the **Hypothesis Test Summary** information (i.e., point the mouse within the **Hypothesis Test Summary** table and double-click). The Model Viewer window will open to display the detailed information.

Findings for Part 1 of 2

The graph displayed on the right side of the Model Viewer window shows a normal distribution. The distribution represents the probability of observing the calculated number of runs based on the expected number of runs given the selected cut point. The red line on the graph shows the calculated number of runs and the vertical dashed line shows the expected number of runs. Therefore, the farther away the red line is from the dashed line, the less likely the observations for Rater_score are random.

The Runs test for Rater_score, using the sample mean as the cut point, suggests that you do not have enough information to reject the possibility that Rater_score values are random. In statistical terms, the Runs test of the null hypothesis that the Rater_score values are random, based on the sample mean of 16.448, fails to reject the null hypothesis at $\alpha = .05$ ($p = .577$ in this example).

Practically, patterns in values are sometimes difficult to recognize, especially when the sample size is relatively small. In this situation, the Runs test did not have enough evidence to suggest a pattern. If a pattern does not exist, the natural conclusion is that the values are possibly random, which is the case for Part 1. However, do keep in mind that another explanation is that the small sample size simply did not have enough information to recognize a pattern. Remember, use all the information available (e.g., data distribution, graphs, categorical coverage, etc.) to draw a conclusion for your analysis.

Box 2.7 Research Question #5, Part 2

After reviewing the results from Part 1 above, Dr. Chonody wants you to do one additional test to check her assumption that the order of the youths reporting to have or not have a confidante was random. If Dr. Chonody discovers that the *youths reporting to have a confidante is not random — for example, the first 15 report "no" and the next 15 report "yes" — then she will need to review other aspects of her study (e.g., how she selected respondents, when the respondents filled out their questionnaires) to learn the potential influences causing the nonrandom respondent order.*

Part 2 of 2

Unlike investigating randomness with a scale variable, the runs test on a dichotomous variable does not require the selection of a cut point. With the dichotomous variables' two response options (e.g., *yes* or *no, positive* or *negative*), the runs test will examine the number of times the values switch from one to another. For Part 2 of Research Question #5, the dichotomous variable to investigate is Do you have a confidante? (*1-yes, 0-no*).

To begin this part of your analysis,

- ☐ Select=> **Analyze** => **Nonparametric Tests** => **One Sample**.
- ☐ Select => **Fields** tab to display the list of available variables.
- ☐ Click to select **Use custom field assignments** option.
- ☐ Click to highlight **Do you have a confidante?** in the **Fields:** area.
- ☐ Click the move arrow to move **Do you have a confidante?** to the **Test Fields:** area.

Now that SPSS knows what variables to use,

- ☐ Select => **Settings** tab to see the list of available nonparametric tests.
- ☐ Select => **Choose Tests** to highlight it in the **Select an item:** area.
- ☐ Click on **Customize tests** option and then click the checkbox for the **Runs test** (uncheck other **Customize tests** options if checked).
- ☐ Select => **Options**. . . to open the Runs Test Options window.

When the Runs Test Options window opens,

- ☐ Verify that the option selected for **Define Groups for Categorical Fields** is **There are only 2 categories in the sample**.
- ☐ Once verified, select => **OK** button at the bottom to close the Runs Test Options window.

- ☐ When you have carefully reviewed your choices, select => **Run** to begin the analysis

See Appendix A for complete SPSS syntax.

SPSS Output

Shortly after you select => **Run,** SPSS's output window will display and show the **Hypothesis Test Summary** results for the analysis. To see the details of the test, double-click the **Hypothesis Test Summary** information (i.e., point the mouse within the **Hypothesis Test Summary** table and double-click). The Model Viewer window will open to display the detailed information.

Note: If you left the Rater_score variable from Part 1 in the **Fields:** area when Do you have a confidante? was added, the **Hypothesis Test Summary** will present information on two analyses, one for Rater_score and one for Do you have a confidante? Simply click on the second test (i.e., the second line in the **Hypothesis Test Summary** area) to display the results of the runs test for the dichotomous variable.

Findings for Part 2 of 2

For Confidante, the red line in the graph on the right is farther away from the dashed line than it was for Rater_score, but the distance is still not far enough to suggest that Do you have a confidante? is not random. Clicking between the two tests in the **Hypothesis Test Summary** area, if you tested both Rater_score and Do you have a confidante? at the same time, will show the differences in distance between the two lines, which is reflected in the **Sig.** values. For Rater_score, **Sig.** = .577, and for Do you have a confidante? **Sig.** = .164. Therefore, once again, you do not have enough information to reject the possibility that youths who have a confidante is random (i.e., you have no information to question Dr. Chonody's assumption of youth randomness).

Practically, it is not a good idea to assume something like randomness and then not check to see if it is true. Part 1 was clear in that there was not enough information to suggest a pattern in rater scores. In Part 2, a variable somewhat related to the concern of patterns in rater scores was tested for randomness. Neither variable indicated a

Box 2.8 Research Question #6
Based on the data from her interviews, Dr. Chonody suspected that, for her participants who had not come out to anyone, the older they became, the lower they would score on a self-esteem measure. However, for those who had come out to at least one person, she hypothesized that the older they became, the higher they would score on self-esteem. Dr. Chonody included a validated measure of self-esteem on the follow-up questionnaire to test her hypotheses. She grouped the responses into two groups — those who had come out to at least one person, and those who had not come out to anyone. Because of the small number of participants and the possibility of non-normal data, she should not conduct a Pearson correlation to analyze the data. Dr. Chonody asks you to conduct the analysis.
Research Question #6
Does a relationship (i.e., correlation) exist between age and the youth's level of self-esteem, and does the relationship differ between youths who have not come out and youths who have come out to at least one person?

possible pattern, so the findings indicate that there is no pattern in question quality or youth order. Of course, the runs test works only on unsorted data.

SPEARMAN'S RHO AND KENDALL'S TAU-B COEFFICIENTS

Spearman's rho, also called the Spearman's rank correlation coefficient, and Kendall's tau-b, also called the Kendall rank correlation coefficient, are measures of the relationship between two scale variables. However, unlike the commonly used Pearson correlation, these two coefficients use ranking in the calculation of the correlation value, rather than the assumed linear relationship used in Pearson correlation calculations. The reason for using Spearman's rho and Kendall's tau-b is to avoid the restriction of the assumed linear association and because the Pearson correlation does not effectively deal with outliers, departure from normality, unequal variances, and nonlinearity. In addition, the interpretation of the coefficient values for these two procedures is the same as for the Pearson correlation — e.g., if *y* increases as *z* increases, it is a positive relationship; if *y* increases as *z* decreases, it is a negative relationship. Another strength of the two coefficients is that you can include ordinal variables in the procedure. Note that SPSS will provide a

Pearson correlation value that involves an ordinal variable, but if either of the more appropriate coefficients is requested, the Pearson correlation is not included in the results.

The assumptions for Spearman's rho and Kendall's tau-b are that the variables in the analyses are not nominal and that the participants' responses are independent. In addition, if you are including an ordinal variable, you should determine if you have enough categories to provide conceptual support for calculating a correlation value — e.g., does a correlation with an ordinal variable with three categories make sense? The assumptions for the Pearson correlation are the absence of outliers, equal variances among variables, linear association, and bivariate normality. If you are dealing with a smaller data set, verifying the multiple assumptions may make the decision to use the Pearson correlation risky. When comparing the results of Spearman's rho and Kendall's tau-b, keep in mind that Kendall's tau-b is more efficient at adjusting for matching values (i.e., ties in rank-ordered values) than Spearman's rho.

Other examples are as follows:

1. In your small inpatient unit, you wish to learn if patients who are more depressed have less self-esteem than those who are not depressed, so you can improve treatment protocols.
2. You are analyzing data from a small pilot study, and you wish to explore if general anxiety is associated with different levels of socioeconomic status.

Note from SPSS: "Before calculating a correlation coefficient, screen your data for outliers (which can cause misleading results) and evidence of a linear relationship. Pearson's correlation coefficient is a measure of linear association. Two variables can be perfectly related, but if the relationship is not linear, Pearson's correlation coefficient is not an appropriate statistic for measuring their association."

Note from authors: Research Question #6 illustrates a situation in which the desire is to understand the relationship between two concepts. However, it is always good practice to run Pearson correlations, Spearman's rho, or Kendall's tau-b among all the continuous variables used in a study. The primary reason is that these values will identify if two variables are mimicking each other and thus duplicating the information for a respondent. If you don't fully understand how similar two variables are, using two almost identical variables will bias the findings and invalidate analysis results. Generally, examining bivariate and univariate scatter plots for outliers and linearity is always a good practice when first examining the data. Another reason to run correlations on all the variables is to help validate the data, which provides another opportunity to find data-entry mistakes or to catch an error in how a variable was operationalized.

For Research Question #6, the variables from the data set are Age and Self_esteem. Because Dr. Chonody is interested in learning more about the relationship between Age and Self_esteem, and whether this relationship differs for youth who have come out, a good way to begin is to examine the information contained in the two variables. You can do this by having SPSS create histograms of their distributions.

SPSS Process

To request histograms,

- □ Select => **Analyze** => **Descriptive Statistics** => **Frequencies** to open the Frequencies window.
- □ Click to select Age in the list of available variables on the left side of the window, then click the move arrow to move the variable to the **Variable(s):** area.
- □ Click to select Self_esteem in the list of available variables on the left side of the window, then click the move arrow to move the variable to the **Variable(s):** area.
- □ Select => **Charts**. . . to open the Frequencies: Charts window.
- □ Click to select the **Histograms:** option and then click to select the **Show normal curve on histogram** checkbox.
- □ Select => **Continue** button to close the Frequencies: Charts window.
- □ Select => **OK** button at the bottom to close the Frequencies window and generate the histograms.

Shortly after you select **OK,** SPSS's output window will display a few tables, and if you scroll down the window you will see the two histograms for Age and Self_esteem. These two histograms show the variable distributions for all the youths who completed the follow-up questionnaire. However, along with these two graphs, you want histograms for the two groups of youths. To do this, you need to set a data filter in SPSS. First filter to get only youths who have not come out to anyone, request histograms, and then filter to get youths who have come out to one or more people.

To filter respondents to select only the youths who have not come out to anyone, and ignoring youths who have come out to one or more people,

- □ Select => **Data** => **Select cases**. . . to open the Select Cases window.
- □ Click to select the **If condition is satisfied** option in the **Select** area.
- □ Select => **If**. . . button to open the Select Cases: If window.
- □ Click to select the Coming_Out variable on the left side of the window, then click the move arrow to move the variable to the white space to the right of the move arrow.
- □ Right after the variable name Coming_Out in the white space, click to move the focus to right after the variable name.
- □ Type =1 (i.e., an equals sign and a 1). This tells SPSS to use cases that have a value of 1 for the variable Coming_Out. Recall, a 1 represents youths who have not come out to anyone.
- □ Select => **Continue** button to close the Select Cases: If window.
- □ Select => **OK** button to close the Select Cases window and use only the cases that have a 1 for Coming_Out.

You can verify that the filter is working by looking at SPSS's cases view datasheet. You can identify the cases that have been filtered by the diagonal line that appears across some of the line numbers on the left side of the datasheet. Another way to verify that the filter is working as expected is to review the value for *N* (i.e., number of cases used in the analysis) on the notes in the histogram.

Now that the youths who have come out to one or more people have not been included, you can rerun the steps for requesting histograms (see "To request histograms"). After you have obtained the histograms for your first group of youths, you need to change the SPSS data filter to select the other group of youths.

To filter respondents to select youths who have come out to one or more people, ignoring youths who have not come out to anyone,

- ☐ Select => **Data** => **Select cases**. . . to open the Select Cases window.
- ☐ Click to select the **If condition is satisfied** option in the **Select** area.
- ☐ Select => **If**. . . button to open the Select Cases: If window.

The white area should still have the previous filter command (i.e., Coming_Out = 1), but, if not,

- ☐ Click to select the Coming_Out variable on the left side of the window, then click the move arrow to move the variable to the white space to the right of the move arrow.
- ☐ Right after the variable name Coming_Out in the white space, click to move the focus to right after the variable name.

Next, to filter youths who have come out to one or a few people (Coming_Out values of 2) and youths who have come out to everyone (Coming_Out values of 3), you need to use a "not equal to" symbol instead of the equal sign. Therefore, you want the filter command in the white area that will tell SPSS to use all the cases that have something other than a 1 for Coming_Out (i.e., Coming_Out ~=1, where ~= is the same as ≠).

- ☐ Once you set the filter command to use all the cases that do not have a 1 for Coming_Out, select => **Continue** button to close the Select Cases: If window.
- ☐ Select => **OK** button to close the Select Cases window and filter out all the cases that do not have a 1 for Coming_Out.

Now that your filter is set, you can again rerun the request to get histograms for the second group of youths by following the same steps as before (i.e., steps under "To request histograms").

SPSS Output

Figures 2.4, 2.5, and 2.6, show the histograms produced by completing the steps above (i.e., Age and Self_esteem histograms for all the youths, youths who have not come out to anyone, and youths who have come out to one or more people). A review of the values for *N* in each of the

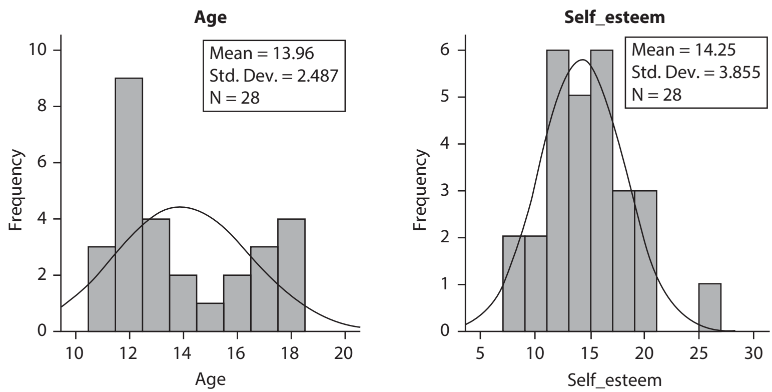


Figure 2.4. Age and Self_esteem histograms for all youths.

figures shows the number of youths represented by the histograms (i.e., 28, 9, and 19, respectively).

The reason for reviewing all the histograms is to fully understand the information captured by the two variables, Age and Self_esteem, by the different categories for youths coming out. A review of the histograms for Age does not show signs of a normal distribution. In addition, the histogram for Self_esteem for all the youths appears to follow closely a normal distribution, but when only youths who have not come out to

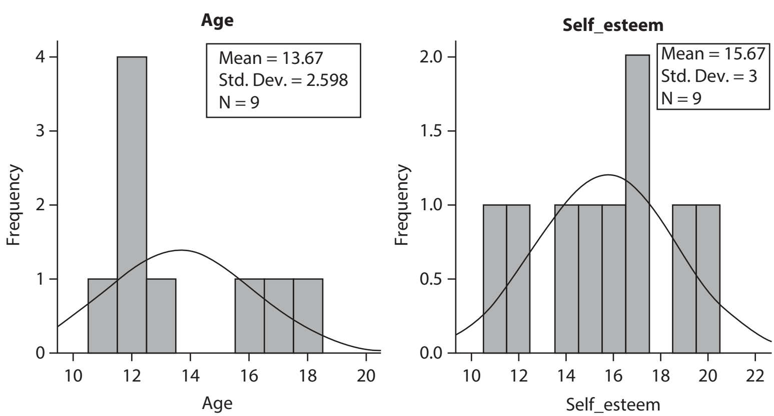


Figure 2.5. Age and Self_esteem histograms for youths who have not come out to anyone.

Figure 2.6. Age and Self_esteem histograms for youths who have come out to one or more people.

anyone are reviewed (i.e., Figure 2.5), the distribution appears to follow a uniform distribution.

Another reason a close review of the data using histograms is important is that a single statistical test alone is not enough to support a strong conclusion. For example, if the above distributions for all respondents were tested for normality (don't forget to remove the filter) using the Kolmogorov-Smirnov test (see Research Question #3), both normal and Poisson distributions are possible for Self_esteem. The reality is that, when dealing with small samples, the amount of available information to reject a hypothesis is small, making it more of a challenge to find significance (i.e., more difficult to reject a null hypothesis). Therefore, a researcher must use different approaches to draw conclusions about analysis findings as well as information to validate assumptions. Because the histograms for Age are nonnormal, the procedures to answer Research Question #6 must be nonparametric.

Note: When using filters, you must pay extra attention to make sure your filters are working properly before conducting any analysis or requesting graphs. SPSS has no way to warn you if your filter is not working as you intended, so you must look at what cases are and are not filtered out. In addition, make sure you remove any filters when you are finished using subsets of your dataset. Forgetting to remove a filter can be very confusing if you move forward with other analyses and cannot understand the unexpected analysis results!

SPSS Process

If you still have an active filter from requesting the histograms above,

- ☐ Select => **Data** => **Select cases**. . . to open the Select Cases window.
- ☐ Click to select **All cases** option in the **Select** area.
- ☐ Select => **OK** button to close the Select Cases window and remove any filter on the dataset.

Now that there is no active filter,

- ☐ Select => **Analyze** => **Correlate** => **Bivariate** to open the Bivariate Correlations window

Note that Spearman's rho and Kendall's tau-b are not found in the **Nonparametric Tests** menu structure**.**

After the Bivariate Correlations window opens,

- ☐ Click to select Age in the list of variables on the left side of the window, then click the move arrow to move the variable to the **Variables:** area.
- ☐ Click to select Self_esteem in the list of variables on the left side of the window, then click the move arrow to move the variable to the **Variables:** area (both Age and Self_esteem appear in the **Variables:** area).
- ☐ Click the check boxes next to both **Kendall's tau-b** and **Spearman** tests. (*Note:* Even if **Pearson** is checked, it is ignored if either of the other two is selected.)
- ☐ Click to select the **Flag significant correlations** checkbox.
- ☐ After reviewed your choices carefully, select => **OK** button at the bottom of the Bivariate Correlations window to run the correlation analyses

See Appendix A for complete SPSS syntax.

SPSS Output

Shortly after you click **OK,** SPSS's output window will display and show the **Correlations** results for the analyses (see Table 2.3).

Findings

Table 2.3 shows that both Spearman's rho and Kendall's tau-b measure a negative relationship between Age and Self_esteem. Note that the Kendall's tau-b finds a significant correlation ($p = .042$) between the two variables, while the Spearman's rho fails to find significance ($p = .061$) at $\alpha = .05$. You can attribute the difference between the two tests to their differences in dealing with ties within the ranking process. To verify that you have ties in the data, you can sort the data by Age and look for multiple occurrences of Age and Self_esteem combinations.

Table 2.3. Correlation Results for Spearman's rho and Kendall's tau-b

Correlations

			Age	*Self_esteem*
Kendall's tau-b	Age	Correlation Coefficient	1.000	−.296*
		Sig. (2-tailed)	.	.042
		N	28	28
	Self_esteem	Correlation Coefficient	−.296*	1.000
		Sig. (2-tailed)	.042	.
		N	28	28
Spearman's rho	Age	Correlation Coefficient	1.000	−.359
		Sig. (2-tailed)	.	.061
		N	28	28
	Self_esteem	Correlation Coefficient	−.359	1.000
		Sig. (2-tailed)	.061	
		N	28	28

* Correlation is significant at the 0.05 level (2-tailed).

The dot in SPSS represents a non-reported value.

Note: A highly respected statistician whose specialty is nonparametric statistics, Dr. Myles Hollander, believed that when using nonparametric statistics, the strict adherence to $\alpha = .05$ was too aggressive.

He believed that with a thorough understanding of the data, the use of $\alpha = .10$ was just as informative (M. Hollander, personal communication, 2003). Best practice, of course, is to identify the alpha you will use prior to beginning the analyses. In the example above, we believe that concluding that you have a difference between a $p = .042$ and .061 would be a bit arbitrary and that both are sufficient to conclude a significant correlation finding.

To sort data in SPSS,

- ☐ Select => **Data** => **Sort cases**. . . to open the Sort Cases window.
- ☐ Click to select Age in the list of variables on the left side of the window, then click the move arrow to move the variable to the **Sort by:** area.
- ☐ Select => **OK** button to close the Sort Cases window.

Note: When planning to sort cases, it is a good practice to have one variable that will allow you to return the cases to their original order. You can accomplish this by adding a variable to the dataset that assigns an ID to the cases, counting from 1 through *k*. The variable in the data set for Chapter 2 that identifies the original order is ID.

After you sort the data by Age, a review of the values for Age and Self_esteem shows multiple occasions in which cases have the same Age and Self_esteem values — suggesting that the tests managed multiple ties. Because Kendall's tau-b handles ties more efficiently than Spearman's rho, it makes logical sense that Kendall's tau-b calculated a smaller *p*-value than Spearman's rho. Again, you should always pursue multiple ways to verify your findings and the performance of your test so that you can have confidence in your conclusions.

However, in Research Question #6, Dr. Chonody is interested in more than just the relationship between Age and Self_esteem for all the youths. She is interested in the possible differences in the relationships between youths who have not come out and youths who have. Therefore, you must run the correlation analysis two more times, once after filtering the cases so that youths who have come out are excluded, and then again after the youths who have not come out are excluded (i.e., filter on the variable Coming_Out).

Using the steps discussed earlier to filter out certain cases and the steps to conduct the Spearman's rho and Kendall's tau-b tests, you collect the **Sig.** values for the multiple tests into a single table for review (see Table 2.4).

Although the negative correlations still exist when you divide the youths into two groups, youths who have come out and youths who have not, the **Sig.** values in Table 2.4 show that the negative relationship between Age and Self_esteem is being heavily influenced by the youths who have come out (i.e., **Sig.** values are significant at .032 and .043). In addition, Table 2.4 shows that the relationship between Age and Self_esteem is not significant for youths who have not come out, and the negative relationship between Age and Self_esteem is moderate (i.e., correlation value is between -.3 and -.6) for youths who have come out. Therefore, the **Sig.** values for youths who have come out to anyone suggest that there is only a 3% to 4% chance that there is no relationship between Age and Self_esteem.

A more statistical explanation of the results of the correlation analysis is that a test of the relationship between Age and Self_esteem using Kendall's tau-b is -.384 and found to be significant (p = .032) for youths who have come out to one or more people (n = 19).

Practically, if the gay youths in Dr. Chonody's study are typical of all gay youths, the findings suggest that self-esteem is not necessarily linked to age for youths who have not come out. However, for youth who have come out to anyone, the findings suggest that their self-esteem will decrease as they get older. This finding should concern you and push you to conduct a closer examination of, among other things, your data, your sample, how the data were collected, and the wording of the questions provided to the respondents. The reason for the closer examination is that these findings seem to be in conflict of what you

Table 2.4. Correlation for Age and Self_esteem for Two Youth Groups

	Youths who have not come out to anyone			*Youths who have come out to one or more people*		
Correlation test	Correlation	Sig. (2-tailed)	*N*	Correlation	Sig. (2-tailed)	*N*
Kendall's tau-b	.031	.912	9	.384	.032	19
Spearman's rho	.057	.885	9	—.469	.043	19

might expect — that individuals who have come out would experience improved self-esteem as they grow older. Remember, your responsibility as a researcher goes beyond just following the statistical steps to test a hypothesis. Your responsibility includes understanding all the potential threats, limitations, and possible biases that are present in your study. Examining all possible issues is the only way you can be sure that your findings are ready to present to others.

Research Question #6 explored one way to compare two groups of youths. To explore further the possible ways to compare groups, Chapter 3 introduces nonparametric procedures that involve two or more independent groups, and Chapter 4 introduces nonparametric procedures that involve related or dependent groups. Knowing these nonparametric group-related procedures will greatly expand your ability to analyze data, especially when your data do not meet the assumptions for parametric statistics.

3

Comparing Two or More Independent Groups

CHAPTER OBJECTIVES

This chapter covers analyses that compare two or more independent groups, and the topics presented are as follows:

- Determining dependent (related) and independent (unrelated) groupings
- Selecting tests for two independent groups (similar to *t*-test for independent samples)
- Mann-Whitney U test
- Kolmogorov-Smirnov test
- Wald-Wolfowitz runs
- Moses test of extreme reactions
- Hodges-Lehmann estimate
- Selecting tests for more than two independent groups (similar to ANOVA and MANOVA)
- Kruskal-Wallis one-way ANOVA
- Jonckheere-Terpstra for k samples
- Median test

INTRODUCTION

Recall that Chapter 1 describes the importance of understanding the data and individual variables used in statistical tests, and Chapter 2 explains how to examine closely the individual variables and how to conduct nonparametric tests that involve variables for a single group. This includes discussion of how to inspect individual variables to gain a deeper understanding of exactly how much information the variable contributes to its conceptual construct and characteristics. Most importantly, a deeper understanding of the individual variable will help you select variables for other purposes, such as analyses that involve two or more groups of respondents.

Chapter 3 explores independent groups and how to use nonparametric statistics that are specific to group comparisons. The term *independent groups* refers to the situations in which a response from an individual in one group is not influenced by a response from an individual in another group. In comparison, *related groups*, discussed in Chapter 4, involve responses that have a connection (e.g., examining the difference between two responses from the same individual, so that the individual is the connection).

Many of the data concerns from previous chapters remain when moving from single to multiple groups, but the concerns are more pronounced in this chapter because the information available from the data is now being shared among a number of groups. This becomes more of a challenge when dealing with the limited information available in a small sample. The limited information shared among the groups makes meeting parametric analysis assumptions much more difficult, if not impossible. Thus, knowing how to use nonparametric procedures becomes critical to good research. Becoming familiar with nonparametric procedures for groups will help avoid the potential mistake of reporting invalid results due to running a parametric procedure on data that do not meet the necessary assumptions. In addition, verifying that the assumptions are met is more challenging when the data are broken up into even more pieces when involving even more groups.

Research Questions 7 to 11 illustrate situations that involve comparing independent groups and the manner in which data are analyzed.

After each research question, detailed steps are provided for how to perform an analysis, issues to be aware of when conducting the analysis, and how to interpret test results. Simply stated, Chapter 3 covers nonparametric statistics for independent group comparison, including the researchers' responsibility to understand the variables individually as well as how variables are used to investigate potential group differences.

Box 3.1 Scenario for Chapter 3
Dr. Rutledge conducted a small feasibility study of an anticipated clinical trial for healthy, sexually active HIV + men and women. The intervention, a single motivational interviewing session, targets the reduction of HIV transmission by addressing participants' ambivalence about safer sexual practices. Dr. Rutledge recruited potential participants through advertisements in newspapers and flyers placed in physicians' offices. Participants attended an initial screening interview, during which they answered questions to determine the appropriateness of their inclusion for the study. At this time, they also responded to a computer-based pretest questionnaire that included behavioral questions and validated measures of knowledge, attitudes, and beliefs. Dr. Rutledge randomly assigned a total of 60 men and women into one of two groups — a treatment group that would receive the intervention immediately, and a control group that would receive the intervention after the completion of the feasibility study. The 30-member treatment group participated in the motivational interview and completed the posttest, a computer-based questionnaire. The control group completed the posttest as well. You will act as Dr. Rutledge's statistical consultant. His research questions follow.

Like Chapter 2, Chapter 3 begins with a scenario describing a hypothetical line of study that can be explored with specific research questions.

Chapter 3 contains five research questions that guide the presentation of the nonparametric procedures related to the research scenario above. Each research question illustrates the reasoning for using nonparametric procedures. Table 3.1 lists the nonparametric procedures presented in Chapter 3, the kinds of variables for each procedure, and the names of the variables found in the database for Chapter 3 that is available online.

Table 3.1. Summary of Analyses and Variables in Chapter 3

Research Question #	*Nonparametric Test*	*Variable Measure*	*Variable Name in DB*
7	Moses extreme reaction, Hodges-Lehmann confidence intervals	One dichotomous and one scale	Received_MI and Percent_Pre
8	Kolmogorov-Smirnov, Mann-Whitney U, Wald-Wolfowitz	One dichotomous and one scale	Received_MI and Percent_DIF
9	Kruskal-Wallis	One categorical with 3 or more groups, either nominal or ordinal, and one scale	Received_MI, Percent_Post, and Partner
10	Jonckheere-Terpstra	One categorical with 3 or more groups, either nominal or ordinal, and one scale	Stage_of_Change and Percent_DIF
11	Median	One categorical with 3 or more groups, either nominal or ordinal, and one scale	Condom_Request and Self_efficacy

Box 3.2 Research Question #7

Dr. Rutledge's main research question is whether the Motivational Interviewing (MI) intervention reduced the percentage of times the participants had unprotected sex in the two months following the intervention. The first step was to place the participants randomly into two groups (control and treatment) in which the treatment group will receive the intervention. Next, to be certain that he has comparable groups prior to conducting the intervention, Dr. Rutledge wants you to compare the two groups on their pretest scores for their report of the approximate percentage of time they use condoms when having sex. He wants to be sure that the median value of each group is similar so he can properly assess the influence of the MI intervention.

Research Question #7

Are the median and distribution of percentage values for the treatment and control groups similar?

MOSES EXTREME REACTION TEST

The Moses test of extreme reactions focuses on the tail ends of the range of scores to analyze whether two distributions are similar. As in other nonparametric tests, a ranking strategy is used while focusing on the lower and upper values to test the hypothesis that extreme values are equally likely in both distributions. The test is restricted to two groups, with the assumption that the group values captured by a single scale variable are independent. A common way to compare distributions for extreme values is to graph the distributions using either histograms or box-and-whisker plots, and if you assume the distributions are normal, you have the option of testing for equal means. However, other than a visual examination of the data, SPSS does not offer a parametric test for equal probability of extreme values. Thus, the Moses test of extreme reactions is conducted in concert with the Hodges-Lehmann estimate, as follows.

HODGES-LEHMANN ESTIMATE (CONFIDENCE INTERVALS)

Multiple Hodges-Lehmann estimates exist, but described here is the estimate of a confidence interval for the median difference between two distributions. Thus, the Hodges-Lehmann estimate returns a calculated confidence interval for median difference values, which provides an opportunity to examine a potential range of median differences. However, the estimate assumes that the distributions are similar (i.e., kurtosis and skewness values are similar), so the estimate is valid only after it is verified that the two distributions are similar in shape, which is especially challenging when sample sizes are small.

Note: The Hodges-Lehmann assumes that individual group distributions are similar and provides a confidence interval on the plausible difference in median values for a continuous variable (e.g., percentage of times using a condom) across two groups (e.g., treatment and control).

Other examples for use of the Moses test are as follows:

1. You believe the arrival times of visitors to a small agency are similar on Monday and Thursday, and you need the information for scheduling.

2. You believe that depression levels are similar in your small sample of men and women, and you need to know this before developing an intervention plan.

Running the Moses extreme reaction test and requesting a confidence interval from the Hodges-Lehmann procedure together provide a statistical test of equal medians and give you information on the possible range of the median difference. Of course, if the Moses extreme reaction test rejects the null hypothesis (i.e., finds that the two groups have different medians), the assumption for the Hodges-Lehmann procedure is not met. However, if the Moses extreme reaction test does not reject the null hypothesis, the Hodges-Lehmann procedure provides information on a plausible range of median difference values that gives you more insight into how similar the two groups are.

The variables needed for the Moses extreme reaction test and the Hodges-Lehmann confidence interval are a dichotomous variable representing two groups and a scale/continuous variable. For Research Question #7 (see Box 3.2), the dichotomous variable is Received_MI, a grouping variable to distinguish who did and did not receive the Motivational Interviewing (MI) intervention. The scale variable is Percent_Pre, the percentage of unsafe sex episodes measured prior to giving one group the intervention.

SPSS Process

To begin your analysis,

- □ Select => **Analyze** => **Nonparametric Tests** => **Independent Samples**.

Once the Nonparametric Tests: Two or More Independent Samples window opens, SPSS needs to know which variables to use in the analysis and which analysis routines to run.

- □ Select => **Fields** tab to see the list of available variables for the analyses.

☐ Click to select **Use custom field assignments** option.
☐ Click to select the grouping variable Received_MI.
☐ Click the move arrow closest to the **Groups:** area to move the variable to that area.
☐ Click to select the scale variable Percent_Pre that contains percent values for respondents' condom use.
☐ Click the move arrow closest to the **Test Fields:** area to move the variable to that area.

Now that SPSS knows what variables to use,

☐ Select => **Settings** tab to see the list of available nonparametric tests.
☐ Select => **Choose Tests** from the **Select an item:** area.
☐ Click to select **Customize tests** to enable the ability to select specific tests.
☐ Click the check boxes next to **Moses extreme reaction.**
☐ Click the check box next to **Hodges-Lehmann estimate**.

For Research Question #7, the only test option to consider is the Moses extreme reaction for identifying outliers. Use the default option (i.e., **Compute outliers from sample**) to use the observed data.

☐ When you have reviewed your choices carefully, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click). The Model Viewer window will open to display the detailed information for the analysis.

Findings

The information for the Moses extreme reaction test shows that the two groups have similar values for percentage of time using a condom (i.e., **Sig.** = .665 is larger than the α = .05 significance level). The Moses extreme reaction test provides an exact significance (**Sig.** is the notation SPSS uses to represent statistical significance), indicating that there is a 66.5% chance that the ranges are the same. By tradition, statistical significance must be small, with only a 5% or smaller chance that the medians are similar, before concluding that the groups are different.

Note: Not all statistical tests provide exact significance values. Therefore, be sure to translate a **Sig.** value to "percent chance" interpretation only when the test allows this.

To see the Hodges-Lehmann results, you must change the view. At the bottom left of the Model Viewer window, one of the dropdown lists is named **View:**.

- □ Click on the **View:** dropdown list and click to select **Confidence Interval Summary View**.

After you select the **Confidence Interval Summary View**, the information for the **Confidence Interval Summary** is displayed. It reveals a 95% confidence interval for the difference between median values range from -.20 to .10. This range includes a plausible difference of zero and, therefore, is in line with the conclusion drawn by the Moses extreme reaction test.

In addition, a box-and-whiskers plot associated with the test is included with the other analysis results. Box-and-whiskers plots are a graphical representation of a distribution of values. The box part of the graph represents the first and third quartiles, called the interquartile range or IQR, and the thicker dark line represents the median value. The lines extending from the box are called whiskers and are calculated differently by different applications. SPSS draws the end of each whisker as the minimum and maximum values in the data that are not identified as statistical outliers. SPSS identifies outliers as values that fall farther away from the first and third quartiles than 1.5 times the size of the IQR. SPSS identifies outliers with small circles and with asterisks when outliers are three times beyond the IQR. Figure 3.1 does not identify the presence of any outliers, but Figure 3.5 does. The similarity of the vertical positions of these areas for the two groups illustrates and supports the values for the confidence interval and the hypothesis test finding of similar range values (see Figure 3.1).

In summary, the Moses extreme reaction test did not find a significant difference in the range of percentage of condom use between the case and control groups. In addition, the Hodges-Lehmann confidence interval for plausible median difference identified a lower and upper bound of –.20 and .10, respectively. Therefore, the two analyses support the conclusion that there is not enough information to suggest that the median value or range of values for condom use is different.

Figure 3.1. Box-and-whiskers plot for Moses test of extreme reactions.

In statistical terms, the Moses extreme reaction test (p = .648) did not reject the null hypothesis at α = .05 of similar range values for percent condom use between case and control groups. In addition, the Hodges-Lehmann confidence interval found a 95% confidence interval of –.20 to .10. These two values indicate the range in which we are 95% confident that the true percent difference falls. It is not a coincidence that the null hypothesis was not rejected and the confidence interval includes a zero percent difference.

Practically, the above tests verify that the randomization process did not create dissimilar groups that would bias the examination of the intervention. If the tests had suggested a possible difference, then Dr. Rutledge would need to conduct a new randomization of the respondents until the two groups were similar in their condom use. Without a comparison of the two groups prior to conducting the intervention, the results could have been inaccurate without any evidence to inform Dr. Rutledge otherwise. Of course, a statistical test using α = .05 suggests a 5% probability that a difference between the case and control groups is identified when the truth is that there isn't a difference (i.e., .05 probability of incorrectly rejecting the null).

Box 3.3 Research Question #8

Now that Dr. Rutledge has verified that the two groups are similar in their condom use, he can proceed to conducting the MI intervention on the treatment group members and then test to find out if the intervention makes a difference. To examine this, he created change scores for each group. The change scores represent the difference in the percentage of time having unprotected sex from pre-intervention condom use to post-intervention condom use. Dr. Rutledge asks you to compare the changes scores for the treatment group with those of the control group to answer his question.

Research Question #8

Are the changes in percentage of time having unprotected sex different between those who have received the intervention and those who have not?

Note: A nonparametric ANCOVA procedure does exist (Tsangari & Akritas, 2004), and you could use it to answer this research question. However, it is not a procedure available in SPSS, so it is not discussed herein.

KOLMOGOROV-SMIRNOV TEST

The Kolmogorov-Smirnov test calculates the amount of separation of the empirical distributions for the two groups, and it uses the amount of separation to decide if the two distributions are significantly different. Said differently, it tests to find out if it is possible that the values for the two groups come from the same distribution. Because the focus is on the separation of the two distributions, no prior assumption about the distribution exists, unlike the *t*-test, in which the data are assumed to come from a normal distribution. Recall that when the data clearly meet all parametric assumptions, the parametric procedure does have more power to find significance than its nonparametric counterpart (emphasis on the word *clearly*). Importantly for the Kolmogorov-Smirnov test, not having a normality assumption does make the test less able to find significant differences between distributions than the parametric *t*-test when the data clearly meet the *t*-test's normality assumption. In addition, SPSS restricts the Kolmogorov-Smirnov test to two groups.

Other examples for uses of the Kolmogorov-Smirnov test are as follows:

1. You suspect a difference in PTSD between men and women clients at your small treatment center; therefore, you have everyone complete a standardized measure of PTSD so you can examine the differences by gender.
2. You hypothesize that those in your youth group who identify as being spiritual are happier than those who do not, so you ask them to complete a small survey to investigate this.

Remember, prior to conducting any hypothesis test, you must examine the information provided by the variables. For Research Question #8 (see Box 3.3), one variable represents group membership (i.e., participated or not in MI) and the other represents the percent change in having unsafe sex. We assume that you have already used the steps discussed in Chapter 2 to examine the variables individually.

SPSS Process

To begin your analysis,

- ☐ Select => **Analyze** => **Nonparametric Tests** => **Independent Samples**.

Once the Nonparametric Tests: Two or More Independent Samples window opens, SPSS needs to know which variables to use in the analysis and which analysis routines to run.

- ☐ Select => **Fields** tab to see the list of available variables for the analyses.
- ☐ Click to select the **Use custom field assignments** option.
- ☐ Click to select the grouping variable Received_MI.
- ☐ Click the move arrow closest to the **Groups:** area to move the variable to that area.
- ☐ Click to select the scale variable (Percent_DIF) that contains change in percentage values.
- ☐ Click the move arrow closest to the **Test Fields:** area to move the variable to that area.

Now that SPSS knows what variables to use,

- ☐ Select => **Settings** tab to see the list of available nonparametric tests.
- ☐ Select => **Choose Tests** from the **Select an item:** area.
- ☐ Click to select **Customize tests** to enable the ability to select specific tests.
- ☐ Click the check box to the left side of **Kolmogorov-Smirnov** test in the **Compare Distribution across Groups** list of tests area.

You may run into situations in which you want to change the significance level for your test. In addition, some of the tests within this list of nonparametric procedures can estimate confidence intervals. The steps below show how to change the options associated with significance level and confidence intervals.

☐ Select => **Test Options** from the **Select an item:** area.
☐ Decide on the significance level (default is 0.05) for the test.
☐ Decide on the confidence interval range (default is 95.0%).
☐ Decide how to handle missing data (default is Exclude cases test-by-test).
☐ After reviewing your choices carefully, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click). The Model Viewer window will open to display the detailed information for the analysis.

Findings

The Kolmogorov-Smirnov test shows **Sig.** = .056, which is very close to the **Significance level:** set in the **Test Options** from the **Select an item:** area. In the **Hypothesis Test Summary** information, SPSS is not indicating that this test rejects the null hypothesis (i.e., the null hypothesis is that distributions between the two groups are the same). We discuss in Chapter 2 how nonparametric procedures are more conservative when finding test significance. Therefore, a case can be made that this test is close enough to rejecting the null hypothesis to report a difference in group distributions, although some statistical purists will definitely disagree. Ideally, you would make the decision to select a more nontraditional significance level (e.g., $\alpha = .10$) before conducting the analysis. We have learned that clearly describing your reasoning for determining significance is sufficient information to have a manuscript published or a grant proposal funded.

Another approach is to gather more information before drawing a conclusion about the findings. To do this, simply conduct the other two **Compare Distribution across Groups** tests to learn how all three tests report their findings. The two others are the Mann-Whitney U test and the Wald-Wolfowitz test, as follow.

MANN-WHITNEY U TEST

The Mann-Whitney U test, also called the Wilcoxon rank sum test, uses ranking of the observed values to determine if the two groups come from the same distribution. Unlike the parametric equivalent independent *t*-test that assumes a normal distribution, the Mann-Whitney has no distribution assumption. The only restrictions in SPSS are that the test is for only two groups and the comparison variable is on a scale level of measurement. The lack of any distribution assumption does make the Mann-Whitney U appropriate for any type of distribution, but, as with the Kolmogorov-Smirnov and Wald-Wolfowitz tests, the flexibility does make finding significant differences between the distributions a bit more difficult.

WALD-WOLFOWITZ TEST

The Wald-Wolfowitz test analyzes runs for the values while focusing on group membership. The test's structure provides an opportunity to test the hypothesis that values for a scale variable from two groups have similar value distributions. The test is restricted to two groups for comparing the distribution of a scale variable. Because the Wald-Wolfowitz uses runs, no assumption about the distribution is required. However, when a sample size is greater than 30, SPSS uses a normal approximation for a test statistic because as the number of possible runs reaches 30+, the distribution for the number of possible runs is similar to a normal distribution. Another use of this test is to examine if responses from two groups captured by a single scale variable are independent. The execution of the test is the same, only the interpretation of the findings changes (i.e., distributions of a scale variable found to be different for two groups imply that responses between the two groups are independent).

The summary of tests for Research Question #8 is:

- Kolmogorov-Smirnov — uses distribution separation to draw conclusion about similarity between distributions.
- Mann-Whitney U — uses ranking to draw conclusion about similarity between distributions.
- Wald-Wolfowitz — uses runs to draw conclusion about similarity between distributions.

SPSS Process

To begin your analysis using all three tests,

- ☐ Select => **Analyze** => **Nonparametric Tests** => **Independent Samples**.

Once the Nonparametric Tests: Two or More Independent Samples window opens, SPSS needs to verify that the variables to use in the analysis are still selected.

- ☐ Select => **Fields** tab to see the list of available variables for the analyses.
- ☐ Click to select the **Use custom fields assignments** option.

If Percent_DIF and Received_MI are not in the **Test Fields:** and **Groups:** areas,

- ☐ Click to select the grouping variable Received_MI.
- ☐ Click the move arrow closest to the **Groups:** area to move the variable to that area.
- ☐ Click to select the scale variable (Percent_DIF).

☐ Click the move arrow closest to the **Test Fields:** area to move the variable to that area.

Now that SPSS knows what variables to use,

☐ Select => **Settings** tab to see the list of available nonparametric tests.
☐ Select => **Choose Tests** from the **Select an item:** area.
☐ Click to select **Customize tests** to enable the ability to select specific tests.

Figure 3.2 shows many options for nonparametric tests, some for two groups and some for more than two groups. Looking closely at Figure 3.2 will show five tests specific to having only two groups.

☐ Click to select the three check boxes on the left side of the **Compare Distribution across Groups** list of tests — Mann-Whitney U, Kolmogorov-Smirnov, and Wald-Wolfowitz.

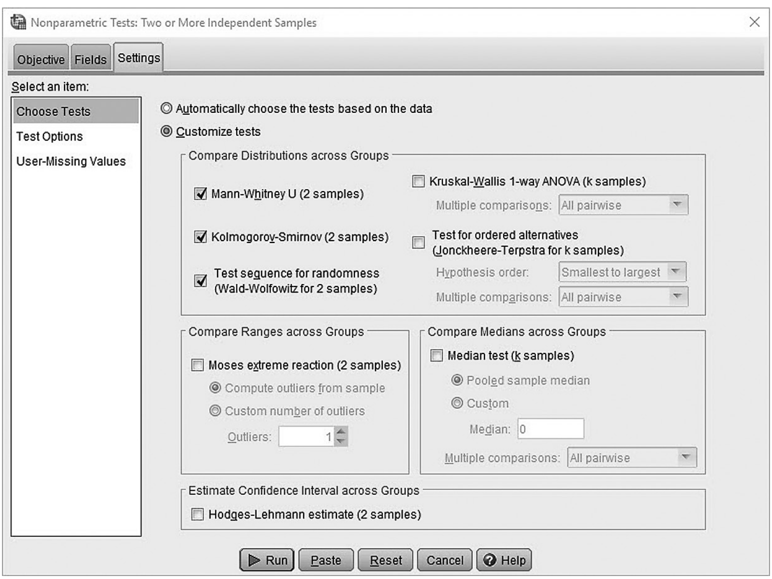


Figure 3.2. Nonparametric Tests: Two or More Independent Samples window.

All three "for 2 samples" tests are selected because, as discussed above, each one approaches the test of similar groups differently. Therefore, to fully understand if the two groups are actually different, the results from each of the three tests need to be reviewed.

Before continuing, review the test options to make sure they are set appropriately.

- ☐ Select => **Test Options** in the **Select an item:** area to see the place where to change default settings.
- ☐ Decide on the significance level (default is 0.05) for the test.
- ☐ Decide on the confidence interval range (default is 95.0%).
- ☐ Decide how to handle missing data (default is Exclude cases test-by-test).
- ☐ After reviewing your choices carefully, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click). The Model Viewer window will open to display the detailed information (see Figure 3.2).

The Model Viewer window has access to more than what is shown in Figure 3.2. By clicking on one of the three test results listed in the **Hypothesis Test Summary** table on the left side of the Model Viewer

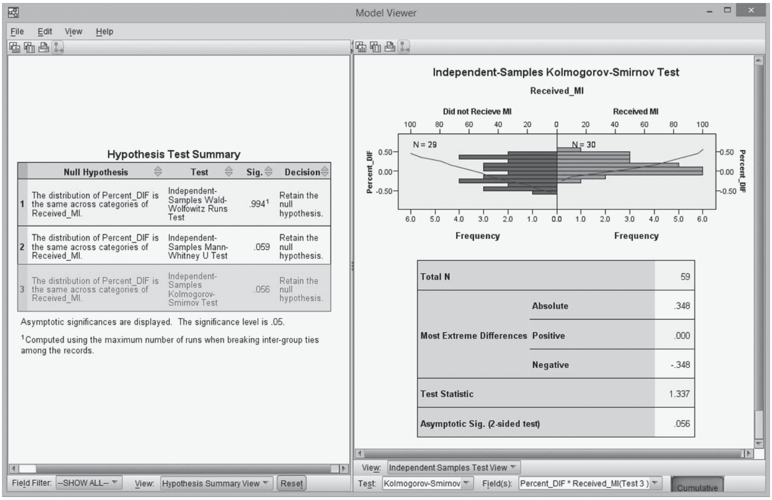


Figure 3.3. Model Viewer window with detailed test results.

window, you change the information on the right side to that particular test information.

Another way to switch between test information is to change the selection from the dropdown lists on the left for **View** and **Field Filter**, and on the right for **Test**, **Field(s)**, and **View**. Figure 3.2 shows the information for the Kolmogorov-Smirnov test because that test is the one highlighted in the list of tests on the right.

Findings

A review of the **Sig.** values finds that the Wald-Wolfowitz runs test is clearly not rejecting the null hypothesis (p = .994). The Mann-Whitley U test (p = .059) and the Kolmogorov-Smirnov test (p = .056) are similar in their findings. In other words, the two are close to rejecting the null hypothesis at α = .05 that the groups are similar in their Percent_DIF values. Differing values for these three tests is common, and this situation demonstrates how important it is to review all three. To complicate matters even further, if you conduct an independent-samples t-test on the data, the parametric test will return a **Sig.** value of .012 — leading you to conclude that the two groups are significantly different for values of Percent_DIF.

When faced with conflicting information, begin with what you know about the tests. You know that the nonparametric tests are less likely to find significance than parametric tests when parametric assumptions are met. You also know that one guideline for deciding if you have enough information about a group is 30 per group, although small group sizes can typically detect only large effect sizes. An a priori power analysis is another method for understanding the strengths of your data, but the discussion on how to conduct a power analysis is left for other resources. For Research Question #8, you again find yourself with a marginal group size (i.e., two groups of 30). You know also that when the data for each group resembles normality, the independent-sample *t*-test may be appropriate (i.e., after looking for outliers, missingness, data value gaps, etc.). Figure 3.4 shows histograms of the distributions for the two groups by filtering on the variable Received_MI. Recall that the process on how to filter data is discussed in Chapter 2, and it is discussed later in this chapter.

Neither of the two histograms shows a large departure from normality, providing additional support for the findings of the independent-samples *t*-test. In addition, the two nonparametric tests related to distribution comparisons were close to being significant, with **Sig.** values of .059 and .056. Thus, Dr. Rutledge has enough information to cautiously reject the null hypothesis and conclude that the MI intervention significantly changed the percentage of unsafe sexual behavior over a two-month period. However, when cautiously rejecting a null

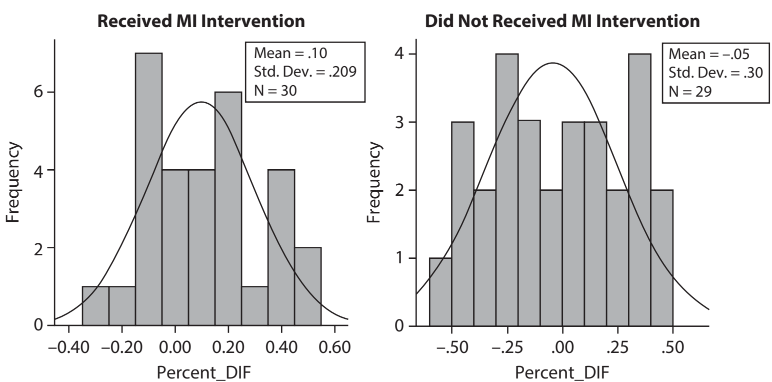


Figure 3.4. Histograms for Percent_DIF values for MI and non-MI groups.

hypothesis, it is very important that he explain to his readers that additional studies with larger samples need to be conducted before making a final determination about the intervention's effectiveness.

In statistical terms, both the Mann-Whitney U test (p = .059) and the Kolmogorov-Smirnov test (p = .056) for distribution comparison reject the null hypothesis of equal distributions at $\alpha < .10$. Additionally, results from conducting a parametric independent-sample t-test also reject the null hypothesis (p = .012)

Practically, the findings suggest that Dr. Rutledge's MI intervention influenced the respondents' unsafe sex behavior over a two-month period. After he divided his respondents into two groups, one group received a motivational intervention (MI) while the other group did not. He measured both groups on their unsafe sexual behavior using pre and post surveys. The analysis provided enough evidence to suggest that the MI intervention was effective in increasing the percent of safe sex activity for the group who received the intervention.

Box 3.4 Research Question #9

Based on the literature, Dr. Rutledge suspects that frequency of condom use is similar among three groups of participants — those who have exclusively same-sex sexual partners, those who have opposite-sex sexual partners, and those who have both same-sex and opposite-sex sexual partners. Dr. Rutledge asks you to test this for the participants in his treatment group, as the findings can be useful when analyzing the post-intervention data, and in determining if this will be a reasonable research question for his clinical trial.

Research Question #9

After participants in the treatment group receive the MI intervention, is condom use different among these participants based on who has exclusively same-sex, exclusively opposite-sex, and both same-sex and opposite-sex sexual partners?

KRUSKAL-WALLIS ONE-WAY ANOVA TEST

The parametric ANOVA analysis is frequently used to investigate group differences, but the independent groups, homogeneity of variances, and normality assumptions are sometimes very difficult to meet, especially when sample sizes are small. The Kruskal-Wallis one-way ANOVA test does assume independent groups but has no normality assumption. This distribution-free characteristic makes the test applicable to many data analyses and it is best suited for testing data that represent groups that are nominal, conceptually (i.e., the compared groups do not have a natural ordering). As in other nonparametric tests, the distribution-free characteristic comes with the cost of being more conservative, or less likely to find significant group differences. The Kruskal-Wallis one-way ANOVA test uses runs of values from two, three, or more independent groups to determine if the groups come from the same distribution. Of course, as the number of groups increases, so does the required size of the sample. The test pools the values from the groups to see how often the values switch from one group to the other to determine group differences.

Other examples for use of the Kruskal-Wallis one-way ANOVA test are as follows:

1. You conducted a pilot study for a larger research study about college drinking. You suspect that risky drinking is different among students in fraternities and sororities than it is among students who are not members, so you test this hypothesis with your small pilot sample before including those variables in your larger study.
2. You plan to conduct a comprehensive analysis of a secondary data set, but you are unsure about some of the relationships among your variables. You use a small subset of the data and place people in groups according to their socioeconomic status (SES) so that you can test if the level of resilience, as indicated by a standardized measure, is different among SES categories.

Box 3.5 Research Question #9, Continued

From the findings of the analysis conducted in Research Question #8, Dr. Rutledge has evidence to suggest that receiving an MI intervention changes the amount of condom use. Now he wants you to verify that the change from the intervention is not related to sexual partner choice. To do this, the first SPSS process to perform is to filter out the control group and then conduct the Kruskal Wallis one-way ANOVA analysis.

SPSS Process

To filter the data so that only respondents who received the MI intervention are analyzed,

- ☐ Select => **Data** => **Select cases**. . . to open the Select Cases window.
- ☐ Click to select the **If condition is satisfied** option in the **Select** area.
- ☐ Select => **If**. . . button to open the Select Cases: If window.
- ☐ Click to select the Received_MI variable on the left side of the window, then click the move arrow to move the variable to the white space to the right of the move arrow.
- ☐ Right after the variable name Received_MI in the white space, click to move the focus to right after the variable name.
- ☐ Type =1 (i.e., equals sign and the number 1). This tells SPSS to use cases that have a value of 1 for the variable Received_MI. Recall, a 1 represents respondents in the treatment group.
- ☐ Select => **Continue** button to close the Select Cases: If window.
- ☐ Select => **OK** button to close the Select Cases window and use only the cases that have a 1 for Received_MI.

To begin your analysis of the treatment group,

- ☐ Select => **Analyze** => **Nonparametric Tests** => **Independent Samples**.

Once the Nonparametric Tests: Two or More Independent Samples window opens, SPSS needs to know which variables to use in the analysis and which analysis routines to run.

- ☐ Select => **Fields** tab to see the list of available variables for the analyses.
- ☐ Click to select the **Use custom field assignments** option.

Note: If the Received_MI variable is in the **Groups:** area, move it back to the list of variables on the left. In addition, if the Percent_Pre variable is in the **Test Fields:** area, click to select the variable and then click the move arrow to move it back to the **Fields:** area.

- ☐ Click to select Partner variable in the list of variables on the left of the window.
- ☐ Click the move arrow closest to the **Groups:** area to move the variable to that area.
- ☐ Click to select the scale variable (Percent_Post) that contains percentage condom use post intervention.
- ☐ Click the move arrow closest to the **Test Fields:** area to move the variable to that area.

Now that SPSS knows what variables to use,

- ☐ Select => **Settings** tab to see the list of available nonparametric tests.
- ☐ Select => **Choose Tests** from the **Select an item:** area.
- ☐ Click to select **Customize tests** to enable the ability to select specific tests.

Figure 3.2 shows many options of nonparametric tests, some for two groups, and some for three or more groups, frequently referred to as k samples in published literature. A close review of Figure 3.2 shows three tests specific to having three or more groups.

- ☐ Click to select the **Kruskal-Wallis one-way ANOVA** check box on the right side of the **Compare Distribution across Groups** area. Make sure that none of the other test check boxes are checked.

Before running the test, you should consider the test option. For the Kruskal-Wallis test, the dropdown list for **Multiple comparisons:** should be set to **All pairwise** to compare the distributions of all the possible pairing of groups in the data. When the number of groups is too large, the comparison of all possible pairwise combinations becomes cumbersome, and the **Stepwise step-down** option is helpful. This option sorts the groups by their calculated group median value and compares only adjacent groups, making the comparison of groups possible. However, with only three groups, the comparison of all pairwise combinations is completely manageable.

☐ After reviewing your choices carefully, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click). The Model Viewer window will open to display the detailed information for the analysis (see Figure 3.5). As in the process discussed earlier, change the settings in the dropdown lists to find detailed information about the data and the analyses.

Findings

Figure 3.5 shows the results of the analyses for Research Question #9. The box-and-whiskers plot in Figure 3.5 shows that the median values among the three groups are very similar, but the group variability appears different. Specifically, the respondents who are exclusively with opposite-sex partners fall much closely around the median value, while the other two groups spread out more (i.e., whiskers indicate a wider range of values for exclusively same-sex and both same-sex and opposite-sex partner groups than for the exclusively opposite-sex partner group).

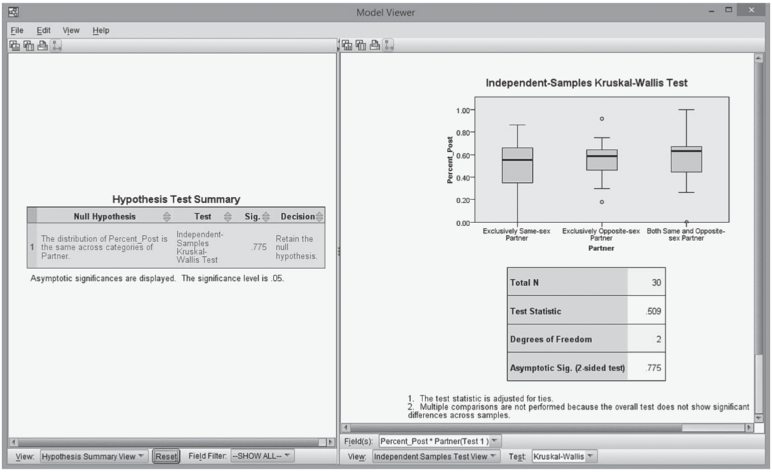


Figure 3.5. Model Viewer window with detailed test results.

Even though the box-and-whiskers plot in Figure 3.5 shows only a slight difference in Percent_Post median values along the horizontal axis for the three Partner groups, you still need to verify observed difference. Another way to examine the difference in median values is to review the calculated mean ranks for the three groups. To view the mean rank values, hover the mouse pointer over the black line in the box-and-whiskers plots. Hovering will cause a Mean Rank value to appear (14.25 for exclusively same-sex, 15.27 for exclusively opposite sex, and 17.43 for both).

The Kruskal-Wallis test did not find significant distribution differences among the three groups. The Kruskal-Wallis **Sig.** value = .775, larger than the traditionally used threshold value of $\alpha = .05$. This nonsignificant finding is graphically supported by the amount of similar vertical placement in the box-and-whiskers plots.

In statistical terms, the Kruskal-Wallis test for distribution comparison ($p = .775$) failed to reject the null hypothesis of equal distributions at $\alpha = .05$.

Practically, this means that the evidence suggests that the influence of an MI intervention is similar among groups of people who are exclusively with same-sex partners, exclusively with opposite-sex partners, and those who have both same and opposite-sex partners. Of course, another explanation is that the small sample simply does not provide enough information to identify differences. Therefore, adding the information from Research Question #8, we can infer that there is not enough information to conclude that the intervention benefits one group over the other.

Box 3.6 Research Question #10

Dr. Rutledge utilized Motivational Interviewing (MI) as the intervention in his feasibility study to reduce unprotected sexual behavior by moving participants through a variety of stages of change. The goal of MI is to facilitate the movement from the lowest stage, precontemplation, through the subsequent stages of contemplation, preparation, action, and maintenance. During the initial screening, Dr. Rutledge assessed his participants' stage of change, and found that all the participants in the treatment group fall into one of three stages of change — precontemplation, contemplation, and preparation for change. Dr. Rutledge suspects that the amount of change in condom use after the intervention differs among the groups based on the participants' stage of change, so he divided his sample into these three groups. He hypothesizes that he will find the lowest percent change in condom use among those in precontemplation, and the highest change in condom use among those in preparation. He asks you to examine this hypothesis, because if it is true, the use of MI to improve the participants' stage of change will be a valuable goal for the clinical trial.

Research Question #10

Is the percentage of change in condom use after receiving the MI intervention different for respondents who are in different stages of change (i.e., precontemplation, contemplation, and preparation)?

Because the grouping of respondents suggests an order, the test to use in this kind of situation is the Jonckheere-Terpstra test. Unlike the other nonparametric independent sample tests, the Jonckheere-Terpstra is better equipped to look for group differences when the groups have an order like the one in Research Question #10.

JONCKHEERE-TERPSTRA TEST

Like the Kruskal-Wallis test, the Jonckheere-Terpstra test uses runs to determine if the values from the groups come from the same distribution. However, the alternative hypothesis is that the distributions are somehow ordered, suggesting that the distributions are naturally decreasing or increasing among the groups. This difference from the Kruskal-Wallis test makes the Jonckheere-Terpstra more powerful in detecting distribution differences when the ordering characteristic is present. Similar to the Kruskal-Wallis test, the Jonckheere-Terpstra test assumes only independent groups and does not assume normality, which makes this test applicable whenever groups have some kind of order. Although the ordering of the groups does increase its strength when seeking group differences, it will still be more conservative than a one-way ANOVA when the data meet the normality assumption.

Other examples for use of the Jonckheere-Terpstra test are as follows:

1. As a supervisor in a small organization, you wonder if your professionals' negative behavior with clients is related to their caffeine intake. You know that measuring the number of cups of coffee is one way to examine this question, but you imagine that the size of their coffee cups is important to measure as well. You divide your workers into groups based on whether people drink from small, medium, and large cups of coffee, and test this with a standardized measure of behavior.
2. A personal trainer suspects that people who exercise a lot may express more narcissistic symptoms than those who don't exercise much or at all. To test this, she asks all her clients to fill out a questionnaire that includes a standardized measure of narcissism along with a measure of how frequently they exercise. She divides her trainees into three groups – those who exercise only occasionally, those who exercise one to three times a week, and those who exercise four or more time a week, to test her hypothesis.

The variables needed for the Jonckheere-Terpstra tests are a categorical ordinal variable representing three or more ordered groups, and a scale or continuous variable. For Research Question #10, the categorical variable is Stage_of_Change and the scale variable is Percent_DIF.

SPSS Process

Similar to the steps for Research Question #9, Research Question #10 involves analyzing only the data that represent those who were part of the treatment group. Therefore, the data filter used earlier on the Received_MI variable needs to be active before beginning the analysis. If you do not recall how to set the filter, see the first few steps under the SPSS Process for Research Question #9.

To begin your analysis,

- ☐ Select => **Analyze** => **Nonparametric Tests** => **Independent Samples**.

Once the Nonparametric Tests: Two or More Independent Samples window opens, SPSS needs to know which variables to use in the analysis and which analysis routines to run.

- ☐ Select => **Fields** tab to see the list of available variables for the analyses.
- ☐ Click to select the **Use custom field assignments** option.
- ☐ Click to select Stage_of_Change variable in the list of variables in the **Fields:** area.
- ☐ Click the move arrow closest to the **Groups:** area to move the variable to that area.
- ☐ Click to select Percent_DIF scale variable that contains percent change difference in condom use.
- ☐ Click the move arrow closest to the **Test Fields:** area to move the variable to that area.

Now that SPSS knows what variables to use,

- ☐ Select => **Settings** tab to see the list of available nonparametric tests.
- ☐ Select => **Choose Tests** from the **Select an item:** area.
- ☐ Click to select **Customize tests** to enable the ability to select specific tests.
- ☐ Click the check box for **Jonckheere-Terpstra for k samples** on the right side of the **Compare Distribution across Groups** area.
- ☐ Make sure that none of the other test check boxes are checked.

Before running the test, you should consider a few options. First, set the dropdown list for **Multiple comparisons:** to **All pairwise** to compare the distributions of all the possible pairings of groups in the data. To take advantage of the strengths of the Jonckheere-Terpstra test and increase the probability to finding significance among the groups when group differences exist, you will likely want to select the best option for **hypothesis order:**. The ordering refers to the theoretical or natural order of the groups identified by the categorical variable. Because the Jonckheere-Terpstra test uses an ordinal variable to identify groups, the ordinal variable places the groups in a sequence or specific order. While examining the groups in this specific order, you can compare the median values among the groups by running a Kruskal-Wallis test.

Note: If the medians are from smaller to larger values, set the **Hypothesis order:** dropdown option to **Smallest to largest**. If the medians are from larger to smaller values, set the dropdown option to **Largest to smallest**.

- ☐ For Research Question #10, click on **Hypothesis order:** dropdown and select **Smallest to largest**.
- ☐ After reviewing your choices carefully, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click). The Model Viewer window will open to display the detailed information for the analysis. As in the process discussed earlier, change the settings in the dropdown lists to find detailed information about the data and the analysis.

Findings

The results displayed in the Model Viewer window for the Jonckheere-Terpstra test did not indicate that the distributions were significantly different among the three groups (**Sig.** = .055). Because SPSS did not find significant differences, the multiple comparisons **All pairwise** option was ignored. However, a **Sig.** value of .055 is very close to the $\alpha = .05$ threshold. Considering the closeness of the **Sig.** value to .05 and your understanding that nonparametric tests tend to be more conservative in finding significance, a .055 may be enough to proceed further. But how?

One option is to adjust the significance level slightly so that SPSS will provide the pairwise comparison output, which in turn will give you more information for you to make a final determination about the test (despite inflating the experiment-wise Type I error rate). To adjust the significance level value,

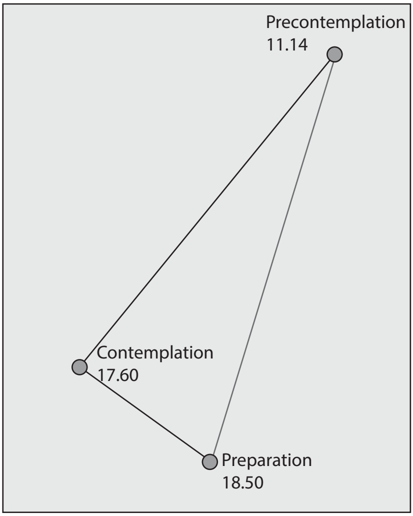
- ☐ Select => **Analyze** => **Nonparametric Tests** => **Independent Samples** to reopen the Nonparametric Tests: Two or More Independent Samples window.
- ☐ Select => **Settings** tab to see the options in the **Select an item:** area.
- ☐ Select => **Test Options** from the **Select an item:** area.
- ☐ Change the value displayed for the **Significance level:** area from 0.05 to 0.10.
- ☐ Select => **Run** button to rerun the analysis.

Now that SPSS recognizes that the **Sig.** value is smaller than the **Significance level** option value, the pairwise comparisons are conducted. Notice that although the significance level value has changed, the **Sig.** value remains the same. The reason is that the decision to set the significance level at a specific value is theoretical, and that decision affects only the choice of rejecting or failing to reject a test's null hypothesis. The chance that the three groups have similar means is still the same.

After double-clicking within the **Hypothesis Test Summary** area in the output window, and then changing the value of the dropdown for **View:** on the right side of the window to **Pairwise Comparisons**, you can see which comparisons are more different than others (see Figure 3.6).

Figure 3.6 shows that the group that is dissimilar to the other two is Precontemplation. In addition, the pairwise comparison that is most dissimilar is between Precontemplation and Preparation, with **Adj. Sig.** = .082. Like the **Sig.** value, the **Adj. Sig.** is used to decide when to

Pairwise Comparisons ...



Each node shows the sample average rank of Stage_of_Change.

Sample1-Sample2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
Precontemplation-Contemplation	78.000	14.090	1.632	.051	.154
Precontemplation-Preparation	74.500	13.018	1.920	.027	.082
Contemplation-Preparation	47.000	12.199	.164	.435	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (1-sided tests) are displayed. The significance level is .10.

Figure 3.6. Pairwise comparison graph of Stage of Change groups.

reject the null hypothesis based on the established significance level. The difference is that there is an adjustment to the value due to the increased likelihood of finding significance just because of conducting multiple tests. Recall, in this figure the reason that the pairwise is showing significance is that earlier you changed the significance level to .10, allowing SPSS to conduct the pairwise comparisons.

Because the initial test was very close to rejecting the null hypothesis and the review of the pairwise comparisons shows a clear difference between Precontemplation and Preparation, there is enough evidence to claim that respondents within this sample that are in different stages of change do, indeed, respond differently to the MI. The stage difference is seen in the percent change in condom use prior to and then after the intervention, with the largest difference being between those in precontemplation and preparation.

In statistical terms, the Jonckheere-Terpstra test (p = .055) rejects the null hypothesis of equal distributions at $\alpha = .10$. A pairwise comparison found adjusted p = .082 between Precontemplation and Preparation, adjusted p = .154 between Precontemplation and Contemplation, and adjusted p = 1.0 between Contemplation and Preparation.

Practically, this means that the evidence suggests that individuals in the precontemplation stage of change decrease their level of unsafe sexual behavior after experiencing an MI intervention, but less of a decrease than that of individuals in the preparation stage. This information makes theoretical sense, and it will greatly benefit the effort to improve the effectiveness of an MI intervention that targets decreasing the amount of unsafe sexual activity.

Box 3.7 Research Question #11

Improving participants' self-efficacy can be a goal for Motivational Interviewing, and Dr. Rutledge suspects that self-efficacy is related to the likelihood that participants will ask their sexual partners to use condoms. Based on their reports in the initial screening, the treatment group participants were divided into three groups — those unlikely to ask a partner to use a condom, those somewhat likely, and those very likely to request condom use. Participants also completed a validated measure of self-efficacy, so Dr. Rutledge requests that you assess if those unlikely to ask a partner to use a condom have lowest self-efficacy, and if those very likely to request condom use have highest self-efficacy scores. This finding can be useful when analyzing *the post-intervention treatment group, and it will support the use of these measures in the clinical trial.*

Research Question #11

Do the median values for self-efficacy differ among respondents who are unlikely to ask their partner to use condoms, those somewhat likely, and those very likely to request condom use?

MEDIAN TEST

The median test uses the frequency in which the values for each group are above and below a median value and then employs these counts to decide if the group medians are significantly different. In SPSS, the median value can be either a pooled value, in which SPSS calculates a median value using all the values from the groups, or a specified value identified by the person conducting the test. Unlike the parametric *t*-test that assumes a normal distribution for the groups, the median test has no assumption about the group's distribution. The median test does calculate a chi-square statistic based on expected vs. observed frequencies, so the test will identify significant median differences when the sample size is large. Note that "large" for three groups is a sample size of 300+ per group. The other median test assumption is that the groups are independent, so the values in one group do not have influence on the values in another group.

Other examples of uses for the median test are as follows:

1. You wish to know if the scale values from a depressive symptoms measure completed by the participants in your pilot study identify median differences for men and women.
2. You wish to verify that median values are not significantly different between subgroups within a sample in your feasibility study.

The variables needed for the median test are a categorical variable (i.e., nominal or ordinal) representing three or more groups of an independent variable, and a scale or continuous variable representing the dependent variable. For Research Question #11, the categorical variable is Condom_Request and the scale variable is Self_efficacy.

SPSS Process

Similar to Research Questions #9 and #10, Research Question #11 is analyzing only the data that represent those who were part of the treatment group. Therefore, the data filter for the Received_MI variable needs to be active before beginning the analysis. If you do not recall how to set the filter, see the first few steps under the SPSS Process for Research Question #9.

To begin your analysis,

- □ Select => **Analyze** => **Nonparametric Tests** => **Independent Samples**.

After the Nonparametric Tests: Two or More Independent Samples window opens, SPSS needs to know which variables to use in the analysis and which analysis routines to run.

- □ Select => **Fields** tab to see the list of available variables for the analyses.
- □ Click to select the **Use custom field assignments** option.
- □ Click to select the grouping variable Request_Condom.
- □ Click the move arrow closest to the **Groups:** area to move the variable to that area.
- □ Click to select the scale variable (Self_efficacy) that contains values representing levels of self-efficacy for the respondents.
- □ Click the move arrow closest to the **Test Fields:** area to move the variable to that area.

Now that SPSS knows what variables to use,

- □ Select => **Settings** tab to see the list of available nonparametric tests.
- □ Select => **Choose Tests** from the **Select an item:** area.
- □ Click to select the **Customize tests** option to enable the ability to select specific tests.
- □ Click to select the check box next to the **Median test** on the lower right side of the window in the **Compare Medians across Groups** area.
- □ Make sure that none of the other test check boxes are checked.

Before running the test, you must consider a few options. You must decide what median value to use when the median test is counting the larger and smaller values. If you do not have a specific value you identified from the literature or another source to use as a cut point, then it is best to allow SPSS to use a pooled sample median, which is the median value calculated from all the values in the groups. Therefore, for Research Question #11, you click to select **Pooled sample median**.

To test a specific median value, you would select the **Custom** option that enables you to enter a specific value in the **Median** area. To compare the medians of all the possible pairings of groups in the data, set the dropdown list for **Multiple comparisons:** to **All pairwise**. When the number of groups is too large, making the comparison of all possible pairwise combinations cumbersome, the **Stepwise step-down** option is helpful. This option sorts the groups by their calculated group median value and compares only adjacent groups, making the comparison of groups possible. However, with only three groups, the comparison of all pairwise combinations is completely manageable.

☐ After carefully reviewing your choices, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click). The Model Viewer window will open to display the detailed information for the analysis. Information for the **Hypothesis Test Summary** is displayed on the left, and box-and-whiskers plots on the right.

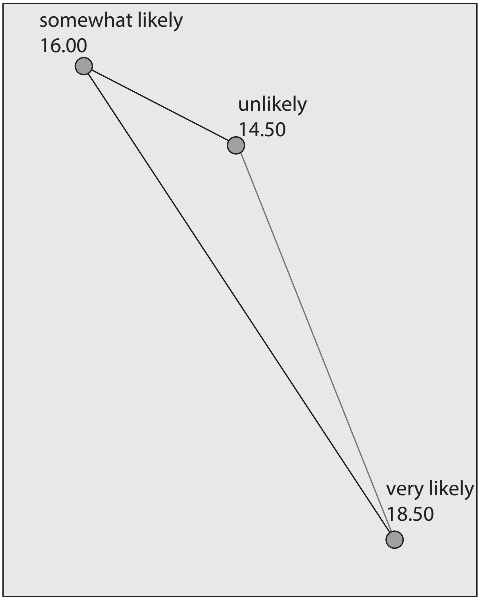
This information is for the median test and indicates enough information to conclude the median values for Self_efficacy are different across Condom_Request categories, because the **Sig.** value of .035 is smaller than the α = .05 significance level. To see the pairwise comparison results,

□ At the bottom right side of the Model Viewer window, click on the dropdown list named **View:** and select **Pairwise Comparisons**.

The information for the pairwise comparisons is displayed and shows **Sig.** and **Adj. Sig.** values for the three-category pairing (see Figure 3.7). The difference between **Sig.** and **Adj. Sig.** values is an important distinction. The **Sig.** values represent a straight comparison test of mean values without an adjustment for the multiple statistical tests that were conducted. However, the family-wise error rate when conducting multiple statistical tests within an analysis inflates the possibility of finding significant differences in median values for a category pair. The **Adj. Sig** values include an adjustment for this family-wise error rate and provide a more accurate view of testing median differences among the category pairs.

Findings

The overall median test indicates sufficient evidence to conclude that there is a difference in median self-efficacy across likelihood of requesting condom use categories (**Sig.** value = .035). The results of the pairwise tests show that median self-efficacy between respondents reporting they are *unlikely* to request a partner to use a condom and *very likely* to request a partner to use a condom is significantly different, with median values of 14.50 for *unlikely* and 18.50 for *very likely* categories.



Each node shows the sample median of Condom_Request.

Sample1-Sample2	Test Statistic	Sig.	Adj.Sig.
unlikely-somewhat likely	.486	.486	1.000
unlikely-very likely	6.600	.010	.031
somewhat likely-very likely	2.423	.120	.359

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

Figure 3.7. Pairwise comparison of Condom_Request.

In addition, the pairwise tests do not show a significant difference between *unlikely* and *somewhat likely*, and *somewhat likely* and *very likely* categories.

A review of Figure 3.7 shows that the *unlikely-very likely* pairwise comparison is significant with an adjusted p = .031, which is smaller than the α = .05 threshold. The other comparisons have adjusted significance values of 1.0 and .359, both larger than the significance threshold. Therefore, the median levels of self-efficacy are significantly different for only one of the three pairwise comparisons.

In statistical terms, the median test (p = .035) rejects the null hypothesis of equal median values at α = .05 for self-efficacy between three respondent categories on likelihood of requesting a partner to use a condom (*unlikely, somewhat likely*, and *very likely*). A pairwise test comparison identified one category pair (i.e., *unlikely* and *very likely*) that are significantly different (adjusted p = .031).

Practically, this means that levels of self-efficacy are different in Dr. Rutledge's sample for people grouped by their willingness to ask their partners to use a condom, and the findings provide support for further research into the association of self-efficacy with risky sex behavior.

4

Comparing Two or More Related Groups

CHAPTER OBJECTIVES

Chapter 4 presents information about comparing two or more groups that relate to one another. The chapter covers the following topics:

- Determining dependent (related) and independent (unrelated) groupings
- Selecting tests for two dependent groups (similar to the paired *t*-test for dependent samples)
- Comparing scale scores from pretests and posttests — Wilcoxon matched pair signed rank and sign tests
- Comparing dichotomous scores/responses for two related groups — McNemar's test
- Comparing ordinal (Likert-type) scores on two related groups — marginal homogeneity test
- Selecting tests for more than two dependent groups (similar to ANOVA and MANOVA)

- Comparing ordinal (dichotomous) scores on three or more related groups — Cochran's Q test
- Comparing scale scores on three or more related groups — Friedman's 2-way ANOVA by ranks test

INTRODUCTION

Chapter 4 covers information about comparing related samples and the way you analyze these data. Before beginning, if you are new to research, it might be helpful to clarify the use of the word *sample*. In this chapter, when comparing samples that are related, the term *samples* refers to collecting data multiple times from the same group of people, not to be confused with the use of the word *sample* to mean a subset of a target population (i.e., a sample of the population). It is very important to recognize when the term *sample* relates to a single data collection effort from one group of respondents, as described in Chapter 3, a sample from a population, and when the term *samples* relates to multiple data-collection activities from one group of respondents.

Prior to analyzing data for related samples, it is very important to make sure your data are identified with the appropriate level of measurement in SPSS and are in the appropriate format for related-samples kinds of tests. For example, in Research Question #12, the data format requires that responses for the two nominal variables for each respondent fall on the same row on the data spreadsheet, as they might for pre and post scores for the same respondent. This lets SPSS know that the pre and post scores are related to a single person or a single respondent.

As a reminder, to avoid any unexpected surprises or worse, report a finding not supported by the data, be thorough in your approach, and examine the information provided by each variable individually before moving forward with an analysis that involves multiple variables. This includes examining categorical variables to verify that the appropriate response coverage is present (i.e., a minimum of 5 responses for each category), and reviewing a graphical depiction of the distribution of scale variables to examine the theoretical/conceptual validity of the data in the variable. Chapter 1 describes a few things to consider when examining variables individually, and Chapter 2 provides multiple analytical procedures to assist the examination.

Like Chapters 2 and 3, Chapter 4 starts with a scenario describing a hypothetical line of research that we can explore with research questions. Box 4.1 presents the hypothetical research study for Chapter 4.

Box 4.1 Research Scenario for Chapter 4
Dr. Fleming's study is a secondary data analysis of a longitudinal study of youth in a major metropolitan area, and you are his statistician on the project. The original researchers utilized a systematic probability sampling strategy to collect data from public school students in four waves, beginning with children in third grade and surveying the students every three years. Dr. Fleming's interest was in the influence of gang membership on several outcomes, so he determined that data from all four waves — waves 1 (grade three), 2 (grade six), 3 (grade nine), and 4 (grade twelve) — would be suited for his needs. Because of attrition, the original sample size of 1000 decreased by 10% with each data collection, so wave 2 had 900 participants, wave 3 had 810 participants, and wave 4 had 729 participants. Of the 729 youth who responded to all four data collections, 3.4% (N = 25) reported that they eventually had joined a gang in one of the four waves of data collection and remained in the gang through wave 4.

Chapter 4 contains seven research questions to guide the presentation of the nonparametric procedures related to the above scenario. Each research question section describes the reasoning for using nonparametric procedures. Table 4.1 lists the nonparametric procedures presented in Chapter 4, the kinds of variables for each procedure, and the names of the variables found in the database for Chapter 4 that is available online.

MCNEMAR'S TEST

McNemar's test uses combination frequencies to determine if *yes/no* kinds of responses at two different times are different. The combinations are constructed from responses captured by two dichotomous variables (i.e., *yes-yes, yes-no, no-yes,* and *no-no*). The 2 × 2 contingency table structure of McNemar's test is similar to Pearson's chi-square test; however, the chi-square test is for independent samples (i.e., unpaired data) and not very effective when dealing with

Table 4.1. Summary of Analyses and Variables in Chapter 4

Research Question	*Nonparametric Test*	*Variable Measure*	*Variable Name in DB*
12	McNemar's	Two dichotomous variables	Joined_gang_w2 and Violence_crime_w2
13	Marginal homogeneity	Two categorical variables (i.e., ordinal variables representing two related responses, such as pre and post values)	Friends_w2 and Friends_w4
14	Sign test	Two continuous variables (i.e., paired scale variables representing two related responses, such as pre and post values for respondents)	Anxiety_w2 and Anxiety_w4
15	Wilcoxon matched-pair signed-rank Hodges-Lehmann confidence interval	Two continuous variables (i.e., paired scale variables representing two related responses, such as pre and post values for respondents)	Anxiety_w2 and Anxiety_w4
16	Cochran's Q	Three or more dichotomous variables	Assault_w2, Assault_w3, and Assault_w4
17	Kendall's coefficient of concordance	Three or more continuous variables (i.e., scale variables representing three or more related responses, such as pre-observation, second observation, third observation, etc., for respondents)	Support_w1, Support_w2, Support_w3, and Support_w4
18	Friedman's 2-way ANOVA by ranks	Three or more continuous variable (i.e., scale variables representing three or more related responses, such as pre-observation, second observation, third observation, etc., for respondents)	Support_w1, Support_w2, Support_w3, and Support_w4

Box 4.2 Research Question #12
Dr. Fleming is curious about the reasons youth join gangs, and previous studies suggest that youth who are victims of a violent crime may be more likely to join a gang for protection. To test this in his sample, Dr. Fleming wants you to compare responses to two questions in wave 2 — "Have you ever been the victim of a violent crime?" and "Have you joined a gang?" None of the respondents at wave 1 selected for this analysis was in a gang or had been a victim of a violent crime, thereby providing an opportunity to explore a relationship between violence and joining a gang.
Research Question #12
Does being a victim of a violent crime increase the probability that you will join a gang?

small sample sizes (i.e., less than 100 records). Another related test is the paired-sample *t*-test, but the paired-sample *t*-test is only for continuous variables that meet the normality assumption. Because McNemar's test uses two dichotomous variables, the null hypothesis is that the marginal probabilities for each of the two data point outcomes are equal. This means, for example, that the probability that a Hispanic/Latino person will answer *yes* to a specific question is equal to the probability that a non-Hispanic/Latino person will answer *yes* to the same question.

Other examples of uses for McNemar's test are as follows:

1. A researcher wants to know if a certain drug can keep people from getting the flu, so she conducts a small pilot study.
2. In a feasibility study for a larger research agenda, a researcher wants to test if those who regularly participate in risky drinking (also known as binge drinking) also regularly participate in risky sexual behaviors so he can decide if these variables are appropriate for the larger study.

The variables needed for the McNemar's tests for Question #12 are two dichotomous variables representing two related responses — e.g., pre and post values for respondents. For this question, the dichotomous variables are *victim of a violent crime* and *joined a gang in wave 2* (i.e., Violent_crime_w2 and Joined_gang_w2).

SPSS Process

To begin the analysis,

- ☐ Select => **Analyze** => **Nonparametric Tests** => **Related Samples**. . ..

After the Nonparametric Tests: Two or More Related Samples window opens, SPSS needs to know which variables to use in the analysis.

- ☐ Select => **Fields** tab.
- ☐ Click to select the **Use custom field assignments** option.
- ☐ Click to select the first variable, Violent_crime_w2, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.
- ☐ Click to select the second variable, Joined_gang_w2, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.

Now that SPSS knows what variables to use, you need to choose the kind of nonparametric test to conduct.

- ☐ Select => **Settings** tab.
- ☐ Select => **Choose Tests** found in the **Select an item:** area.
- ☐ Click to select **Customize tests** to show all the available tests for related samples (see Figure 4.1).

Figure 4.1 shows many different tests, some for two related observations, some for three or more related observations, some for categorical variables, and some for scale variables. A close examination of Figure 4.1 reveals four tests that are specific to having only two related observations. Because Research Question #12 involves binary data (i.e., two dichotomous variables), you should select McNemar's test.

- ☐ Click to select the checkbox next to the **McNemar's test** in the **Test for Change in Binary Data** area.

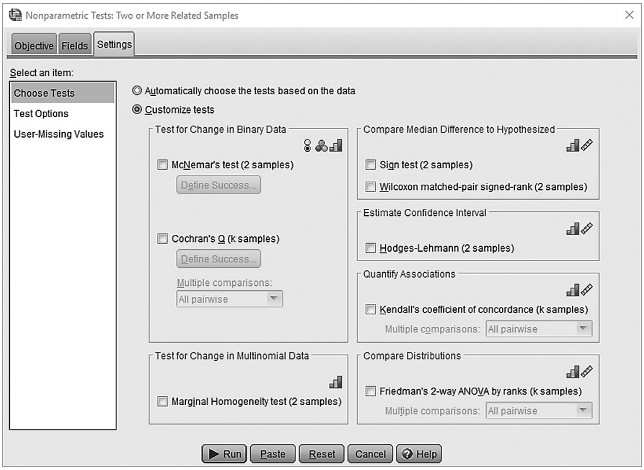


Figure 4.1. Nonparametric Tests: Two or More Related Samples Window.

Note: When you select **Choose Tests** in the **Select an item:** area, you will see one option that allows SPSS to decide automatically which tests are appropriate based on the variable definitions (i.e., identified by the level of measurement). It is prudent to monitor the analysis process closely by always choosing **Customize tests** and then selecting the specific test you want to conduct.

Before running McNemar's test, SPSS needs to know which value in the dichotomous variables is associated with *success*. SPSS uses the terminology of *success* and *failure* to identify the two options in dichotomous variables. You should keep track of the values identified as *success* to verify the use of your data and to understand how best to interpret the test results.

☐ Select => **Define Success**. . . to open the McNemar's Test: Define Success window.

Once the window opens, two options are displayed. The default option is to let SPSS use the **First value found in data**. To be fully aware of how the data are defined and used,

☐ Click to select the **Combine values into success category** option.

After making this selection, you gain access to the **Success:** area.

- ☐ Click on the white space below the word **Value** and enter a 1. Entering a 1 tells SPSS to associate the number 1 with a success.
- ☐ Select => **OK** to close the window.

Note: If the two variables had more than two outcomes that represent success, multiple lines in the white space identify the success outcomes. Recall, McNemar's test is based on dichotomous events. Having the capability to combine values expands the usability of McNemar's test to categorical variables with more than two values.

To continue your analysis,

- ☐ Select => **Test Options** in the **Select an item:** area to see where to change the default significance level of 0.05 for tests, the default 95.0% confidence interval range, and the default on how to handle missing data, which is **Exclude cases test-by-test**.
- ☐ After carefully reviewing your choices, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click). The

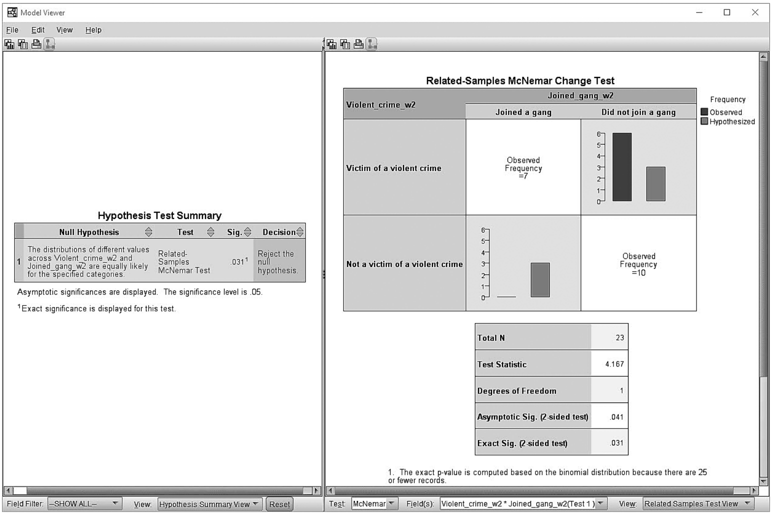


Figure 4.2. Model Viewer window for McNemar's test.

Model Viewer window will open to display the detailed information for the analysis (see Figure 4.2).

The Model Viewer window has access to more than the default view shown in Figure 4.2. At the bottom of the window are dropdown lists on the left for **View:** and **Field Filter:** and dropdown lists on the right for **Test:**, **Field(s):**, and **View:**.

- ☐ Changing the default **View:** on the right from **Related Samples Test** to **Categorical Field Information** is necessary so you can verify that you constructed the data in the analysis as intended — an important part of any analysis before providing any findings.

A review of the Model Viewer information provides a few interesting observations.

- **Total N** has a value of 23 even though the number of respondents for Research Question #12 was 25. This is due to the influence of missing data and the fact that **Exclude cases test-by-test** option was selected. The option for how to handle missing data is set under **Test Options** in the **Select an item:** area.

- The footnote on the right side of the Model Viewer window explains that a different statistical distribution testing for significant differences was used (i.e., Binomial) because there are fewer than 25 records. The footnote explains the use of counts in observed frequency rather than the number of respondents. Due to the use of the Binomial distribution, the analysis provides an exact p-value, giving the probability that the responses for the two variables are random.
- In addition to switching views by changing dropdown values, hovering the mouse over the bar charts shows additional information related to the test. Figure 4.2 displays information for combination of *victim of a violent crime* and *did not join a gang* if you hover the mouse over the area. The difference between the observed and expected values is used to determine if the frequencies for the combination of responses are not random, suggesting a relationship between the two variables.

Findings

The **Hypothesis Test Summary** information on the left side of the Model Viewer window reveals the exact significance of the test, $p = .031$. The exact significance values suggest only a 3.1% chance that being a victim of a violent crime is unrelated to the probability of joining a gang. The values in the bar charts in Figure 4.2 reveal that every respondent who joined a gang in wave 2 was a victim of a violent crime. Therefore, the data suggest that joining a gang is related to being a victim of a violent crime.

In statistical terms, McNemar's test rejects the null hypothesis of equal marginal probabilities at $\alpha = .05$ ($p = .031$). The small sample and use of the Binomial distribution in the statistical analysis provide an exact significance value ($p = .031$). The p-value indicates that if the null hypothesis is true, there is only a 3.1% chance of equal success probabilities.

Practically, this means that you found, in Dr. Fleming's sample of adolescents, a significant relationship between adolescents' joining a gang and having been a victim of a violent crime. Not explained in this analysis is if their reason for joining the gang is for protection.

Box 4.3 Research Question #13

Dr. Fleming has noticed that other researchers use a friendship measure as a proxy for the degree to which a youth is integrated in a gang. His data set includes an item that asks respondents, "How many of your friends are in a gang?" with response options of a few, many, *or* most. *He expects that, in his sample, the numbers of participants reporting* many *and* most *should increase over time, but he isn't certain. To test this, he asks you to compare the data from wave 2 and wave 4. If the numbers of gang friends increase as he suspects, this may be a valuable variable to use in subsequent analyses (e.g., investigating social support, peer influence).*

Research Question #13

Are the number of adolescents reporting that their friends are in a gang increasing over time?

MARGINAL HOMOGENEITY TEST

The marginal homogeneity test is an extension of McNemar's test in that it addresses two ordinal categorical variables rather than two dichotomous variables. The number of categories can be three or more, but, as described in Chapter 2, the more categories you have, the more responders you need to be sure each category has enough coverage to be used in an analysis. The marginal homogeneity test uses response combination frequencies to determine if the number of *positive* outcomes has changed, and because the test uses two categorical variables, the null hypothesis is that the marginal probabilities for each of the two data point outcomes are equal.

Other examples of uses for the marginal homogeneity test are as follows:

1. A clinician believes that an event in his therapy group increased the participants' level of anxiety. Prior to the group's first session, he had participants respond to an anxiety measure with 4-point Likert-like response options. He asked them to respond to the same measure after the event to learn if his participants' level of anxiety increased.
2. A social services administrator believes that she has improved her organization's services, and that clients' satisfaction should reflect this. At regular intervals, clients fill out satisfaction surveys

based on the following response options: *not satisfied, somewhat satisfied, mostly satisfied,* and *very satisfied.* She wants to compare earlier responses with the most recent responses to find out if the clients are more satisfied now.

For Research Question #13 (see Box 4.3), the categorical variables are Friends in wave 2 and wave 4 (Friends_w2 and Friends_w4). These variables identify the number of the respondent's friends who are in a gang, with categorical response options of *few, many,* and *most*. A review of the two categorical variables shows that coverage is just above the five responses per category guideline. However, although the data meet the coverage guideline, the frequency for each category is small, which does not provide much information for identifying differences. Nevertheless, when analyzing small datasets, your only option is to use the best test that matches your situation while making sure the data meet all the related assumptions.

SPSS Process

To begin the analysis,

☐ Select => **Analyze** => **Nonparametric Tests** => **Related Samples**.

Once the Nonparametric Tests: Two or More Related Samples window opens, SPSS needs to know which variables to use in the analysis.

☐ Select => **Fields** tab.
☐ Click to select **Use custom field assignments** option.
☐ Click to select the first variable, Friends_w2, found in the **Fields:** area.
☐ Click the move arrow to move the variable to the **Test Fields:** area.
☐ Click to select the second variable, Friends_w4, found in the **Fields:** area.
☐ Click the move arrow to move the variable to the **Test Fields:** area.

Now that SPSS knows what variables to use, the type of nonparametric test to conduct needs to be chosen.

- ☐ Select => **Settings** tab.
- ☐ Select => **Choose Tests** from the **Select an item:** area.
- ☐ Click to select **Customize tests** to show all the available tests for related samples (see Figure 4.1).
- ☐ Because Research Question #13 involves categorical data (i.e., two ordinal variables), click to select the checkbox next to the **Marginal Homogeneity test** in the **Test for Change in Multinomial Data** area.

Unlike the important step in Research Question #12 to identify values related to *success*, the marginal homogeneity test does not use a *success* distinction.

- ☐ After carefully reviewing your choices, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window.

To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click). The Model Viewer window will open to display the detailed information for the analysis. Use the dropdown lists at the bottom of the Model Viewer window to verify that the data used in the analysis are what you expected and supported theoretically. Hovering the mouse over the bar charts shows observed frequencies related to the test.

Findings

The **Hypothesis Test Summary** information on the left side of the Model Viewer window shows a **Sig.** value of .414. This **Sig.** value is larger than the α = .05 threshold for significance. Therefore, the data suggest that there is not enough information to determine that the number of friends in a gang has changed from wave 2 to wave 4. However, it is important to note that, with so little information for the combinations of categories, you would most likely need a sizable difference (effect size) for any statistical test to find significant differences. With so few respondents, you simply do not have enough power.

In statistical terms, the marginal homogeneity test fails to reject the null hypothesis of equal marginal probabilities at $\alpha = .05$ ($p = .414$).

Practically, this means that there is not enough information in the responses from the few adolescents to conclude that there is an increase in the number of reported friends in gangs over time.

Box 4.4 Research Question #14

Dr. Fleming has reviewed studies investigating anxiety among gang members. Some studies show that anxiety increases with the length of gang membership, while other analyses demonstrate that anxiety decreases over time. To investigate this question with his sample, Dr. Fleming asks you to compare participants' scores on a validated anxiety measure at wave 2 (i.e., early gang membership) with their scores on the anxiety measure at wave 4 (i.e., later gang membership).

Research Question #14

Does the amount of reported anxiety among gang members change over time?

SIGN TEST

The sign test uses the positive and negative differences for the matched pairs of values from two continuous variables to determine if the two sample medians have equal value. The null hypothesis is that the median of the calculated differences for the two samples is zero (i.e., the initial premise is that the median of one group is equal to the median of the second group). The equivalent parametric test is the related-sample *t*-test, but this test assumes that the difference scores are normally distributed in the population, whereas the sign test has no distribution assumption. The lack of a distribution assumption does make the sign test more applicable to a variety of data situations, especially with smaller samples for which you cannot verify the normality assumption. However, this also makes the test more conservative in finding significant differences between the two samples, meaning that the test is less likely to detect a statistically significant difference.

Other examples of uses for the sign test are as follows:

1. A researcher testing a new depression intervention on a small sample wishes to learn if his participants' depressive symptoms scale scores have changed from six months earlier.
2. A hospital administrator wishes to find out if the frequency of different diagnoses among his inpatient mental health unit has changed from the previous year.

The variables needed for the sign test are two paired scale or continuous variables representing two related responses, such as pre and post values for respondents. For Research Question #14 (see Box 4.4), the scale variables are Anxiety in waves 2 and 4 (i.e., Anxiety_w2 and Anxiety_w4).

SPSS Process

To begin the analysis,

☐ Select => **Analyze** => **Nonparametric Tests** => **Related Samples**.

Once the <u>Nonparametric Tests: Two or More Related Samples</u> window opens, SPSS needs to know which variables to use in the analysis.

☐ Select => **Fields** tab.
☐ Click to select **Use custom field assignments** option.
☐ Click to select the first variable, Anxiety_w2, found in the **Fields:** area.
☐ Click the move arrow to move the variable to the **Test Fields:** area.
☐ Click to select the second variable, Anxiety_w4, found in the **Fields:** area.
☐ Click the move arrow to move the variable to the **Test Fields:** area.

Now that SPSS knows what variables to use, the type of nonparametric test to conduct needs to be chosen.

- □ Select => **Settings** tab.
- □ Click to select **Choose tests** from the **Select an item:** area.
- □ Click to select **Customize tests** to show all the available tests for related samples (see Figure 4.1).
- □ Because Research Question #14 involves scale data (i.e., two continuous variables) click to select the checkbox related to the **Sign test** in the **Compare Median Difference to Hypothesized** area.

Unlike the step in Research Question #12 in which you identify values related to *success*, the sign test does not use a *success* distinction. Therefore, no options button is available other than the **Test Options** found in the **Select an item:** area.

- □ After carefully reviewing your choices, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click). The Model Viewer window will open to display the detailed information for the analysis.

- ☐ Use the dropdown lists at the bottom of the Model Viewer window to verify that the data used in the analysis are what you expected and are supported theoretically.

Findings

The histogram shows that nine of the paired values had a positive difference and 10 had a negative difference. In the total of 25 records, one pair had missing data and five pairs were equal. Note that the statistical calculations do not include pairs with zero differences. The number of positive-difference pairs to negative-difference pairs suggests that there is no information to indicate an increase in anxiety levels. In addition, the sign test shows a **Sig.** value (1.000) that is the largest significance value a test can return. Therefore, the sample of adolescents does not indicate any increase in anxiety from wave 2 to wave 4.

However, if you look at the histogram on the right side of the Model Viewer window (see Figure 4.3), the bars for the negative differences are in a very different pattern from the bars for the positive differences. This does suggest that something is happening related to anxiety levels.

Figure 4.3 reveals that the distribution of negative differences is much wider than the distribution of positive differences. Knowing that the sign test uses rankings and not difference values to test for equal medians, you might consider that another test may offer more insight into anxiety values over time (see Research Question #15).

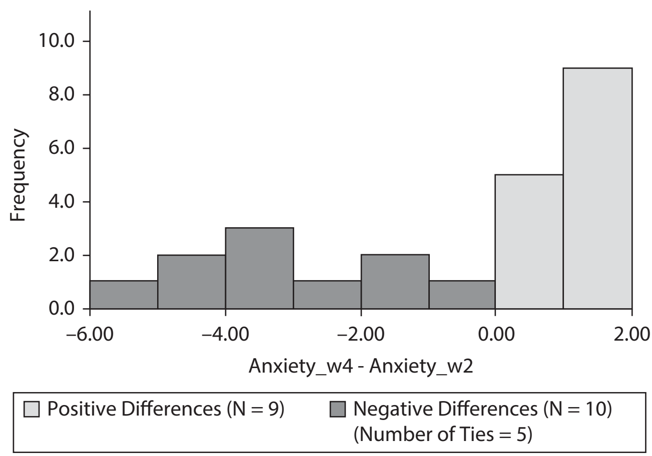


Figure 4.3. Histogram for anxiety levels for wave 2 and wave 4.

For Research Question #14, using statistical terms, the sign test fails to reject the hypothesis of equal medians at $\alpha = .05$ ($p = 1.00$), indicating that there is not enough information to indicate that the median anxiety levels for wave 2 are different from the median anxiety levels for wave 4.

Practically, this means that no information suggests that adolescent anxiety levels are changing. However, a review of the histogram in Figure 4.3 reveals that adolescents who report drops in their anxiety level experience a larger drop, on average, than the increase for adolescents reporting an increase in anxiety. Further analysis of these data is warranted to fully understand what the data suggest.

Box 4.5 Research Question #15

After reviewing the findings from Research Question #14 about anxiety in waves 2 and 4, Dr. Fleming wants to take the additional step of using a different statistical procedure more suited to variations in difference values between related observations. Thus, he requests that you return to SPSS to rerun the analysis using a different statistical procedure.

Research Question #15

Is the median difference in anxiety level between waves 2 and 4 equal to zero?

WILCOXON MATCHED-PAIR SIGNED-RANK TEST

The Wilcoxon matched-pair signed-rank test, also called Wilcoxon signed-rank test, is different from the Wilcoxon rank-sum test. The Wilcoxon matched-pair signed-rank test determines if median values from two samples are equal. The null hypothesis is that the median difference between related pairs is zero. Like the sign test in Research Question #14, the Wilcoxon matched-pair signed-rank test uses the frequency of positive and negative matched pairs value differences, but includes the ordering of these differences to test if median values are equal. This additional step makes the Wilcoxon test more powerful in detecting differences, but it does have an added assumption that you need to test. The assumption is that the median differences values are symmetrical about the median of the difference values.

Other examples of uses for the Wilcoxon matched-pair signed-rank test are as follows:

1. A researcher conducts a pilot study of a self-esteem intervention, and she wishes to test if the self-esteem scores immediately after the intervention are the same as when measured three months after the intervention ends.
2. A new manager believes he is doing well in his new role, but he wonders if the clients have any new concerns about his organization. He decides to examine the differences between the clients' satisfaction levels while his predecessor was in the managing role and their satisfaction levels now that he has been manager for six months.

HODGES-LEHMANN CONFIDENCE INTERVAL

The Hodges-Lehmann confidence interval is based on the Wilcoxon matched-pair signed-rank procedure that calculates an upper and lower limit of a confidence interval for median difference values. The confidence interval range is based on an alpha, typically of 0.05, representing a 95% confidence interval.

The variables needed for the Wilcoxon matched-pair signed-rank test and the Hodges-Lehmann confidence interval estimate are two continuous variables — paired scale variables representing two related responses, such as pre and post values for respondents. For Research Question #15 (see Box 4.5), the scale variables are Anxiety_w2 and Anxiety_w4.

Along with rerunning the analysis using the Wilcoxon matched-pair signed-rank test, Dr. Fleming wants a confidence interval. The confidence interval will provide a potential range in median differences between the two waves of data. In addition, knowing that the sign test **Sig.** value was 1.00, Dr. Fleming requests that you adjust the **Significance level** option from .05 to .10, because SPSS's Hodges-Lehmann procedure produces a confidence interval only if the test finds significance.

SPSS Process

To begin the analysis,

□ Select => **Analyze** => **Nonparametric Tests** => **Related Samples**.

To adjust the significance level value,

☐ Select => **Settings** tab to see the options in the **Select an item:** area.
☐ Select => **Test Options** from the **Select an item:** area.
☐ Change the value displayed for the **Significance level:** area from 0.05 to 0.10.

Next, SPSS needs to know which variables to use in the analysis.

☐ Select => **Fields** tab.
☐ Click to select the **Use custom field assignments** option.

If the anxiety variables are no longer in the **Test Fields:** area,

☐ Click to select the first variable, Anxiety_w2, found in the **Fields:** area.
☐ Click the move arrow to move the variable to the **Test Fields:** area.
☐ Click to select the second variable, Anxiety_w4, found in the **Fields:** area.
☐ Click the move arrow to move the variable to the **Test Fields:** area.

Now that SPSS knows what variables to use, the type of nonparametric test to conduct needs to be chosen.

☐ Select => **Settings** tab.
☐ Select => **Choose tests** from the **Select an item:** area.
☐ Click to select **Customize tests** to show all the available tests for related samples (see Figure 4.1).

Because Research Question #15 involves two scale/continuous variables,

☐ Click the checkbox next to the **Wilcoxon matched-pair signed-rank test** in the **Compare Median Difference to Hypothesized** area.
☐ Click the checkbox next to **Hodges-Lehmann** to request a confidence interval of the median difference between the two variables.

Unlike the step in Research Question #12 in which you identify values related to *success*, the Wilcoxon matched-pair signed-rank test does not use a *success* distinction.

☐ After carefully reviewing your choices, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click). The Model Viewer window will open to display the detailed information for the analysis (see Figure 4.4).

☐ Use the dropdown lists at the bottom of the Model Viewer window to verify that the data used in the analysis are as you expected and are conceptualized correctly for the analysis.

Findings

The histogram in Figure 4.4 is the same as the histogram in Figure 4.3. As before, the number of positive difference pairs to negative difference pairs suggests that there is no increase in median anxiety levels. However, instead of looking at the pairs, a review of the value differences suggests a different finding. The Wilcoxon signed-rank test finds

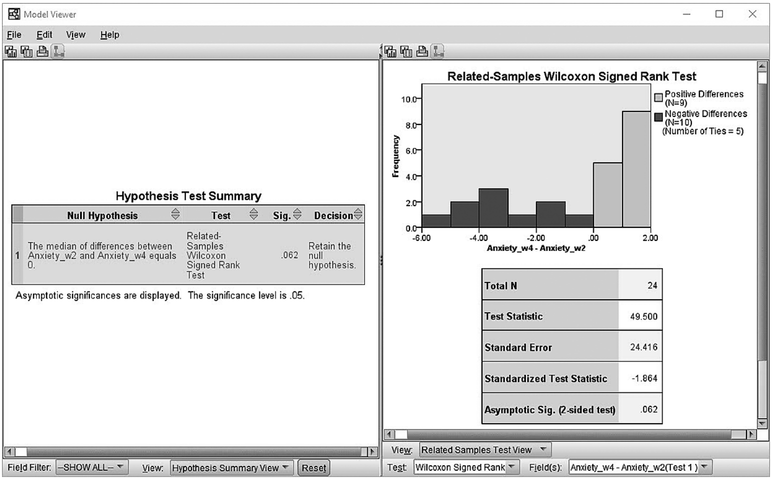


Figure 4.4. Model Viewer window for the Wilcoxon matched-pair test.

that median of the differences does not equal zero, indicating that there is a difference between anxiety levels for wave 2 and wave 4 (**Sig.** value of .062, which is smaller than $\alpha = .10$).

To view the upper and lower bounds of a 95% confidence interval for median difference values,

- ☐ Click on the dropdown list next to the **View:** area on the bottom left of the Model Viewer window.
- ☐ Click to change **Hypothesis Summary View** to the **Confidence Interval Summary View**.

After changing to the **Confidence Interval Summary View**, you will find the lower and upper values for a 95% confidence interval for the difference between the median values. The **Confidence Interval Summary** indicates that the Hodges-Lehmann procedure calculated a median difference estimate of –1.00 and a 95% confidence interval of –2.00 to 0.50. These results suggest that the median drop in reported adolescent anxiety levels from wave 2 to wave 4 is –1.00, and that the plausible range of median changes in anxiety is from –2.00 to 0.50.

Before reporting on the test results, you must verify that the data meet the median difference assumption for the Wilcoxon matched-pair signed-rank test. Recall, this assumption is that the distribution for the median difference values is symmetrical about the median.

To verify that the data meet the assumption, you need to generate a new variable from the two samples and review its distribution.

SPSS Process

- ☐ Select => **Transform** from the main SPSS menu.
- ☐ Select **Compute Variable**. . . to open the Compute Variable window.
- ☐ Enter Median_DIFF in the white space below the **Target Variable:** area.
- ☐ Click to select the variable Anxiety_w2 from the list of variables on the left side of the window.
- ☐ Click the move arrow to move the variable to the **Numeric Expression:** area.
- ☐ After Anxiety_w2 in the **Numeric Expression:** area, enter a space, minus sign, and a space.
- ☐ Click to select the variable Anxiety_w4 from the list of variables on the left side of the window.
- ☐ Click the move arrow to move the variable to the **Numeric Expression:** area.

The above steps tell SPSS to create a new variable called Median_DIFF by calculating the anxiety difference value for each respondent (i.e., Anxiety_w2 – Anxiety_w4)

- ☐ Select => **OK** to compute the new variable.

After the new variable is computed, you need to ask SPSS to graph the new computed values.

- ☐ Select => **Analyze** => **Descriptive Statistics** => **Frequencies** to open the Frequencies window.
- ☐ Click to select the new computed variable Median_DIFF from the list of variables on the left side of the window.
- ☐ Click the move arrow to move the variable to the **Variable(s):** area.
- ☐ Select => **Statistics**. . . to open the Frequencies: Statistics window.

- ☐ Click to select the **Median** option in the **Central Tendency** area.
- ☐ Select => **Continue** to close the window.
- ☐ Select => **Charts**. . . to open the Frequencies: Charts window.
- ☐ Click to select the **Histograms:** option in the **Chart Type** area.
- ☐ Click to select the **Show normal curve on histogram** option.
- ☐ Select => **Continue** to close the window.
- ☐ Select => **OK** to ask SPSS to create the graph of the distribution.

Note: The steps above include a request to draw a normal curve on the graph. The assumption you are attempting to verify refers to the symmetry about the median and NOT an assumption of normality. The reason for requesting SPSS to draw a normal curve is only to assist in visually interpreting the symmetry and NOT to compare the distribution against a normal curve.

SPSS Output

Figure 4.5 shows the graph of the computed variable Median_DIFF. A visual examination of the graph reveals a departure from symmetry

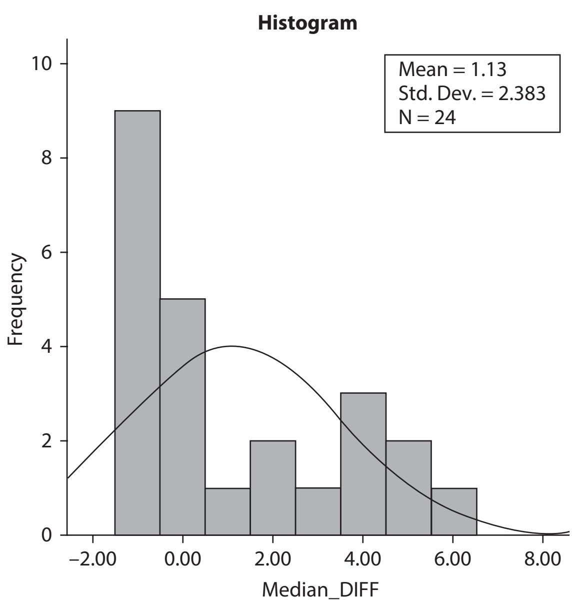


Figure 4.5. Distribution of value differences for two related samples.

about the median value of 0.00. Is the departure enough to determine that the data do not meet the assumption? This kind of question can be challenging, and this example is no exception. If you are certain that the data do not meet the assumption, then a different analytic process must be used to answer the research question. However, if you are uncertain, then use additional procedures to find out if they agree. In the case of the Wilcoxon matched-pair test, the sign test is a close relative and can give you additional insight into the data. The bottom line is that it is your responsibility as a researcher to gather enough statistical information to be comfortable with your results. Most of the time, this will include multiple statistical procedures that examine your data from different angles, because one statistical test rarely provides enough evidence for a conclusion. Even running the same statistical procedures on a subset of your data is more informative than a single stand-alone test.

In this situation, the distribution for Median_DIFF is not symmetrical about the median. Therefore, the relatively balanced number of values on each side of the median, the upper and lower values of the Hodges-Lehmann confidence interval, and the size of the *p*-value from the sign and Wilcoxon signed-rank test indicate that there is not enough evidence to conclude the two samples are different. If the distribution for Median_DIFF were closer to being symmetrical about the median, this decision would have been much more difficult to make. However, the skewness shown in Figure 4.5 and the significance level of the sign test at 1.00 are clear indications that you cannot reject the hypothesis of similar median values.

In statistical terms, the data do not meet the required assumption of a symmetrical distribution of difference values for the Wilcoxon signed-rank test. Therefore, the findings of significance at $\alpha = .10$ ($p = .062$) may not be valid. The Hodges-Lehmann confidence interval procedures provided an estimated mean difference of –1.00 and upper and lower bound values for a 95% CI for median difference of –2.00 to 0.50, which includes the possibility of a median difference value of 0, thereby failing to reject a hypothesis of equal median values. Recall, the confidence interval values provide a range in which we are 95% confident that the true median value falls between –2.00 and 0.50.

Practically, this means that you do not have enough information to suggest that adolescent anxiety has changed from wave 2 to wave 4.

Note: Do not forget to reset your **Significance level:** value from .10 to .05 if you prefer a more restrictive alpha level.

Box 4.6 Research Question #16
Dr. Fleming wonders if personal assaults committed by youth in his sample of gang members increase, decrease, or show no pattern across data collection waves. To investigate this, he asks you to analyze the participants' responses to the question, "Have you been involved in the commission of an assault since the last time we interviewed you?" for waves 2, 3, and 4.
Research Question #16
Is the frequency of committed assaults changing over time in this sample of gang members?

COCHRAN'S Q TEST

Cochran's Q test uses frequency of *success* and *failure* for k-related samples/observations to determine if the number of successes/failures is different among the samples. For Conchran's Q, success could represent testing positive on a medical test, answering yes to a yes/no question, or endorsing, rather than not endorsing, a candidate. Comparing samples to find potential differences is similar to repeated measures ANOVA, but with dichotomous variables. Remember that repeated measures ANOVA compares mean values and assumes normality for the sample values within groups. Before conducting Cochran's Q test, you must verify that you have enough data for all the responses to have enough information for a statistical test, and that the values are independent within groups (i.e., the relationships among the values exist only across groups/waves of data). The variables needed for Cochran's Q test are three or more dichotomous variables representing three or more related response values for respondents (e.g., observation 1, observation 2, observation 3). Because Cochran's Q test uses dichotomous variables, the null hypothesis is that the distribution of frequencies for each of the k data point outcomes is similar.

Other examples of uses for Cochran's Q test are as follows:

1. A researcher conducts a pilot study of a brief alcohol-related intervention for college students. To test the intervention's effectiveness after the conclusion of the intervention, he collects data on their drinking at three, six, and twelve months.
2. A researcher is training four research assistants to score a recorded interview, and she wishes to test if the raters

agree with each other or if they are more often not in agreement.

For Research Question #16 (see Box 4.6), the dichotomous variables are Committed an Assault for waves 2, 3, and 4 (i.e., Assault_w2, Assault_w3, and Assault_w4). These three variables represent three time points in which data were collected from adolescents, asking them if they had participated in an assault on another individual since the last time interviewed.

SPSS Process

To begin your analysis,

- ☐ Select => **Analyze** => **Nonparametric Tests** => **Related Samples**.

After the Nonparametric Tests: Two or More Related Samples window opens, SPSS needs to know which variables to use in the analysis.

- ☐ Select => **Fields** tab.
- ☐ Click to select the **Use custom field assignments** option.
- ☐ Click to select the first variable, Assault_w2, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.
- ☐ Click to select the second variable, Assault_w3, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.
- ☐ Click to select the third variable, Assault_w4, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.

Note: Another method to move variables from one side of the window to the other side is to point your mouse at the variable name and double-click. SPSS interprets the double-click as a message to move the variable to the opposite area.

Now that SPSS knows what variables to use, you must choose the kind of nonparametric test to conduct.

☐ Select => **Settings** tab.
☐ Select => **Choose Tests** from the **Select an item:** area.
☐ Click to select **Customize tests** to show all the available tests for related samples (see Figure 4.1).

Because Research Question #16 involves binary data (i.e., three dichotomous variables),

☐ Click to select the checkbox next to **Cochran's Q**.

Before running the Cochran's Q test, SPSS needs to know which value in the dichotomous variables are associated with a *success*.

☐ Select => **Define Success**. . . to gain access to the Cochran's Q Test: Define Success window.

After the window opens, two options are displayed. The default option is to let SPSS use the **First value found in data**.

☐ Click to select the **Combine values into success category** option to be fully aware of how the data are defined and used.

After making this selection, you gain access to the **Success:** area.

☐ Click on the white space below the word **Value** and enter a 1. Entering a 1 tells SPSS to associate the number 1 with a success.
☐ Select => **OK** to close the window.

Note: If the variables have more than two outcomes that represent success, then the success outcomes are identified on multiple lines in the white space. Recall, Cochran's Q test is based on dichotomous events. Having the capability to combine values expands the usability of Cochran's Q test to categorical variables with more than two values.

You must set one more option for the Cochran's Q test, the **Multiple comparisons:** option. The default is for **All pairwise** and is shown in the white area for the dropdown list under **Multiple comparisons:**. The other options are **Stepwise step-down** or **None**; click on the dropdown to show all three options. To compare the medians of all the possible waves in the data, set the dropdown list to **All pairwise**.

Note: When the number of comparisons is too large, requesting all possible pairwise combinations may be problematic due to the number of comparisons to review. When this is the case, the **Stepwise step-down** option may be a better choice. This option sorts the groups (e.g., waves of data) by their calculated group median value and compares only adjacent groups, making the comparison of groups manageable. However, with only three groups, the comparison of all pairwise combinations is completely manageable.

☐ After carefully reviewing your choices, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

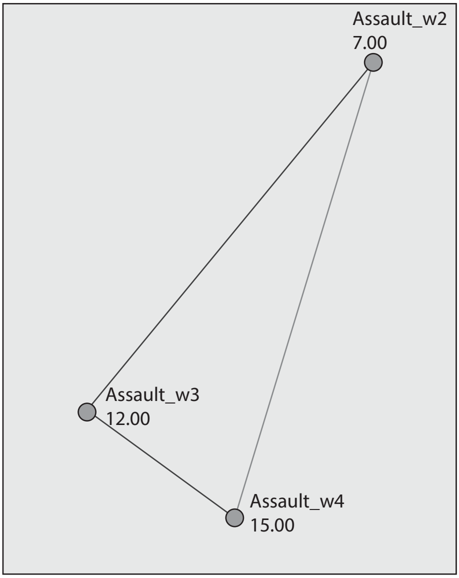
After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click). The Model Viewer window will open to display the detailed information for the analysis.

☐ Using the **View:** dropdown list at the bottom right of the Model Viewer window, select **Pairwise Comparisons** to display the information shown in Figure 4.6.

Pairwise Comparisons



Each node shows the sample number of successes.

Sample1-Sample2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
Assault_w2_Assault_w3	−.208	.136	−1.531	.126	.377
Assault_w2_Assault_w4	−.333	.136	−2.449	.014	.043
Assault_w3_Assault_w4	−.125	.136	−.919	.358	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

Figure 4.6. **Pairwise Comparisons** for Cochran's Q.

Findings

The **Adj. Sig.** values shown in Figure 4.6 indicate that the only significant pairwise difference for assault frequencies is from wave 2 (i.e., Assault_w2) to wave 4 (i.e., Assault_w4). The other pairwise differences are not significant (**Adj. Sig.** values of .377 and 1.00, both greater than the α = .05 threshold).

Hovering the mouse over the lines in the Model Viewer's **Pairwise Comparisons** graph will display the related **Adj. Sig.** values. Therefore, what you can draw from the **Adj. Sig.** values is that the number of respondents committing assaults increased at each wave and by wave 4 adolescents committing assaults had significantly increased from wave 2.

In statistical terms, Cochran's Q test ($p = .047$) rejects the null hypothesis that the number of respondents committing assaults has not changed at $\alpha = .05$ for the three waves of collecting data from the sample of gang members. In addition, a test of equal probabilities was conducted for all pairwise comparisons. The adjusted significance for multiple tests was calculated, and it revealed that a single pair — wave 2 and wave 4 — was significantly different ($p = .043$). The remaining pairs, wave 2 and wave 3 ($p = .377$), and wave 3 and wave 4 ($p = 1.00$) were not significant at $\alpha = .05$.

Practically, this means that as a group in the sample, the number participating in assaults significantly increased from wave 2 to wave 4. The frequency changes from wave 2 to 3 and wave 3 to 4 are not large enough to claim a significant difference. In addition, viewing the frequencies at each wave indicates the change is from fewer assaults in wave 2 to significantly more in wave 4. Another interesting relationship to investigate would be to look at joining a gang and committing assaults.

Box 4.7 Research Question #17

Dr. Fleming learned from the literature that gang members report staying in gangs because of the social support they offer. He hypothesized that this feeling of support would increase over time, as the youth became more integrated with their gang. To investigate this hypothesis, Dr. Fleming asks that you compare the participants' responses to a validated measure of social support at baseline (i.e., in grade three) with the responses at the subsequent three time points.

Research Question #17

Is the level of social support reported by the adolescents changing across the four waves of data collection?

KENDALL'S COEFFICIENT OF CONCORDANCE

Kendall's coefficient of concordance, also called Kendall's *W*, is a calculated value or coefficient that represents the amount of similarity among samples.

The value range for the coefficient is from 0 for no similarity to 1 for complete similarity. Kendall's coefficient requires three or more continuous/scale variables that represent three or more samples, in which the value (i.e., row) in the sample is somehow related to the same row in the other samples. The commonly used parametric procedures associated with Kendall's coefficient is the Pearson correlation coefficient, but the Pearson correlation coefficient assumes that the variables come from a normally distributed population of values, and it can compare only two variables in a test. As mentioned, Kendall's coefficient handles three or more variables and its only assumption is that the samples are dependent. Kendall's coefficient uses rankings and the number of ties to look for trends among the observations to determine if the trends in the ranks are different among the samples. The null hypothesis is that the distribution of the samples is the same.

Other examples of uses for Kendall's coefficient of concordance are as follows:

1. A researcher asks three or more raters to score a small group of research participants. Each rater (i.e., sample) scores each individual in a group (i.e., row in sample data). The researcher knows that the scores for each rater are related because they are assigning scores to the same group of participants. The researcher uses Kendall's coefficient of concordance because it provides a process to determine how much agreement exists among the raters.
2. An instructor conducts satisfaction surveys for three different classes that she is teaching. She wishes to know if the satisfaction scores for her three classes are similar or different.

For Research Question #17 (see Box 4.7), the scale variables are Levels of Support in waves 1, 2, 3, and 4 (i.e., Support_w1, Support_w2, Support_w3, and Support_w4), which represent the values recorded for social support for the four waves of data collections.

SPSS Process

To begin your analysis,

- □ Select => **Analyze** => **Nonparametric Tests** => **Related Samples**.

Once the Nonparametric Tests: Two or More Related Samples window opens, SPSS needs to know which variables to use in the analysis.

- ☐ Select => **Fields** tab.
- ☐ Click to select **Use custom field assignments** option.
- ☐ Click to select the first variable, Support_w1, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.
- ☐ Click to select the second variable, Support_w2, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.
- ☐ Click to select the third variable, Support_w3, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.
- ☐ Click to select the fourth variable, Support_w4, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.

Note: Another method to move variables from one side of the window to the other side is to point your mouse at the variable name and double-click. SPSS interprets the double-click as a message to move the variable to the opposite area.

Now that SPSS knows what variables to use, you need to choose the type of nonparametric test to conduct.

- ☐ Select => **Settings** tab.
- ☐ Select => **Choose Tests** from the **Select an items:** area.
- ☐ Click to select **Customize tests** to show all the available tests for related samples (see Figure 4.1).
- ☐ Because Research Question #17 involves scale data (i.e., continuous variables), click to select the checkbox next to the **Kendall's coefficient of concordance** procedure.

Unlike the step in Research Question #12 in which you identified values related to *success*, Kendall's coefficient does not use a *success* distinction. Therefore, no buttons to select options are available. However, you can change the **Multiple comparisons** option. The default is **All pairwise,** shown in the white area for the dropdown list under **Multiple comparisons:** for the test. The other options for **Multiple comparisons:** are **Stepwise step-down** or **None**. When the number of groups is too large, requesting the comparison of all possible pairwise combinations can be cumbersome. The **Stepwise step-down** option is helpful in these situations. This option sorts the groups by their calculated group median value and compares only adjacent groups, making the comparison of groups possible. However, with only three groups, the comparison of all pairwise combinations is completely manageable.

□ After carefully reviewing your choices, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click). The Model Viewer window will open to display the detailed information for the analysis.

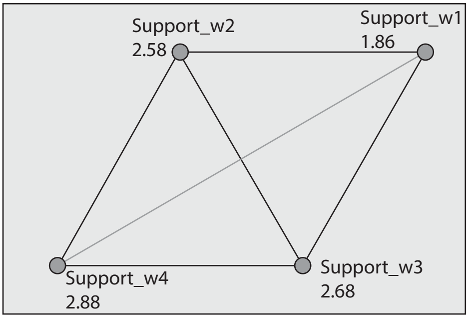
To change the detailed information displayed in the Model Viewer window, you change the selected values in the dropdown lists at the bottom of the window.

☐ Click the dropdown list for the **View:** dropdown list at the bottom right of the Model Viewer window and select **Pairwise Comparisons** to display the information shown in Figure 4.7.

Findings

The **Adj. Sig.** values shown in Figure 4.7 indicate that one of the six pairwise comparisons is significant — Support_w1 vs. Support_w4 (p =.031).

Pairwise Comparisons



Each node shows the sample average rank.

Sample1-Sample2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
Support_w1_Support_w2	−.720	.365	−1.972	.049	.292
Support_w1_Support_w3	−.820	.365	−2.246	.025	.148
Support_w1_Support_w4	−1.020	.365	−2.793	.005	.031
Support_w2_Support_w3	−.100	.365	−.274	.784	1.000
Support_w2_Support_w4	−.300	.365	−.822	.411	1.000
Support_w3_Support_w4	−.200	.365	−.548	.584	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

Figure 4.7. **Pairwise Comparisons** for Kendall's coefficient of concordance.

The other five pairwise comparisons have **Adj. Sig**. values larger than the **Significance level:** set in **Test Options**. Hovering the mouse over the lines in the **Pairwise Comparisons** graph shown in the Model Viewer window will cause the related **Adj. Sig.** values to display. Therefore, from the **Adj. Sig.** values, you can conclude that the level of support reported by the adolescents is significantly different from wave 1 to wave 4. If you review the descriptive information for the four variables, you will see that reported support levels are larger in wave 4 than in wave 1, indicating that the difference is an increase in reported support. In addition, you can see this increase in the **Pairwise Comparisons** graph under each variable name.

In statistical terms, Kendall's coefficient of concordance ($p < .016$) at $\alpha = .05$ rejected the null hypothesis of similar observed trends for the four samples. This analysis says that if the null hypothesis is true, there is less than a 1.6% chance that all four trends are similar. The test of equal responses was conducted for all pairwise comparisons. Results of the statistical analysis identify only one pair that is significantly different (Support_w1 and Support_w4), with adjusted $p = .031$.

Practically, this means that reported support levels are changing over time for the adolescents, and that a significant increase in the support levels exists for the adolescents in the sample between wave 1 and wave 4. Understanding why the support levels increase over time will require additional data collection and analyses.

Box 4.8 Research Question #18

Now that Dr. Fleming has learned from his previous analysis that support levels are not remaining the same over time for the adolescents in the study, he wonders if, at each wave, the reported levels are similarly distributed. You recognize that, to investigate this, you can utilize another nonparametric procedure to examine social support in the same four data collection waves, this time comparing distributions rather than changes over time.

Research Question #18

Are the reported support levels for each wave similarly distributed?

FRIEDMAN'S TWO-WAY ANOVA BY RANKS TEST

Friedman's two-way ANOVA is a statistical procedure that tests if related samples come from the same population. A ranking strategy, along with a calculated mean rank for each sample, is used to draw a conclusion about the null hypothesis that the sample distributions are similar. Because the comparable related-samples ANOVA has a normality assumption that is difficult to meet and sometimes to verify when samples are small, you can use the Friedman's test in a wider range of data situations. Both Kendall's coefficient of concordance and Friedman's two-way ANOVA test examine sample similarities. However, Friedman's test is calculating the possibility that the samples all come from one population, and Kendall's coefficient is looking only for similarities among the samples, a slight but recognizable difference when selecting which procedures to use when analyzing your data. The variables needed for Friedman's two-way ANOVA test are three or more continuous/scale variables representing three or more related responses (e.g., pre-observation, second observation, third observation, fourth observation for respondents).

Other examples of uses for the Friedman's two-way ANOVA test are as follows:

1. A college professor teaches a statistics course, and he wishes to determine if his students' test scores on three different quizzes are comparable.
2. A researcher preparing a large-scale research proposal wonders if the time of day will have an influence on how participants report their depressive symptoms. She measures depressive symptoms in the participants in her feasibility study in the early morning, late morning, early afternoon, and late afternoon.

For Research Question #18 (see Box 4.8), the scale variables are reported Support Levels at waves 1, 2, 3, and 4 (i.e., Support_w1, Support_w2, Support_w3, and Support_w4).

SPSS Process

To begin your analysis,

☐ Select => **Analyze** => **Nonparametric Tests** => **Related Samples**.

If you just completed the analysis for Research Question #17, the variables may still be in the **Test Fields:** area. If they are, skip down to the *Note:* below. If they are not, SPSS needs to know which variables to use in the analysis.

- ☐ Select => **Fields** tab.
- ☐ Click to select **Use custom field assignments** option.
- ☐ Click to select the first variable, Support_w1, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.
- ☐ Click to select the second variable, Support_w2, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.
- ☐ Click to select the third variable, Support_w3, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.
- ☐ Click to select the fourth variable, Support_w4, found in the **Fields:** area.
- ☐ Click the move arrow to move the variable to the **Test Fields:** area.

Note: Another method to move variables from one side of the window to the other side is to point your mouse at the variable name and double-click. SPSS interprets the double-click as a message to move the variable to opposite area.

Now that SPSS knows what variables to use, you need to choose the type of nonparametric test to conduct.

- ☐ Select => **Settings** tab.
- ☐ Select => **Choose Tests** from the **Select an items:** area.
- ☐ Click to select **Customize tests** to show all the available tests for related samples (see Figure 4.1).

Because Research Question #18 involves scale data (i.e., continuous variables),

☐ Click the checkbox next to the **Friedman's 2-way ANOVA by ranks** procedure.

Unlike the step in Research Question #12 in which you identified values related to *success*, Friedman's test does not use a *success* distinction. Therefore, no buttons to select options are available. However, you can change the **Multiple comparisons** option. The default is **All pairwise,** shown in the white area for the dropdown list under **Multiple comparisons:** for the test. The other options for **Multiple comparisons:** are **Stepwise step-down** or **None**. When the number of groups is too large, requesting the comparison of all possible pairwise combinations can be cumbersome. The **Stepwise step-down** option is helpful in these situations. This option sorts the groups by their calculated group median value and compares only adjacent groups, making the comparison of groups possible. However, with only four groups, the comparison of all pairwise combinations is completely manageable.

☐ After carefully reviewing your choices, select => **Run** to conduct the analysis.

See Appendix A for complete SPSS syntax.

SPSS Output

After you click **Run**, a variety of information shows up on SPSS's output window, including:

- Syntax used by SPSS (i.e., log)
- Location of the data set used (i.e., Active Dataset)
- **Hypothesis Test Summary** (i.e., Model Viewer)
- List of items in the output index on the left side of the output window

Clicking on an item in the list initiates a scroll to that information on the right side of the output window. To see the details of the test, double-click within the area outlined by the Model Viewer (i.e., point the mouse at the **Hypothesis Test Summary** information and double-click).

Within the Model Viewer window, on the right side, you will see the distributions for each sample and their individual **Mean Rank** score. Also on the right side, you will see the **Total N** count for the number of responders, the **Test Statistic**, the **Degrees of Freedom**, and the **Asymptotic Sig. (2-sided test)** value. You can use this information in the test to decide if the distributions are dissimilar enough to reject the possibility that they come from the same population.

Findings

The **Asymptotic Sig.** value (.016) tells us to reject the possibility that they come from the same population. A review of the horizontal histograms in the Model Viewer window clearly reveals different distributions for the four variables because the shapes of the histograms are different. However, there is more analysis information to review than just the significance test. To see the descriptive information and the distribution for the individual waves, change the selected option in the dropdown list for **View:** at the bottom of the window.

- ☐ Click the **View:** dropdown list at the bottom of the Model Viewer window.
- ☐ Select **Continuous Field Information** to display the overall distribution for the samples.
- ☐ Click the dropdown list of variables **Field(s):** to change the selected variable and the display from one variable to another.

To see the results of the pairwise comparisons, change the selected option in the dropdown lists for **View:** at the bottom of the window.

- ☐ Click the **View:** dropdown list at the bottom right of the Model Viewer window.
- ☐ Select **Pairwise Comparisons** to display the pairwise information.

In statistical terms, Friedman's Two-way ANOVA test ($p < .016$) rejected the null hypothesis at $\alpha = .05$ of similar distributions. The test of equal distributions was conducted for all pairwise comparisons. Results of the statistical analysis identify only one pair that is significantly different (Support_w1 and Support_w4), with adjusted $p = .031$.

Practically, this means not only that the support levels from one wave to another increased over time, but also the reported levels clustered differently. More adolescents reported lower levels of support in wave 1 and more reported higher levels in wave 4.

5

Predicting with Multiple Independent Variables

CHAPTER OBJECTIVES

This chapter presents nonparametric models that allow for prediction of a dependent variable based on values for a set of independent variables. The chapter covers the following:

- Nonparametric regression (LOWESS)
- Nonlinear regression
- Factorial design regression

INTRODUCTION

Chapter 5 covers prediction using both dependent and independent variables in a nonparametric model. Similar to parametric models like linear regression, nonparametric models are available that allow for prediction of a dependent variable based on values for a set of independent variables. As in the earlier chapters, the reason for selecting a nonparametric procedure is because the data do not meet the assumptions of the traditional parametric procedures. For linear regression models, the assumptions involve linearity, independence, homoscedasticity, and normality.

You may recall that, in your beginning statistics course, the instructor showed you two scatter plots on an X-Y axis that included an estimated regression line. One of the scatter plots showed dots following relatively close to a straight line, while the other plot showed dots that followed a curved line. The curved line was an example of a nonlinear relationship between two variables (X and Y). If you are in this kind of situation with your own data, you can consider transforming the X variable to see if linearity is possible (see Chapter 6 in Chatterjee & Hadi, 2012, for information on transformations). Alternatively, you can consider using nonparametric regression that does not have the linearity assumption (i.e., the assumption that the relationship among the variables can be represented by a linear combination of parameters).

The independence assumption, when violated, means that the model is more than likely not fitting the data or that the *p*-values are smaller. In other words, the errors in the model have a pattern or trend (e.g., serial correlation). One way to test the independence assumption (i.e., check for serial correlation) is to conduct the Durbin-Watson test, which is a test available in the SPSS list of statistics options on the Linear Regression: Statistics window. A YouTube video that offers additional explanation can be found at https://www.youtube.com/watch?v=Lyk_S9T-oCw (Grande, 2015). Homoscedasticity is another assumption for linear regression, and it requires no relationship between the residuals and predicted values. To check for homoscedasticity, you can plot standardized residuals against standardized predicted values to look for patterns. SPSS provides options for requesting different combinations of prediction and residuals plots to test for linearity, homoscedasticity, and normality.

Normality is another necessary assumption when performing parametric regression. Violations of normality can occur due to the presence of a few extreme outliers that cause the distribution of the error terms to be skewed, not representing a normal distribution. Variables in the model that do not represent normal distributions also have the potential to cause a violation of normality. Thus, it is important to examine the distribution of residuals for your regression model.

You must explore all potential violations of assumptions before running a parametric test. This is not to say that nonparametric regressions are exempt from variable examination. On the contrary, you must closely examine every variable so you can make an informed selection of the appropriate statistical test. When you are faced with the need to conduct a regression and the parametric assumptions cannot be met by transformation, by dropping cases with inappropriate outliers, etc., the following nonparametric regression procedures are your next option.

Note: Over time, many people have asked IBM-SPSS to add a nonparametric regression test. Unfortunately, at the time we wrote this chapter, SPSS did not offer a menu-driven option for conducting a nonparametric regression analysis. Therefore, we use a second software package to conduct some of the statistical tests in Chapter 5. The software package, XLSTAT, and is an add-in to Excel (see www.xlstat.com for more information on XLSTAT).

Another note: R^2 is the amount of variance explained in your regression model. For social science research, in general, R^2 values are smaller than in other disciplines due to the complexity of topics/issues under investigation. For example, consider the complexity and number of different issues that may be involved in understanding burnout or self-efficacy. R^2 values in the range of 20% to 30% are common, while values above 60% are extremely rare. However, if find that your analysis does not explain much of the variation, then there is something your model does not capture. Departures from what you hypothesize are always a good indication to verify the statistical assumptions.

Chapter 5 describes situations in which a researcher wishes to predict a value or level based on a set of independent variables, but the data do not meet the assumptions for least-squares regressions. Following the same format used in the previous chapters, a representative scenario related to prediction is presented, along with research questions that can be answered by using nonparametric methods for regression.

Box 5.1 Scenario for Chapter 5

Dr. Northcott's area of study is adoption and foster care, and she also provides training for practitioners working in this field. Recently, in response to comments she received in several of her training sessions, Dr. Northcott decided to conduct *a survey about the personal and occupational stressors that may influence the quality of these social workers' practice. After assembling her questions, receiving comments from expert content and measurement reviewers, and revising her items, she was ready to pilot test her questionnaire. She sent a questionnaire to every k^th^ person on the membership list of a professional organization for adoption and foster care workers, following up with reminder postcards and duplicate questionnaires at regular intervals, as recommended in the survey research literature (e.g., Dillman, 2014). She mailed questionnaires to 50 professionals and achieved a 70% response rate (N = 35). However, eight professionals did not answer all the questions, reducing her usable sample to 27. Dr. Northcott included many variables for the study, knowing from the literature that both personal and occupational issues might be relevant. A sample size of 27 does not provide sufficient power to conduct the planned regression analyses, nor does it meet the guidelines that states 10 to 15 cases are needed for each independent variable in the regression model. She hires you to assist her in answering her questions.*

Chapter 5 contains three research questions that guide the presentation of the nonparametric procedures related to the above scenario. Each research question provides the details for using the nonparametric procedures. Table 5.1 lists the nonparametric procedures presented in Chapter 5, the kinds of variables for each procedure, and the names of the variables found in the database for Chapter 5 that is available online.

Table 5.1. Summary of Analysis and Variables in Chapter 5

Research Question	*Nonparametric Test*	*Variable Measure*	*Variable Name in DB*
19	Nonparametric regression	IV — categorical or continuous DV — continuous	Impairment, Stressors, Burnout, Pred_Stressors, and Pred_Burnout
20	Nonlinear regression	IV & DV — categorical (ordinal) or continuous	Edu_Level and Depression
21	Factorial design	IV — categorical DV — continuous	LifeSatisfaction, Gender, and Employment

Box 5.2 Research Question #19

Dr. Northcott's training participants reported that the reason the quality of their work was suffering was because they were dealing with too many stressors at work and that they were experiencing burnout. Dr. Northcott created a professional impairment instrument that encompassed several measures of impaired practice (e.g., missing appointments, missing days from work, employer disciplinary action). She asks you to find a regression model that could predict level of professional impairment from standardized measures of burnout and work stressors.

Research Question #19

Can levels of burnout and work stressors predict levels of professional impairment?

ROBUST LOWESS REGRESSION

Robust LOWESS regression is a nonparametric regression modeling process available in XLSTAT that you can use when your data do not meet the assumptions for linearity, independence, homoscedasticity, and normality for the comparable parametric modeling process, multiple regression. Robust LOWESS regression, like all nonparametric regression models, does not involve any assumptions about normality. However, unlike multiple regression, which can examine individual variable effects on a dependent variable and make predictions based on a set of independent variables, the robust LOWESS can be used for prediction only. LOWESS stands for LOcally WEighted regression and Smoothing Scatter plots (Cleveland, 1979; Cleveland & Devlin, 1988). Robust LOWESS is an updated version. Simply stated, the process uses a proportion of the available data to begin estimating prediction values for the dependent variable, and through many computations identifies a prediction model. Calculations for smoothing are related to softening any sharp changes in the prediction model due to gaps between data points that naturally occur in scale variables, thereby creating a smoother prediction line that can be viewed in graphical form.

Other examples of nonparametric regression are as follows:

1. A researcher wishes to know if measures of self-efficacy and gender predict levels of life satisfaction in his small clinical sample.

2. The researcher above also wants to know if demographics are related to depressive symptoms in his clinical sample. Specifically, he wants to know if current employment, age, and socioeconomic status predict any part of depressive symptoms.

For Research Question #19 (see Box 5.2), three variables are used in the analysis — two continuous independent variables, Burnout and Stressors, and one continuous dependent variable, Impairment.

Before beginning a nonparametric procedure, you want to check if the data do, indeed, violate parametric assumptions, which would prevent you from using the customary linear regression. This is especially true when choosing between linear regression and nonparametric regression. The major objective when building a nonparametric regression statistical model is to find the model that best fits your data. Thus, unlike linear regression, nonparametric regression does not provide beta values that you can use to interpret variables' contribution to the model. The reason lies in the technique used to estimate the requested model for prediction. Therefore, viewing the results from a nonparametric regression analysis is very different, and the process of identifying the best model with the available variables is a bit more repetitive, as it requires manually moving variables in and out of the model while monitoring R^2 and prediction charts.

The first step is to examine the variables individually to fully understand the quality and type of information they contain. Begin by running frequency distributions on each so you can review, among other things, levels of skewness and kurtosis. For Research Question #19, the distributions of both Burnout and Stressors are left-skewed and the distribution of Impairment is somewhat bimodal (see Figure 5.1).

SPSS Process

To obtain frequency distributions,

- □ Select => **Analyze** => **Descriptive Statistics** => **Frequencies**. . ..
- □ Click to select Impairment in the list of variables on the left side of the window.
- □ Click the move arrow  to move the variable to the **Variable(s):** area.

- ☐ Click to select Burnout in the list of variables on the left side of the window.
- ☐ Click the move arrow to move the variable to the **Variable(s):** area.
- ☐ Click to select Stressors in the list of variables on the left side of the window.
- ☐ Click the move arrow to move the variable to the **Variable(s):** area.

Note: Another method to move variables from one side of the window to the other side is to point your mouse at the variable name and double-click. SPSS interprets the double-click as a message to move the variable to opposite area.

Now that SPSS knows what variables to use in the descriptive analysis,

- ☐ Select => **Charts**. . . to open the Frequencies: Charts window.
- ☐ Click to select the **Histograms:** option.
- ☐ Click to select the checkbox next to **Show normal curve on histogram**.
- ☐ Select => **Continue** to close the window.
- ☐ Select => **OK** to run the analysis, which will open the Model Viewer window.

See Appendix A for SPSS syntax.

SPSS Output

The histograms in Figure 5.1 are displayed in the Model Viewer window.

If these variables were normally distributed, you would continue assessing the assumptions. However, given the variables' departure from normality and the small data set (N = 27), using parametric regression is not an option. In addition, one of the cases has a missing value (ID 11), and XLSTAT requires you either to drop the case from the analysis or to select a missing value replacement option (e.g., mean, nearest neighbor replacement). Because of the repetitive nature of conducting nonparametric regression, it is probably best to

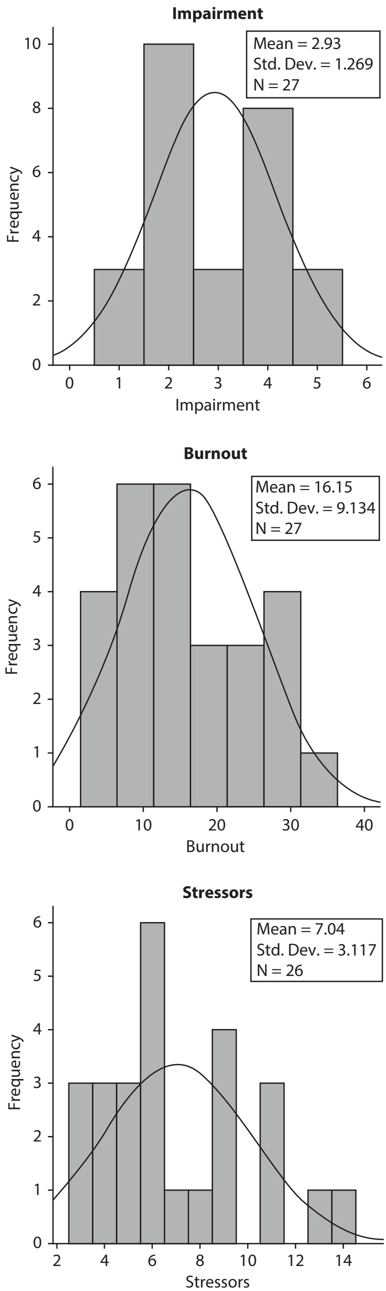


Figure 5.1. Histograms of continuous variables.

examine model fit first by dropping case(s) with missingness. Next, compare the model fit when using the different replacement options. If the model fit is appropriate, the best option is to choose the model that dropped case(s), as this model will avoid adding bias into your prediction values. See Appendix B for more information on handling missing values.

ANALYZING THE DATA IN XLSTAT

Prior to running an analysis, you must arrange the data to meet the requirements of the software you are using. XLSTAT expects that you have arranged the variables in columns, and that each row represents a unique case or respondent. Therefore, for this example, you should arrange values for Burnout, Stressors, and Impairment in three separate columns. SPSS uses the same layout for the data as XLSTAT, with one important difference — how to identify missing values. SPSS provides a method in which to name multiple values as identifiers for missingness, but XLSTAT identifies missing values only wherever a blank or empty cell exists.

After installing XLSTAT on your PC, you need to make XLSTAT active. To activate XLSTAT within Excel,

- ☐ Select => **ADD-INS** tab at the top of the Excel window.
- ☐ Click to select the **ToolBar Commands** button to open the XLSTAT toolbar.

After you click on the button, Excel will take a few minutes to bring up the toolbar and add a new tab named XLSTAT to the list of tabs at the top of the Excel window.

XLSTAT Process

The nonparametric regression analysis to run is found under the **Modeling Data** button.

- ☐ Select => **XLSTAT tab** in the Excel menu options at the top of the Excel window.

- ☐ Select => the **Modeling Data** button to display the list of analyses associated with modeling data.
- ☐ From the list, select => **Nonparametric regression**, which opens the Nonparametric regression window.

At the top of the Nonparametric regression window are seven tabs (**General, Options, Validation, Prediction, Missing data, Outputs**, and **Charts**). The information and selections you provide within each of these tabs tell XLSTAT how to conduct the nonparametric regression (see Figure 5.2). For robust LOWESS, XLSTAT dims some of the options to indicate that this specific nonparametric regression model does not use them.

To begin the analysis,

- ☐ Select => **General** tab.
- ☐ Be sure to check the option **Variable labels** on the right side of the window if variable names appear in the first row of the Excel spreadsheet above the data.
- ☐ Click on the white space for the field name **Y/Dependent variables:**, right below the word **Quantitative:**.

By clicking in the white space, you draw focus to that area, which allows for the identification of the dependent variables. The dependent variable in the spreadsheet for Impairment is in column B, so find that column on the spreadsheet and

- ☐ Click on the **B** at the top of the column.

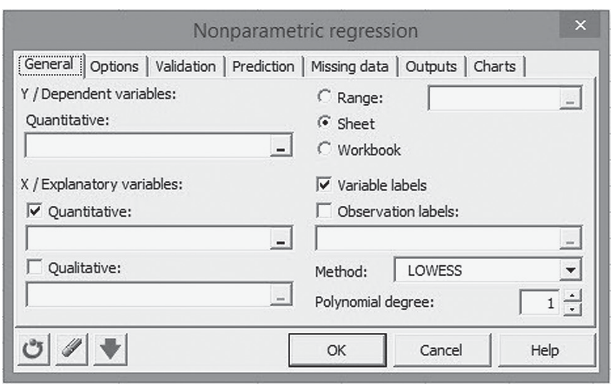


Figure 5.2. Nonparametric regression window.

When you click on the **B**, you tell XLSTAT where to find the values for the dependent variable, Impairment. Your click will both highlight the column in the traditional way Excel selects a column and add text to the **Y/Dependent variables:** field (i.e., Data!$B:$B). Data!$B:$B tells XLSTAT to use all the data in that column for your dependent variable.

- ☐ Click in the white space for the **X/Explanatory variables:** field just below **Quantitative:**, drawing focus to that area.

You can find the independent or explanatory variables in our example in columns **C** and **D**.

- ☐ While holding down the "Ctrl" key on the keyboard, click and drag the mouse from **C** to **D** at the top of the two columns.

Doing this tells XLSTAT which data to use for the independent variables and will place text in the field to identify the two columns (i.e., Data!$C:$D).

Note: Using the mouse to select the data for the analysis is only one method of telling XLSTAT what to use. If you want to avoid using the mouse, you can type the information directly into the fields that represent the dependent and independent data.

Because this research question does not include categorical data, the **Qualitative** checkbox should be unchecked. Next,

- ☐ Click to select the **Sheet** option to tell XLSTAT to place the analysis results in its own Excel sheet.
- ☐ Click to select the dropdown list for **Method:** to select **Robust LOWESS** from the dropdown list.

Robust LOWESS uses a form of curve smoothing in a scatter plot to identify the relationship between the dependent and independent variables.

The other options within this dropdown are for other nonparametric regression model procedures. Selection of a model option is based on the kind of data you are modeling and can become quite complicated.

For Research Question #19, and for many data situations, the robust LOWESS will meet the prediction requirements.

Note: For more information on the above options available, including more information on model procedures, select => **Help** button at the lower right side of the Nonparametric regression window. Selecting **Help** will open a separate window named HTML Help that provides access to help information for XLSTAT and available procedures.

To proceed with the analysis,

- □ The default for the **Polynomial degree:** value is 1 (one). If it is not 1, click to change the **Polynomial degree:** value back to 1.

The reason for this is that the first time you run your model, you start with a linear polynomial (i.e., value of 1). Starting with 1 in this example tells XLSTAT not to investigate the squared values of Burnout or Stressors for predicting Impairment. Once you review the fit indices, you can return to this option and try higher levels of polynomials to see if model fit can be improved, as long as increasing the polynomial makes theoretical sense. For example, does adding the squared values of Stressor (polynomial = 2) improve model fit? However, keep in mind that variables with influential outliers can suggest a curvilinear relationship, especially with small samples. This is another example of how important it is to understand the data thoroughly before building regression models or testing hypotheses.

As you explore all the options provided by XLSTAT for nonparametric regression, you will recognize many different analytic possibilities. However, for this analysis most of the default options fit with Research Question #19.

- □ Select => **Options** tab to see where in the Nonparametric regression window you can change the proportion of the data used to estimate prediction values (i.e., proportion of the sample for "learning").

"Learning" is how XLSTAT describes the process of building a model that fits the data. XLSTAT will use some of the data to build the model and other data to predict model fit. The default is 50% and is a good place to start, but if you are experiencing poor model fit, you may adjust this option up or down to see if you can improve the fit.

Another option is **Tolerance:**, which has as a default to check and set to .01. This option tells XLSTAT how much a variable can be out of conformity with the estimated model before it is ignored altogether. If you want the variables to be more in line with the model, you can decrease the default of .01. In other words, if a variable is not contributing enough to the model, XLSTAT will decide to move on without it. For Research Question #19, we accept the default.

The final option is the **Interactions/Level:**, which is not checked as a default. If you want to include variable interactions in your model, you can check this box. For Research Question #19, we did not include variable interactions in the model.

Note: Dimmed options on the window indicate that they are not used by the robust LOWESS procedure.

- ☐ Select => **Validation** tab if a subset of your data set is to be used to validate the model.

With larger data sets, validation is an option to further validate the fit of a model. This is like splitting a data set so you can conduct an exploratory factor analysis (EFA) on one half and then validate the model with a confirmatory factor analysis (CFA) using the other half. Unfortunately, with smaller data sets, taking some of the records to validate the estimated model is impractical. This is the case for Research Question #19; therefore, due to the limited number of records, you should leave the **Validation** option unchecked.

- ☐ Select => **Missing data** tab to verify that cases with missing data will be dropped.

Dropping the cases with missing data is the safest way to avoid adding unrecognized bias. If you replace the missing values, you introduce the possibility of unnaturally increasing the power of the model to predict the independent variable, and it may not be a true reflection of the concepts you are analyzing.

- ☐ Select => **Outputs** tab to verify that all four options are selected.

For Research Question #19, we want to request all the information available to help understand the data and the model.

- ☐ Select => **Charts** tab to request that the analysis output includes charts.

Be sure that the **Data and predictions** option is checked and that the **As a function of X1** option is selected. In addition, be sure to select the **Residuals** option to add residual information to the analysis output.

Note: It is important that you set these options as described above for both the **Outputs** and **Charts** tabs, because these options provide a clear graphical view into the quality of the data in the model as well as model fit (e.g., predicted vs. observed).

Now that the options are set,

- ☐ Select => **OK** button at the bottom of the window to execute the analysis.

Once **OK** is selected, a final verification window (i.e., XLSTAT — Selections) appears. This gives you one last look at the data and options selected for the analysis. After reviewing your choices carefully,

- ☐ Select => **Continue** button and a new Excel sheet will appear with all the requested analysis output.

Note: If some analysis information is missing, or if XLSTAT recognizes a conflict in your analysis request, an XLSTAT — Message window will appear with an explanation of the issue instead of the XLSTAT — Selections window.

XLSTAT OUTPUT — VERIFYING MODEL FIT

Remember that the major objective when building a nonparametric regression statistical model is to find the model that best fits your data.

The search for the best-fitting model involves trying different model settings (i.e., options) while monitoring the model fit information. Before using any results from a statistical model, you must assess model fit to make sure your model represents your data. If you are familiar with factor analysis or latent class analysis (LCA), this exploratory approach will sound familiar.

The first step in an examination of nonparametric regression model fit in XLSTAT is to consider the amount of variance for the dependent variables explained by the independent variables, identified as the determination coefficient R^2. R^2 is part of the **Goodness of fit statistics:** information found on the new Excel tab at the bottom of the Excel window (see Figure 5.3). R^2 is a calculated value that falls between 0 and 1. Zero represents the worst possible fit, in which the data do not follow the model at all, and 1 represents a model that supports every data point exactly.

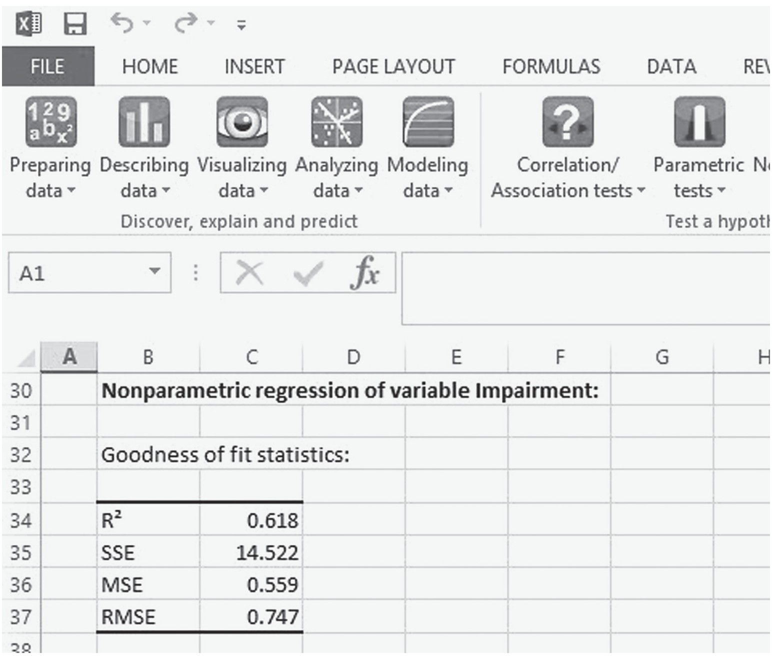


Figure 5.3. Fit indices.

R^2 for Impairment in Research Question #19 is 0.618, which indicates that 62% of the observed variability in Impairment is explained by Burnout and Stressors. Figure 5.3 shows that the robust LOWESS results provide three other fit statistics — SSE (sum of squares of the errors), MSE (means of the squares of the errors), and RMSE (root mean squares of the errors). All four of the fit indices relate in some way to one other. For example, RMSE is the square root of MSE, but because of their interpretation, the two most often reported are R^2 and RMSE. RMSE is the standard deviation of the value differences between the data and the model, thereby providing information on the variations of differences between the two. Because the best-fitting models are those with the smallest errors and that explain the most observed variability, we can use R^2 and RMSE values to identify which model has the best fit.

Note: Because R^2 will, in almost all cases, increase as you add more variables to the model, whether or not the added variables are significant predictors of the dependent variable, R^2 alone is insufficient in determining fit. In addition to following RMSE, a close review of variable information is important. Following the variables individually, as discussed in Chapter 1, is just as important when building prediction models.

So just how much variation (i.e., R^2) needs to be explained before the model can be used? Unfortunately, the response to this question is that it depends on the research question and the specific analysis. For nonlinear regression in which you have the flexibility to try different models (i.e., functions), it comes down to the variables in the study, to how much variation other similar studies have identified, and what would be considered enough based on theoretical understanding of the issues being measured. No matter how you make the decision, the responsibility falls on you, the researcher, to justify that the variance amount explained is enough before sharing your findings with other researchers.

Another important responsibility is to make sure you have eliminated other model options and to verify that you have found the appropriate model for your data. This is an iterative process, so, as mentioned earlier, this includes trying a few other XLSTAT options prior to running the analysis. Another way to eliminate other models is to try other combinations of independent variables that have theoretical support for predicting the dependent variable — in this case, Impairment. In addition, exploring other possible models will help you understand your data at a deeper level, give you an opportunity to build confidence that you have selected the right variables for your final model, and help you draw a conclusion that you feel comfortable sharing with other researchers.

In addition to presenting fit indices, XLSTAT for LOWESS regression provides graphs to examine model fit (see Figures 5.4, 5.5, and 5.6). The first graph, Figure 5.4, is a scatter plot of the model prediction for the relationship between Burnout and Impairment as a function of X1. Burnout was used because it was the first listed independent variable (i.e., X1). If you want to see a graph for Stressors and Impairment, you must switch the order of the independent variables and rerun the analysis. The second graph, Figure 5.5, is a scatter plot of predicted vs. observed values for Impairment. Because LOWESS assumes that the errors for a model are random, an appropriate fitting model should show random error values on the predicted vs. observed plot. Nonrandom patterns along the diagonal line may suggest poor model fit to the data. The third graph, Figure 5.6, is the residuals bar chart showing the amount of error/ residual for each case. Residuals are the differences between observed

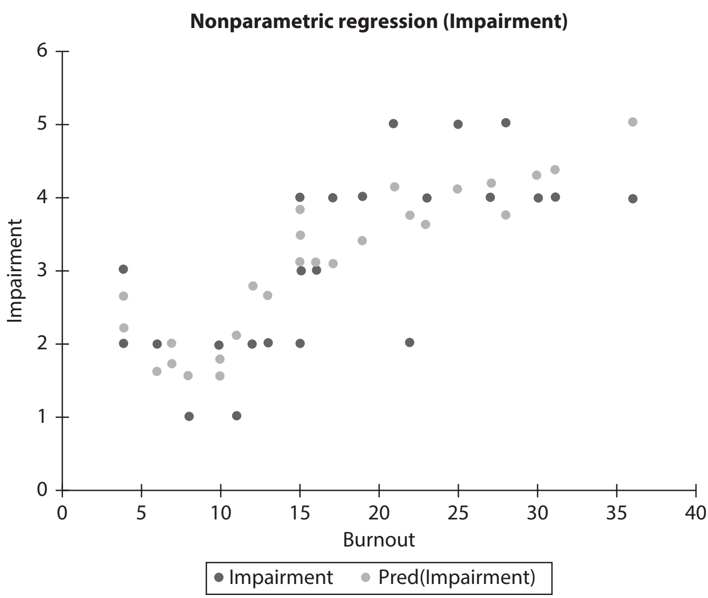


Figure 5.4. Fit assessment graph as a function of X1.

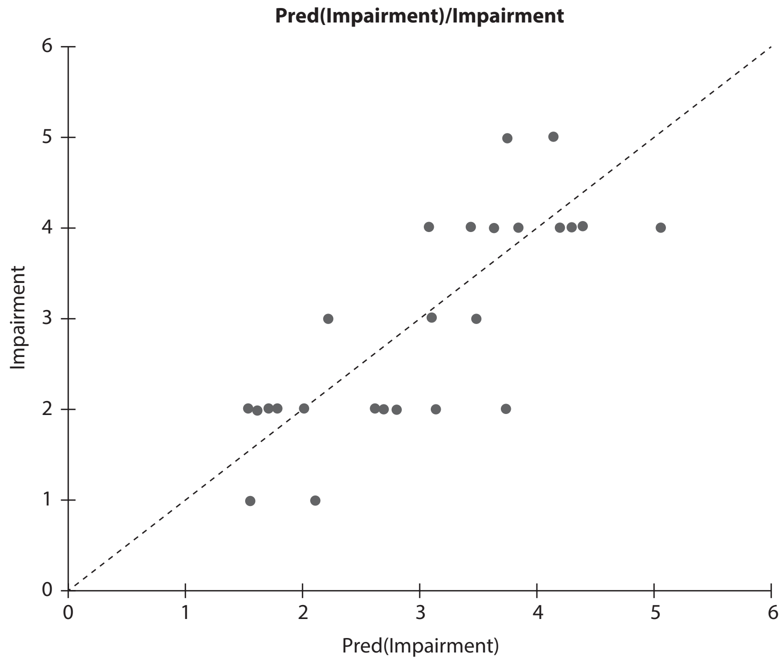


Figure 5.5. Fit assessment graph for predicted vs. observed.

and model prediction values. Like the scatter plot, histograms that show a pattern as you review the bars from left to right suggest the possibility of a poor-fitting model.

In addition, the bar chart provides the opportunity to identify cases that have unusually large error (i.e., a bar on the chart that is much larger than other bars). In these cases, you need to verify that you do not a have a data-entry problem and that the case does follow all the original guidelines for inclusion in the target population. In extreme situations, you may remove the case if the case clearly does not belong theoretically or conceptually, but we do not recommend this practice unless you have irrefutable information to support the fact that the respondent is outside of the parameters of the study.

Fortunately, for Research Question #19, the scatter plot in Figure 5.4 does not show a pattern along an envisioned diagonal line. In addition, the bars in Figure 5.6 appear relatively random and do not indicate a pattern or problem with cases' being dependent. Therefore, you can be

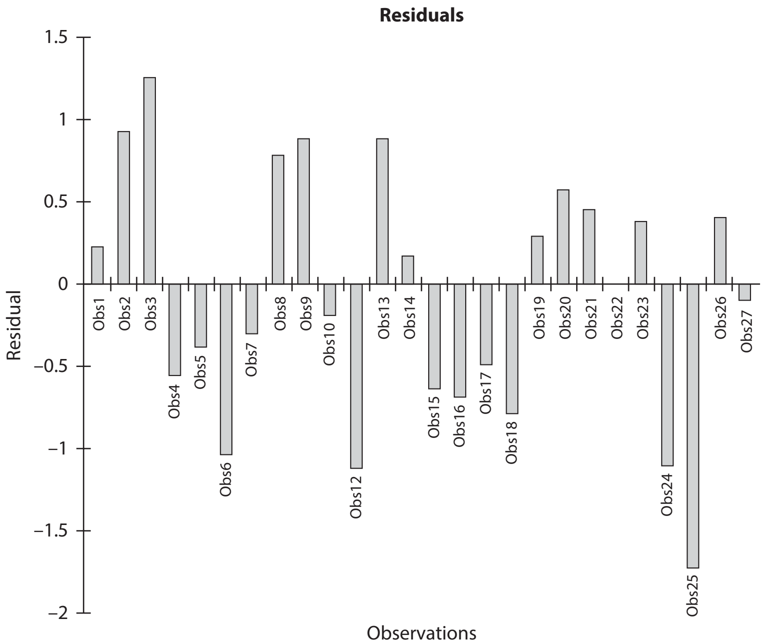


Figure 5.6. Fit assessment graph using residuals.

comfortable using the selected model to predict levels of Impairment with values of Burnout and Stressors.

Once you are comfortable that you have identified the appropriate model for the data, the next step is to add the predictor values for Burnout and Stressors into the nonparametric regression model. To add these values,

☐ Select => **Modeling data** => **Nonparametric regression** to open the Nonparametric regression window.

Verify that the information for each tab is correct in that it matches what you intended for the final model, and that options are set the same as when you found the best-fitting model.

☐ Select => **Prediction** tab and click to select the checkbox next to the **Prediction** option.

□ Click on the white space right below **Quantitative:** to move focus to the **Quantitative:** area.

By clicking in the white space, you draw focus to the area that allows the identification of the predictor values. The predictor values represent values for Burnout and Stressors variables to predict impairment. In the data set, the values are stored in two columns named Pred_Burnout and Pred_Stressors.

□ Find these columns on the spreadsheet. Next, left-click and hold while dragging the mouse from the first value for Pred_Burnout to the last value for Pred_Stressors.

Doing this will place text in the field to identify the data to use for the prediction (i.e., Data!$J$2:$K$3). Unlike the process used to select a column for the independent and dependent variables, the selection of the predictor values cannot include the column headings.

The use of the mouse to select the data for the predictors is only one method of telling XLSTAT what to use. If you want to avoid using the mouse, you can type the information directly into the field.

Important: The order from left to right of the independent variables in the spreadsheet must match the order of the predictor values. For example, the Burnout variable column in the spreadsheet is to the left of the Stressors column, which means that the values for Burnout and Stressors in a prediction model must be Burnout values on the left and Stressors values on the right.

Now that XLSTAT knows what information to use for the prediction,

□ Select => **OK** button at the bottom of the window.

The XLSTAT — Selections window appears, giving you one last chance to see what data you have selected for the prediction.

□ After reviewing your selections carefully, select => **Continue** button at the bottom of the window to conduct the analysis and open a new Excel sheet with the results.

XLSTAT Output

A new Excel sheet appears that contains all the LOWESS results requested and a scatter plot graph that shows the two prediction values along with the other data points. Figure 5.7 shows the two predicted Impairment values (5.038 and 4.466) for the two sets of values for Burnout (27.000 and 15.000) and Stressors (5.000 and 15.000). A review of the predicted values shown in the scatter plot reveals that two are on the fringe of other predicted values, but not so far out that they appear to be outliers. When predicted values are not within the proximity of other data values, this suggests the predictor values may be outside of the model's prediction capability. When this happens, you need to use caution in using these predictions for decision-making purposes.

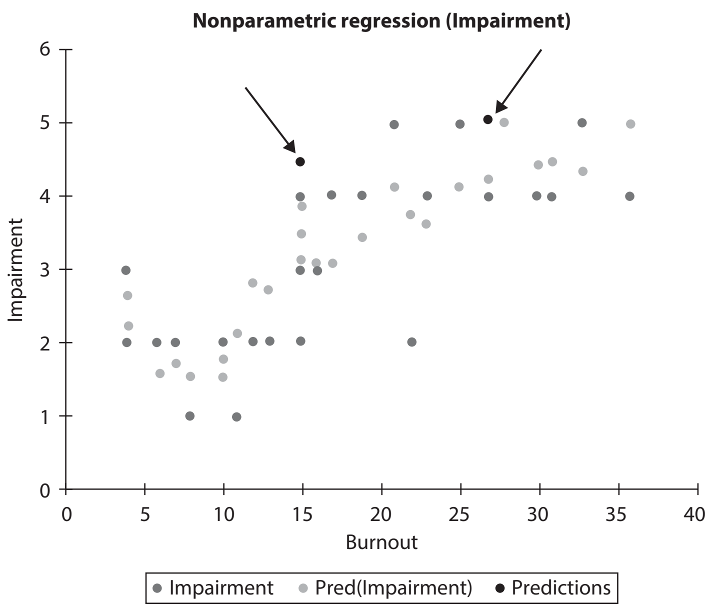


Figure 5.7. Two predicted values.

Findings

In statistical terms, a nonparametric regression model using the robust LOWESS estimation method achieved model fit with $R^2 = .618$ explaining 62% of the variance in Impairment using two independent variables, Burnout and Stressors. Two predictions for Impairment with values for Burnout (27 and 15) and Stressors (5 and 15) were 5.038 and 4.466 and were found to be within the model's range for effective prediction.

Practically, this means that the new Impairment scale is showing signs of reliability for use with two previously validated scales of Burnout and Stressors. This means that you can estimate level of Impairment with values for Burnout and Stressors.

Note: For comparison, we ignored the issue around meeting assumptions and conducted a parametric linear regression on the Research Question #19 data. The results established an R^2 of 0.604, indicating that the parametric model explained 60% of the variance in Impairment, in comparison to the 62% we found for the nonparametric model. In addition, the RMSE for the linear regression was 0.809, while the RMSE for the nonparametric regression as 0.747, meaning that there is a bit more variation in error values for the linear regression model than the error in the LOWESS model. The difference may be due to a variety of things, including how close the data are to violating assumptions. However, this finding does not mean that nonparametric models will always show improved model fit when the data violate the assumptions. Rather, it just means that you must consider many important factors when selecting a statistical approach so that you are not presenting results that are not supported statistically.

Box 5.3 Research Question #20

Dr. Northcott found evidence in the literature of a relationship between depressive symptoms and education level in social service professionals, which is important because rates of depressive symptoms are quite high among this professional group and can be a barrier to good practice. Soon to teach a master's level social work course, Dr. Northcott is interested in exploring the predictability of education level using scores on a standardized depressive symptoms measure. Therefore, her dependent variable is an ordinal measure of education and training (bachelor's degree only, master's degree with no license, and master's degree with clinical license) and depressive symptom scores are *her independent variable. Dr. Northcutt asks if you can find a regression model that could verify this relationship in her pilot sample. If she finds a relationship, she will use this information in her upcoming class.*

Research Question #20

Are depressive symptoms significantly related to an ordinal measure of education and training?

NONLINEAR REGRESSION

Nonlinear regression is a method that models a nonlinear relationship between a dependent variable and a single variable or multiple independent variables. Unlike the frequently used multiple regression, the relationship is not assumed to be linear — the relationship between the dependent and the independent variables is not assumed to represent a straight line. The dependent and independent variables can be either categorical/ordinal or continuous, and it is assumed that the responses captured by the dependent variable are independent and that the independent variables are not overly correlated (i.e., multicollinearity). As in nonparametric regression, the goal in conducting a nonlinear regression is to find a well-fitting model.

Other examples of uses for nonlinear regression are as follows:

1. A researcher conducting a feasibility study for a clinical trial wants to estimate the number of participants she can recruit related to the number of advertising spots she runs that are aimed at recruiting viable participants meeting the inclusion criteria for her trial.
2. An administrator is considering opening her agency one hour later during the weekend. She wants to predict the number of clients who will arrive in that 60-minute window based on data collected on number of arrivals throughout the day for 365 days.

Note: Another frequently used procedure is the multinomial regression procedure, which requires a categorical dependent variable with no implied order to the categories. If you have non-normal continuous variables, highly correlated continuous variables (i.e., $\rho > .8$), or implied order in your dependent categorical variable, then multinomial regression is not an option. Both multinomial regression and nonlinear regression are procedures available in SPSS.

So, what is the difference between nonparametric and nonlinear regression? Because there are many variations of both, answering this question is very complex. An oversimplification is that nonparametric regression builds a model based on some of the observed data, and nonlinear regression uses numerical optimization and a predetermined function (i.e., the formula that describes the probability of a relationship between the dependent and independent variables) to best fit parameter estimates to the data. Identifying the predetermined function for a specific data set can be challenging and may require more statistical background or assistance from a statistician, but if a function can be identified, your estimate errors more than likely will be smaller than when using a nonparametric approach. However, not all dependent to independent variable relationships will effectively fit, even a nonlinear relationship, which leaves the nonparametric regression procedure important to have as an option.

XLSTAT provides preprogrammed functions to help you begin identifying possible functions. However, to conduct a nonlinear regression in SPSS, you will need additional knowledge on function development that reaches beyond the scope of this book, therefore we use XLSTAT for Research Question #20 (see Box 5.3).

For Research Question #20, the dependent variable, Edu_Level, is ordinal (*1 = bachelor's degree, 2 = master's degree with no license*, and *3 = master's degree with clinical license*) and the independent variable, Depression (representing depressive symptoms) is continuous.

Before performing the nonlinear analysis, you must verify that the data support a nonlinear approach. The assumptions of constant error variance and independence are the same as they are for linear regression. However, you now replace the linear relationship assumption between the dependent and independent variable(s) with an assumption that an underlying probability model (i.e., function) exists for the observed data. Identifying the kind of underlying probability model is the most important task in developing a nonlinear regression model — a task demonstrated in Research Question #20.

Just as in Research Question #19, first you must carefully examine the variables individually to fully understand the quality and type of information they contain. Because this example uses a categorical variable, Edu_Level, and one continuous variable, Depression, you will use box-and-whiskers and Q-Q plots to examine the data. A review of the two sets of plots, along with other things, will help identify if the distributions among the three Edu_Level groups (i.e., *bachelor's degree, master's degree with no license*, and *master's degree with clinical license*) are possibly different from each other. If they are, you cannot use Edu_Level as your dependent variable in nonlinear regression. The nonlinear regression process requires that the distributions among the categories for the independent variables be similar.

You can use XLSTAT to review your data. If the XLSTAT menu bar is not visible,

☐ Select => **ADD-INS** tab at the top of the Excel window.
☐ Click on the **ToolBar Commands** button to open the XLSTAT toolbar.

After you click on the button, Excel will take a few minutes to bring up the toolbar and add a new tab named XLSTAT to the list of tabs at the top of the Excel window.

To request the plots,

☐ Select => **Visualizing Data** => **Univariate plots** next to the box-and-whiskers icon to open the Descriptive statistics window.

After the Descriptive statistics window opens,

☐ Select => **General** tab.
☐ Click to select **Quantitative data:** (if not selected) and then click in the white space representing the **Quantitative data:** area.

Now that the quantitative data field has focus,

☐ Click on the variable column for the Depression values (i.e., click on the letter "E" at the top of the column). Once you click on the "E," Data!EE appears in the field.
☐ Click to select **Subsamples:** (if not selected) and then click in the white space below **Subsamples:** to direct focus to that area.

Once the focus is on the **Subsamples:** area,

- ☐ Click on the column in Excel that contains the categorical values named Edu_Level (i.e., click on the letter "F").

Before leaving the **General** tab, verify that **Variable-Category** labels and **Sample labels** are both selected. Next,

- ☐ Select => **Options** tab.
- ☐ Make sure **Descriptive statistics**, **Charts**, **Compare to the total sample**, and **Sort the categories alphabetically** are selected.
- ☐ Select => **Chart(1)** tab.
- ☐ Make sure only **Box plots** and **Normal Q-Q plots** are selected.

The information and selections provided within each of these tabs tells XLSTAT how to produce the two plots.

Now that the options are set,

- ☐ Select => **OK** button at the bottom of the window to produce the plots.

Once **OK** is clicked, a final verification window appears (i.e., XLSTAT — Selections) giving one last look at the data selected for creating the plots.

Note: Only a few procedures in XLSTAT do not have ways to handle missing values. Unfortunately, **Visualizing Data** happens to be one of them. If you have missing data, you will need to sort and reselect your data to exclude cases with missingness prior to asking for **Box** or **Normal Q-Q** plots. After carefully reviewing the choices,

- ☐ Select => **Continue** button.

A new Excel sheet will appear with all the requested descriptive output.

EXCEL Output

The Q-Q plots help identify when distributions among the groups are too dissimilar to use in a nonlinear regression. Fortunately, the Q-Q plots in Figure 5.8 show only minor differences among the three groups, which is important because the distribution must be similar for the three groups if you are going to use nonlinear regression with a categorical dependent variable. In addition, the Q-Q plots show a specific

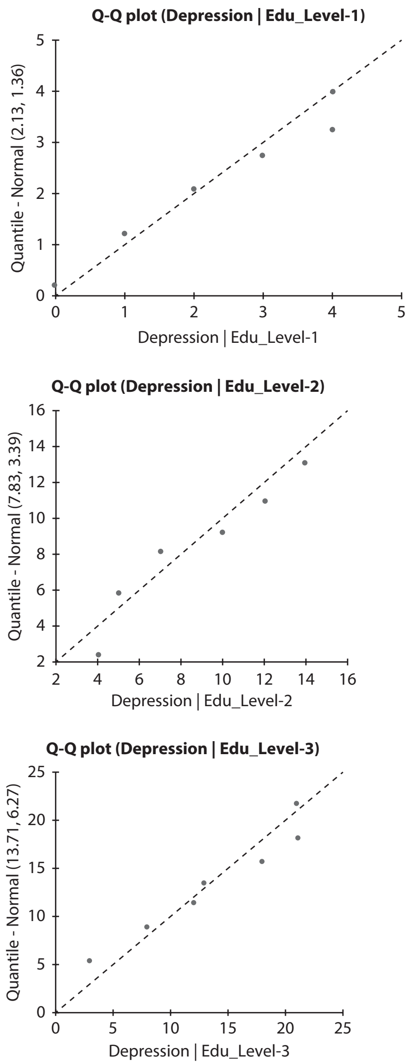


Figure 5.8. Q-Q plots.

pattern along the diagonal line (i.e., first above and then below), which does support the possibility of a nonlinear model.

The requested box-and-whiskers plots show the data distributions for each group, which are the categories for the dependent variable. Like the Q-Q plots, the differences among the box-and-whiskers plots are not enough to indicate that we cannot continue.

As stated for other statistical procedures, another important step before moving on is to make sure you have enough information for each group, which in this situation translates into enough representation in all three groups (i.e., n = 8, 12, and 7 for *bachelor's degree, master's degree with no license*, and *master's degree with clinical license,* respectively). The minimum guideline we recommend is to have at least 5 responses for each category in a statistical procedure, but remember that you should use any parameter estimates with care due to the lack of precision from using a very small sample.

Note: The Q-Q plots presented in Figure 5.8 are those that should be produced prior to conducting a linear regression to verify that the normality assumption is not violated.

Now that you have examined the data closely and have verified similarity among the categories, you can move on to finding a model that fits your data. So how many different models are possible? Unfortunately, the list is extremely long. One of the strengths of nonlinear regression is that a researcher can explore many different models with many different levels of complexity. Up to now, we have avoided discussing formulas, and we will continue to leave formula development to other discussions outside of this book. However, to accomplish the nonlinear regression, you will need to use three different formulas we provide for this procedure. The three formulas represent three possible probability function models that fit nonlinear relationships. Many different probability functions are built into XLSTAT for conducting nonlinear regression. In addition to the built-in functions, XLSTAT offers a process that allows for the development of other function models.

While answering Research Question #20, both the built-in and the user-defined methods are described. You may ask, "With all this flexibility, how do I know where to start?" Fortunately, the three function options cover the use of two-, three-, and four-parameter models (i.e., probability models predicting a value for the dependent variable) that will meet the needs of most people who are not statisticians but have the occasion to use nonlinear regression. These three examples will get you well on your way to conducting a nonlinear regression analysis. We present the functions related to the examples later when discussing the steps for the XLSTAT procedure.

Exploring different models is the main process for identifying an effective nonlinear relationship (unlike nonparametric regression, in which the data identify the relationship). The first model relates to one that uses only two parameters to identify the relationship between a dependent and an independent variable. If you took a basic statistics class, a simple linear regression formula included an intercept and a slope (i.e., β_0 and β_1, respectively). The intercept and slope are two parameters that help explain the relationship between a dependent and independent variable. The two-parameter nonlinear model attempts to do the same thing, which is to find estimates for the two parameters that appropriately explain the relationship. You can add parameters when the two parameters are not enough to explain a more complex situation. Each model has a unique way of using the parameters in a formula and explaining the relationship. As you would expect, using more parameters than are necessary adds complexity and possibly unwanted estimate error to your prediction. Therefore, the goal is to find an effective model by finding the formula that has the fewest number of parameters while still effectively representing the relationship between the variables.

A nice feature of XLSTAT is that two of the three functions that include two and three parameters are already in a list of available functions; however, SPSS offers no list of functions to get you started. In addition, XLSTAT provides a very friendly way to include user-defined functions — which is demonstrated as part of answering Research Question #20.

ANALYZING THE DATA IN XLSTAT AND EXCEL — TWO PARAMETERS

Prior to running the analysis, you must arrange the data to meet the requirements of the software you are using. XLSTAT expects that you arrange the variables in columns, and that each row represents a unique case or respondent. Check that you arrange the values for Depression and Edu_Level in two adjacent columns, respectively. If the XLSTAT menu bar is not available,

- ☐ Select => **ADD-INS** tab at the top of the Excel window.
- ☐ Click on the **ToolBar Commands** button to open the XLSTAT toolbar.

After you click on the button, Excel will take a few minutes to bring up the toolbar and add a new tab named XLSTAT to the list of tabs at the top of the Excel window. You will find the nonlinear regression analysis to run under the **Modeling Data** button.

- ☐ Select => **Modeling Data** => **Nonlinear regression**, which opens the Nonlinear regression window.

At the top of the Nonlinear regression window are seven tabs. The information and selections provided within each of these tabs tell XLSTAT how to conduct the nonlinear regression. For Research Question #20,

- ☐ Select => **General** tab.
- ☐ Click on the white space below **Quantitative:** to move focus to the **Y/Dependent variables:** area.

By clicking in the white space, you shift your focus to the area that allows for the identification of the dependent variable. The dependent variable in the spreadsheet for Edu_Level is in column F, so find that column on the spreadsheet and

- ☐ Click on the "F" at the top of the column.

Doing this will both highlight the column in the traditional way Excel selects a column and add text to the **Y/Dependent variables:** area (i.e., Data!$F:$F). Data!$F:$F indicates that you want to use all the data in that column for the dependent variable.

- ☐ Make sure you check the **Variable labels** option on the right side of the window.

- ☐ Click in the white space below **Quantitative:** for the **X/ Explanatory variables:** area, drawing focus to that area.

You can find the independent or explanatory variable in the example in column E.

- ☐ Find the E column on the spreadsheet and click on the "E" at the top of the column.

Doing this will place text in the field to identify the column (i.e., Data!$E:$E).

Note: The use of the mouse to select the data for the analysis is only one method for telling XLSTAT what to use. If you want to avoid using the mouse, you can type the information directly into the fields.

The next step is to select the probability model to use.

- ☐ Click on the **Functions** tab to display the options for function selection (see Figure 5.9 — Nonlinear regression window).

Remember that **Function** is the formula that describes the probability of a relationship between the dependent and independent variables.

When the information for the **Functions** tab appears, you will notice two sections. The section on the left (**Built-in functions**) lists the

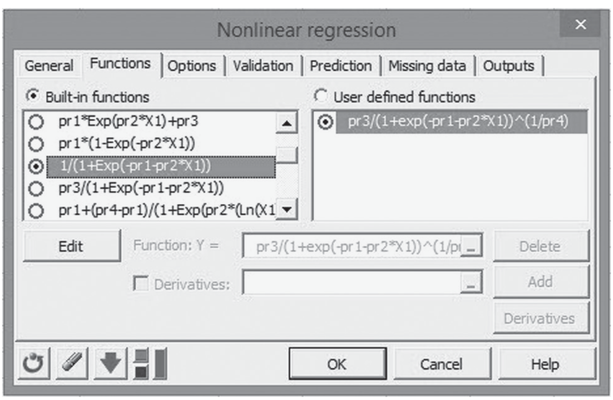


Figure 5.9. Nonlinear regression window.

probability models already defined in XLSTAT. The function highlighted on the left in Figure 5.9 is the first one you should examine to review for model fit.

- ☐ Click to select **Built-in functions**.
- ☐ Scroll down the **Built-in functions** list and click to select the one that is defined as 1/(1 + Exp(−pr1−pr2*X1)).

This function represents a nonlinear regression model that uses only two parameters (i.e., pr1 and pr2). Once the function has been selected, you have chosen to examine the relationship between the variables using a model with only two parameters.

Note: XLSTAT offers many functions that includes the one used first for this example. If you are not comfortable with formulas, then this example may be a bit unsettling. We recommend that, as you experience this nonlinear procedure for the first time, you don't spend time trying to interpret the formulas right away. Simply use this opportunity to work through the entire process for nonlinear regression and see how to perform a nonlinear regression analysis. Later, you can request assistance from a statistician for a deeper understanding.

- ☐ Select => **Options** tab and verify that none of options on the left side is selected.
- ☐ Select => **Validation** and **Prediction** tabs and verify that none of the options under these two tabs is selected.
- ☐ Select => **Missing data** tab and verify that **Remove the observations** option is selected.
- ☐ Select => **Outputs** tab and verify that all the options are selected.

Now that the options are set,

- ☐ Select => **OK** button at the bottom of the window to execute the analysis.

Once **OK** has been clicked, a final verification window appears (i.e., XLSTAT — Selections) giving one last view of the data selected for the analysis.

- ☐ After carefully reviewing the choices, select => **Continue** button and a new Excel sheet will appear with all the requested analysis output.

Excel Output — Two Parameters

At the completion of the analysis, a new Excel tab will appear and will contain the goodness-of-fit statistics as part of the analysis output.

For the function representing two parameters, the R^2 (amount of variation explained) is very small (0.065). A nonlinear regression model that explains only 6% is not a useful model for a study. Therefore, you need to select a different probability model or function in XLSTAT to determine if one of the other models provides an adequate fit. The following information in this section of the chapter refers to three- and four-parameter models, which are models that can add different curves or shapes, along with other planes, to the slope and intercept parameters. Don't worry. We provide the formulas for them.

ANALYZING THE DATA IN XLSTAT AND EXCEL — THREE PARAMETERS

To investigate a different model, you will need to reopen the Nonlinear regression window.

- ☐ Select => **Modeling data** => **Nonlinear regression** to reopen the Nonlinear regression window.
- ☐ Select => **Functions** tab.

Fortunately, XLSTAT also includes the second model. The second function is in the list just below the first one you tried, and is the function pr3/(1 + Exp(-pr1-pr2*X1)). This function attempts to model the data using three parameters (p1, p2, and p3).

- ☐ Click to select the function pr3/(1 + Exp(-pr1-pr2*X1)) from the list.
- ☐ Select => **OK** button at the bottom of the window to execute the analysis.

Once **OK** has been clicked, a final verification window appears (i.e., XLSTAT — Selections) giving one last view of the data selected for the analysis.

- ☐ After carefully reviewing the choices, select => **Continue** button and a new Excel sheet will appear with all the requested analysis output.

Note: Excel will not overwrite the previous output appearing in separate Excel sheets. Each run of an analysis will create a new Excel sheet. Keeping track of the sheets by renaming them will save you time later when you are writing up your findings.

Excel Output — Three Parameters

With the newly selected function that involves three parameters, the R^2 improves to 0.571 (i.e., 57% of the variance in Depression is explained by Edu_Level). In social science settings, 57% is sufficient variance explained to feel positive about the model in a study. However, without further investigation of other possible models, you will not know if a three-parameter model is the best option to explain the variable relationship.

The three graphs in Figure 5.10 show some of the analysis output from nonlinear regression. The new Excel sheet containing the analysis output includes these graphs. You can see that

- The relationship between Depression and Edu_Level is nonlinear — the line showing the relationship is curved.
- In the graph of Edu_Level predictions vs. Edu_Level, the model does better at predicting *bachelor's degree* (category 1) and *master's degree with no license* (category 2) than at predicting the *master's with clinical license* (category 3).
- The residuals graph does not show signs of a pattern in the errors, though one case does have a larger error than the other cases.

It is fair to say that the results from the nonlinear regression model containing three parameters show that you have sufficient evidence to conclude that the levels of depressive symptoms are different for people with different education and training levels. You can draw this conclusion primarily because you could find a model that showed appropriate fit and prediction capabilities. If you were not able to find an effective model, you could not draw this conclusion, and you would need to test the probability model with four parameters. To be thorough in presenting the process and to complete the introduction to nonlinear regression, we next describe the four-parameter model option.

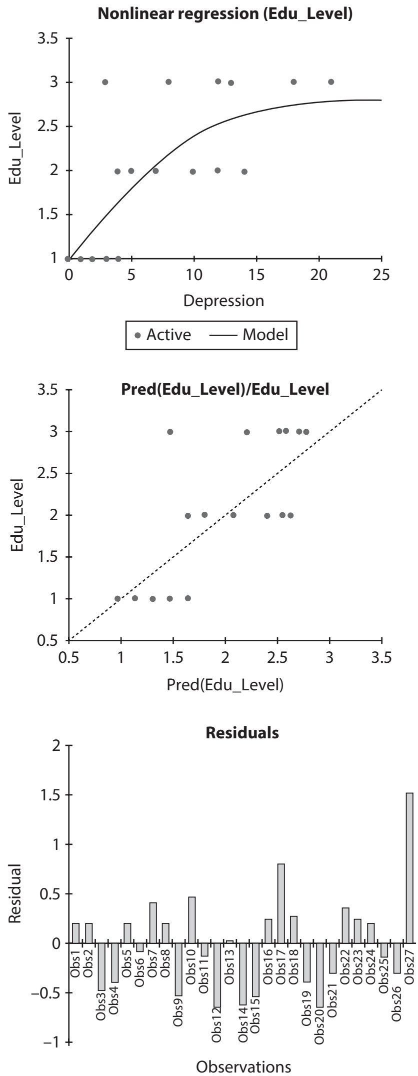


Figure 5.10. Analysis output graphs from nonlinear regression.

ANALYZING THE DATA IN XLSTAT AND EXCEL — FOUR PARAMETERS

To conduct a nonlinear regression with a model with four parameters, you must add a few additional bits of information into Excel, because the probability model with four parameters is not part of the predefined functions within XLSTAT. The additional bits of information are four formulas that represent the derivatives of the four-parameter function. If you do not know what a derivative is, don't worry — it involves calculus and is beyond the scope of this book. You can simply follow the steps, which are very easy to perform, to get XLSTAT ready to use this new model, and learn more about derivatives after you have mastered the steps. One nice feature is that once you add the four-parameter model to XLSTAT, you will not have to do it again when you conduct a nonlinear regression analysis in the future.

By adding the four formulas below, you are providing XLSTAT with the next model option in the set of options for this example. These four model options (i.e., functions with one, two, three, and now four parameters) are part of the exponential family of functions for modeling nonlinear relationships between a dependent variable and independent variable(s). In the future, if you have a prediction model that involves a nonlinear relationship, these four formulas are a great place to start. Simply follow the steps presented in this example for answering Research Question #20 with your own data.

Before returning to the nonlinear regression window in XLSTAT, you must create a new Excel sheet by naming it *derivatives* and entering the following four formulas into column A. However, if you have downloaded the data set for Chapter 5 from the companion website, you will find the sheet named *derivatives* ready for you to use. The four formulas are:

$$(pr3 / pr4) * \exp(-pr1 - pr2 * X1) / (1 + \exp(-pr1 - pr2 * X1))^{(1+1/pr4)}$$
$$(pr3 * X1 / pr4) * \exp(-pr1 - pr2 * X1) / (1 + \exp(-pr1 - pr2 * X1))^{(1+1/pr4)}$$
$$1 / (1 + \exp(-pr1 - pr2 * X1))^{(1/pr4)}$$
$$(pr1 / pr4^2) * \ln(1 + \exp(-pr1 - pr2 * X1)) / (1 + \exp(-pr1 - pr2 * X1))^{(1/pr4)}$$

Said differently, if you have not yet downloaded the Chapter 5 data set, open an Excel sheet that contains no information, type the four formulas exactly as written above into the first four lines in column A in the empty sheet and label it *derivatives* in the label space at the bottom of the sheet. These four formulas are the same formulas you can use with other data when examining a model with four parameters. Again, you do not need to understand the workings of these formulas for now, just know that XLSTAT needs these four to support your effort to examine a four-parameter model that XLSTAT did not predefine. Once you enter the four, and after verifying that you typed them correctly, return to XLSTAT and set up the new probability model.

To continue your analysis in XLSTAT, you'll need to reopen the Nonlinear regression window.

- ☐ Select => **Modeling data** => **Nonlinear regression** to open the Nonlinear regression window.
- ☐ Select => **Functions** tab.
- ☐ This time, click to select the **User defined functions** option on the right side of the Nonlinear regression window.

After the **User defined functions** option is selected,

- ☐ Click on the **Add** button toward the bottom right of the window.

When you click the **Add** button, the two white areas for **Function: Y** = and **Derivatives:** become active.

- ☐ Be sure the **Derivatives:** option is selected.
- ☐ Click to move focus to the **Derivatives:** white space area.

Next, we want to let XLSTAT know where to find the four formulas entered earlier into that blank Excel sheet that you have named *derivatives*.

- ☐ Navigate to where the four formulas are, and click and drag to highlight all four of them (i.e., click and drag adds the location information into the area — derivatives!$A$1:$A$4).
- ☐ Click to move the focus to the white space for the **Function: Y** = area.
- ☐ Type the formula pr3/(1 + exp(-pr1-pr2*X1))^(1/pr4)

Note: The formula starts with "p" and ends with ")" — a closed parenthesis symbol. Also, once you add this as a user-defined function, you will not have to add it again for future analyses unless you update XLSTAT with a new version.

- ☐ After you verify the formula is correct, click the **Save** button to add the formula to the **User defined functions** list.

The new four-parameter function should be highlighted, telling XLSTAT to use it for the analysis. Therefore, adding the four formulas into the blank Excel sheet and adding the new function into the **Function: Y =** area has enabled XLSTAT to use this new probability model with four parameters in an analysis.
After carefully reviewing your work,

- ☐ Select => **OK** button at the bottom of the window to execute the analysis.

Once **OK** is clicked, a final verification window appears (i.e., XLSTAT — Selections), giving you one last view of the data selected for the analysis.

- ☐ Select => **Continue** button and a new Excel sheet will appear with all the requested analysis output.

Excel Output — Four Parameters

A review of the output from the four-parameter model shows an insignificant increase in R^2 from the three-parameter model (from 0.571 to 0.575). Recall that when a less complicated model fits relatively the same as a more complicated model, the less complicated model is the better selection, so you should clearly use the results from the three-parameter model. A review of the three-parameter model fit statistics shows that the RMSE is relatively small, and a review of the estimates of the three parameters seems reasonable. Therefore, our final nonlinear model is within the Excel sheet under **Equation of the model:**. After rounding of a few of the parameter estimates, the final model is:

$$\text{Edu\_Level} = 2.80 / \left(1 + \text{Exp}\left(+0.61 - 0.24 * \text{Depression}\right)\right)$$

The above formula can be found in the Excel sheet that was generated from the analysis. The formula is located on, or close to, line 41 on the Excel sheet. Because the model fit is adequate, you can enter a value for Depression into the above equation to predict what category is most likely — the education level category in which the respondent will most likely be. One way to accomplish this is to use Excel.

- ☐ Open a new Excel sheet.
- ☐ Enter a value for Depression (e.g., 10) into cell A1.
- ☐ Enter the following into cell B1:

$$= 2.80 / \left(1 + \exp\left(.61 - .24 * \text{A1}\right)\right)$$

- ☐ After typing in the formula above, press the Enter key on your keyboard to calculate a value for Edu_Level that will appear in cell B1. The value that appears in cell B1 is most likely a number that you can round to the nearest integer. The integer number is the predicted Edu_Level category for the value of Depression you entered in A1.

Findings

A three-parameter nonlinear model with an ordinal dependent variable Edu_Level and a continuous independent variable Depression was identified for estimating education and training level (Edu_Level) using scores on a depressive symptoms measure (R^2 = .571). The three-parameter fit indices for RMSE and R^2 were much improved over a two-parameter model. In addition, R^2 for a four-parameter model was the same as the three-parameter model, and the RMSE showed only a slight increase, results that support the three-parameter model as most efficient.

Practically, this means that the nonlinear model has identified a relationship between Depression and Edu_ Level (level of education and training). These findings from this sample suggest that levels of depressive symptoms do predict variations in education and training, but more data collection and analyses are necessary to conclude a causal relationship. A different and perhaps more intuitive regression model could reverse the roles of the variables (i.e., education and training as a predictor of depressive symptoms level). Knowing which to choose as the dependent variable centers on the research question you are attempting to answer and the type of data available. This alternative regression model would explore the possibility that education and training lower depressive symptoms levels.

Note: Another SPSS analysis method for models with ordered categorical dependent variables is PLUM (*Pol*ytomous *U*niversal *M*odel). For more information on this nonparametric strategy, see the website https://statistics.laerd.com/spss-tutorials/ordinal-regression-using-spss-statistics-2.php

Box 5.4 Research Question #21

Dr. Northcott is interested in examining professional life satisfaction, as the literature suggests that this is related to turnover, and turnover is a serious problem in adoption and foster care agencies. She suspects that those working in the public sector may experience lower professional life satisfaction than those working in the private sector. Based on studies of other professions, she also suspects that men have higher professional life satisfaction than women. To investigate these relationships in her pilot study, Dr. Northcott wants you to examine the possible relationships between life satisfaction and two factors — working in the private or public sector and if the person is male or female. Note that we are assuming a dichotomous variable for gender for this example only, but that gender, along with biological sex, is more complicated than a dichotomy.

Research Question #21

Are employment setting and gender significant for understanding levels of professional life satisfaction?

FACTORIAL ANALYSIS

Factorial analysis, also called general linear model (GLM) analysis, is a method for conducting a regression analysis in which the dependent variable is continuous (in this example) and the independent variables are categorical (i.e., the factors in factorial analysis). This analysis examines the effects factors have on the dependent variable. In other words, a factorial analysis explores the potential effects on a dependent variable due to being in one group rather than another. The test similar to a factorial analysis is the traditional regression with a continuous dependent variable and with multiple dummy variables as independent variables. The dummy variables represent whether the respondent is or is not in each group for each variable. In some situations, adding the necessary number of dummy variables is problematic and meeting the assumptions of parametric regression is sometimes challenging, especially when dealing with small samples.

Unfortunately, no nonparametric approach is easily available to conduct a factorial analysis that matches Research Question #21 (see Box 5.4). Yes, a few methods that will work have been published, but they require programming/coding experience in SAS or R. Those who do not have that kind of experience need an alternative. Most importantly, the alternative must not be to run an analysis just because you found software that will run an analysis. Without the statistical resources and experience, Research Question #21 becomes much more challenging because you need to lean on the available parametric approach to proceed. Although a nonparametric approach is not easily available, we present a suggested SPSS parametric procedure because of its many applications to small data sets, and knowing what to look for when building factorial models is a very important skill to have when analyzing different kinds of data.

Note: SPSS will be happy to provide you with analysis results, even if the results are fundamentally flawed from running an analysis on data that do not meet the required assumptions, misinterpreting a variable, having the wrong level of measurement, experiencing unrecognized bias, etc. It is the responsibility of every researcher to examine fully all the possible threats and biases to their findings before presenting their results to others.

Factorial models have many assumptions and requirements, most of which are discussed earlier in this book. The assumptions are normality, homoscedasticity, independence, and a continuous variable for the dependent variable. In addition, you need to make sure that you have enough coverage in your categorical variables. Coverage of the variable categories is related to having enough representation in the options to support the use of the variable in the analysis. For example, if a variable for gender has data for 20 females and three males, according to an accepted guideline of five per category, you do not have enough males representing the male category — the variable does not have enough information to compare males and females.

The five-count guideline becomes more difficult to meet as you add more categorical variables into a factorial model. For example, if Research Question #21 aimed at knowing if gender and education level were predictors of depressive symptoms, the data do not have enough coverage to pursue this analysis (see Table 5.2). Remember that Edu_Level is an ordinal measure of education and training (*1 = bachelor's degree only, 2 = master's degree with no license,* and *3 = master's degree with clinical license*).

Review of Table 5.2 reveals that cross-tabulation of males with education level does not meet the five-count guideline. In other words, you do not have enough males in Education Levels 1 and 2 to represent properly all the potential males who are at those education levels, and the same is true for both males and females in the Education Level 3 category.

Sometimes the coverage for the various categories is such that collapsing a category is necessary. For example, one option for Research Question #21 would be to merge education categories 1 and 2 to compensate for the low number of men, but the few numbers and therefore the lack of information for category 3 would still be unresolved, and just managed for analysis purposes. It is important to note that SPSS will perform a factorial analysis on a model with Gender and Edu_Level unchanged, but the potential of coming to a flawed conclusion based on these variables is high. SPSS does not always warn you when cell coverage is too low. Therefore, with small data sets, at times you may not

Table 5.2. Frequency for Gender and Edu_Level — Edu_Level * Gender Crosstabulation

Count		*Gender*		*Total*
		Female	*Male*	
Edu_Level	1	6	2	8
	2	6	4	10
	3	3	4	7
Total		15	10	25

be able to use specific combinations of variables, depending upon the coverage in each cell.

SPSS Process (Using SPSS to Build a Factorial Model)

We now return to using SPSS, because SPSS offers the process needed to analyze Research Question #21. To begin your analysis,

- ☐ Select => **Analyze** => **General Linear Model** => **Univariate**.

You select the **Univariate** option because you have only one dependent variable, named LifeSatisfaction. Once the Univariate window opens, you need to tell SPSS which variable to analyze. The list on the left includes the variables in the data set.

- ☐ Click to select LifeSatisfaction from the list of variables on the left.
- ☐ Click the move arrow closest to **Dependent Variable:** to move the variable to the **Dependent Variable:** area.
- ☐ While holding down the Ctrl key, click to select both the Gender and Employment variables so they are highlighted.
- ☐ Click the move arrow closest to **Fixed Factor(s):** to move the variables to the **Fixed Factor(s):** area.

Why use **Fixed Factor(s):** instead of random (i.e., **Random Factor(s)**)? Variables considered fixed factors are those that contain information on all the possible categories represented by the variable. (For more information on fixed vs. random factors, see Raudenbush, 1993.)

Before you select the **OK** button at the bottom of the window, you must select a few analysis options.

- ☐ Select => **Model**. . . button on the left to open the Univariate: Model window.

Verify that the **Full factorial** option is selected because this will test potential differences for Gender, Employment, and their interaction (Gender * Employment) for LifeSatisfaction.

- ☐ Select => **Continue** to close the window.

Next, we want to request plots to help verify the results, so,

- ☐ Select => **Plots**. . . button on the right to open the Univariate: Profile Plots window.

After the window opens, you will see the two factors Gender and Employment in the **Factors:** area. With only two factor variables, we suggest that you request two plots — one with Gender on separate lines and one with Employment on separate lines.

- ☐ Click to select Gender in the **Factors:** area.
- ☐ Click the move arrow closest to **Separate Lines:** to copy Gender to the **Separate Lines:** area.
- ☐ Click to select Employment in the **Factors:** area
- ☐ Click the move arrow closest to **Horizontal Axis:** to copy Employment to the **Horizontal Axis:** area.
- ☐ Request the plot by clicking on the **Add** button, which will add Employment*Gender to the **Plots:** area. Follow these same steps to add Gender*Employment to the **Plots:** area to see the same information, just a different view.
- ☐ After carefully reviewing your choices, select => **Continue** button to close the window.

Because the factors Gender and Employment are dichotomous, running a post hoc test to examine differences between variable categories is not necessary. However, we do need to select some options to verify a few assumptions about the data.

- ☐ Select => **Options**. . . button on the right to open the Univariate: Options window.
- ☐ Click on the (OVERALL) entry in the **Factor(s) and Factor interactions:** area.
- ☐ Click the move arrow to copy (OVERALL) to the **Display Means for:** area.
- ☐ Select four options in the **Display** area (**Descriptive statistics**, **Parameter estimates**, **Homogeneity tests**, and **Residual plot**).

Note: You can click the **Help** button on this window to get additional information on all the available options.

- ☐ After carefully reviewing your choices, select => **Continue** button to close the window.
- ☐ Select => **OK** to conduct the analysis.

See Appendix A for complete syntax.

SPSS Output

After you select **OK,** a variety of information shows up on the SPSS output window. On the left side of the output window, you will see an index showing the names of the information sections created from running the analysis. The names are arranged in an index, and clicking on the name displays the details of that section on the right. On the right side of the output window, you will see all the detailed information within each section.

If you have been running other analyses before your factorial model, you may have other information in your output window. If this is the case, simply click on **Univariate Analysis of Variance** in the index on the left side of the output window. This will cause the left side of the window to scroll to the detailed information for the analysis output.

Findings

In the **Descriptive Statistics** table in the IBM SPSS Statistics Viewer window, you can verify that you selected the correct variables for the analysis and that the data provide sufficient coverage for the combination of categories (i.e., values for n meet the guideline of five or more per cell).

The next table, **Levene's Test of Equality of Error Variance**, suggests (p = .213) that the statistical assumption about homogeneity of variances among the groups has not been violated. A **Sig.** value smaller than α =.05 would indicate the variances are potentially different among the factors and therefore may require a different analysis approach, because an assumption has been violated.

When you scroll down to the next section of output information, the **Test of Between-Subjects Effects** table shows the results of your general linear model (GLM) analysis — Factorial Models are part of the GLM family. The values in the **Sig.** column show that LifeSatisfaction between private and public employment is not significantly different, and neither is the interaction of male and female with employment type (Employment, p = .951, and Gender*Employment, p = .167). This indicates that there is not enough information to assume that life satisfaction levels for males with public employment are different from life satisfaction levels for females with private employment. However, the **Sig.** value for Gender (p = .043) falls below the commonly used threshold of α =.05. This indicates that information from the data suggests that life satisfaction levels, in general, are different for males and females.

Another important value to review in the **Test of Between-Subjects Effects** table is the **Adjusted R Squared** at the bottom of the table. R^2 is the same as **R Squared** on SPSS's note at the bottom of the table, and recall that R^2 represents the amount of variability in the dependent variable explained by the independent variables. In this analysis, Gender and Employment explain about 15% of the variability in LifeSatisfaction. Many authors elect to report the R^2 value, but the Adjusted R^2 calculation is more informative about the performance of the model because it takes into consideration the number of independent variables used. As more variables are added to a model, the R^2 value naturally increases, and the Adjusted R^2 corrects for this general increase and takes into account the sample size.

The remaining three graphs in the analysis output provide an opportunity to verify that another assumption has not been violated and to see a graphical representation of the **Sig.** values found in **Test of Between-Subjects Effects** table. The first graph, with **Dependent Variable: LifeSatisfaction** at the top, illustrates the observed vs. predicted values for the model, which reveals the amount of difference between the two. A review of this graph should include a search for possible patterns that may indicate a deviation from equal variance for LifeSatisfaction among the groups. For Research Question #21, the graph of predicted vs. observed does not show any patterns, which is what you would expect due to the findings of the **Levene's Test of Equality of Error Variance** discussed earlier.

The last two plots in the **Profile Plots** section show estimated means for the factors, one by Employment and the other by Gender. Recall, we asked for two plots that contained the same information under the

Plots button prior to running the analysis. The differences in how the lines are drawn using the **Estimated Marginal Means** values is larger for Gender than it is for the Employment Settings. Using the vertical case for **Estimated Marginal Means** will help in comparing the two graphs. The observed differences in the two graphs support the *p*-values for Gender ($p = .043$) and Employment Setting ($p = .951$).

Another important observation is the intersection of the two lines for Gender (see Figure 5.11). Intersecting lines can indicate a potential significant factor interaction. However, in our analysis, the *p*-value for Gender*Employment was not significant ($p = .167$). Because the sample is very small, there may not be enough information to identify a significant interaction of Gender and Employment, especially with Employment categories involved ($p = .951$).

A good way to test the effects of the nonsignificant interaction on the model is to remove the interaction term altogether and rerun the analysis. To accomplish this,

☐ Select => **Analyze** => **General Linear Model** => **Univariate** to reopen the Univariate window.

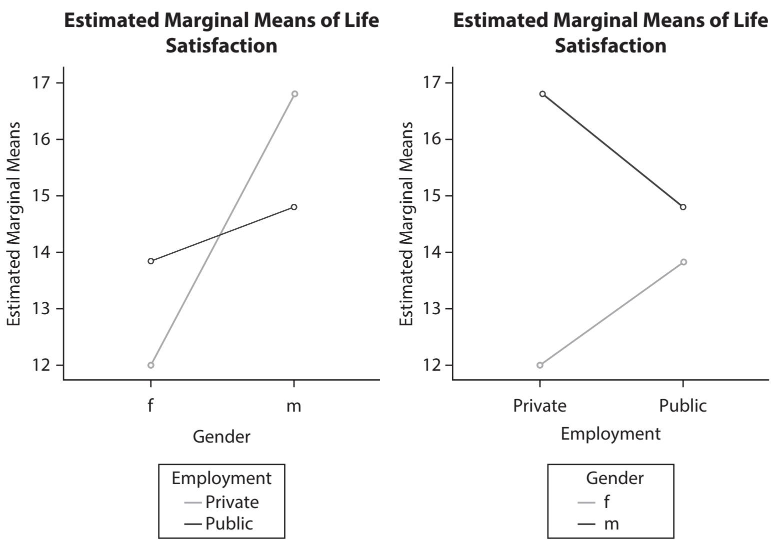


Figure 5.11. **Estimated Marginal Means** of Life_Satisfaction.

- ☐ Select => **Model**. . . button to open the Univariate: Model window.
- ☐ Click to select the **Custom** option to gain access to the **Model:** area.
- ☐ Click to select Gender in the **Factors & Covariates:** area.
- ☐ Click the move arrow to move **Horizontal Axis:** to the **Model:** area.

In other words, because the interaction term Gender*Employment and the factor Employment were found not to be significant, we are going to rerun the analysis with only the factor Gender.

- ☐ After carefully reviewing your choices, select => **Continue** button to close the window.
- ☐ Select => **OK** button to rerun the analysis. If you get a warning message that one of your factors (Employment) is not being used, simply click the **OK** button to continue.

With the updated model, repeat your review of the Levene's test and check the assumptions as previously described. For this analysis, after repeating the process, the **Tests of Between-Subjects Effects** table shows that Gender is significant (p = .031) and that the Adjusted R Squared of .152 is just .003 smaller than the previous model. Therefore, this simpler model, with just a trivial drop in Adjusted R Squared explaining 15% of LifeSatisfaction variance, suggests a more efficient model for analyzing Research Question #21. The **Parameter Estimates** table shows a significant parameter estimate for Gender (β = -3.067, p = .031), providing evidence that mean differences for LifeSatisfaction exist between males and females (see Table 5.3).

Findings

In statistical terms, a GLM model examining LifeSatisfaction for factors Gender and Employment found that Employment and the interaction of Employment and Gender were not significant. A model with only Gender was significant (p = .031) at α = .05 with an adjusted R^2 = .152.

Table 5.3. Parameter Estimates

Parameter	*B*	*Std. Error*	*t*	*Sig.*	*95% Confidence Interval*	
					Lower Bound	*Upper Bound*
Intercept	15.800	1.031	15.323	.000	13.667	17.933
[Gender = f]	–3.067	1.331	–2.304	.031	–5.820	–.313
[Gender = m]	0[a]	.	.	.	.	.

[a] This parameter is set to zero because it is redundant.

[b] Dependent variable: LifeSatisfaction.

Practically, this means that the GLM analysis shows that you do not have enough evidence to suggest a difference in life satisfaction between individuals who have public or private industry employment. However, there is sufficient evidence to suggest that, in general, males and females have different life satisfaction levels, and that females, on average, score three points lower than males when measured by a validated professional life satisfaction scale. In addition, gender accounts for 15% of the variance in life satisfaction in the way life satisfaction was measured for this study.

Appendix A
SPSS Syntax

Appendix A presents the SPSS syntax for each of the procedures presented in this book, identified by chapter and by research question. Working with the syntax is more comfortable for some researchers, as they find it easier to follow the specific characteristics of each analytic procedure. In addition, the capability and usefulness of SPSS can greatly increase when a researcher moves beyond simply using menus and point-and-click procedures.

CHAPTER 2

Research Question #1

```
*Frequency Analysis:
FREQUENCIES VARIABLES=Confidante
/ORDER=ANALYSIS.

*Parametric T-Test
T-TEST
/TESTVAL=0
/MISSING=ANALYSIS
```

```
/VARIABLES=Confidante
/CRITERIA=CI(.95).

*Binomial Analysis:
*Nonparametric Tests: One Sample.
NPTESTS
/ONESAMPLE TEST (Confidante) BINOMIAL(TESTVALUE=0.6
  CLOPPERPEARSON JEFFREYS LIKELIHOOD
  SUCCESSCATEGORICAL=LIST(1) SUCCESSCONTINUOUS=CUTPOINT(MIDP
  OINT))
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Research Question #2

```
*Chi-square test:
*Nonparametric Tests: One Sample.
NPTESTS
/ONESAMPLE TEST (Coming_Out) CHISQUARE(EXPECTED=EQUAL)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Research Question #3

```
*Kolmogorov-Smirnov test:
*Nonparametric Tests: One Sample.
NPTESTS
/ONESAMPLE TEST (Loneliness)
KOLMOGOROV_SMIRNOV(NORMAL=SAMPLE UNIFORM=SAMPLE
EXPONENTIAL=SAMPLE POISSON=SAMPLE)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Research Question #4

```
*Wilcoxon signed-rank test:
*Nonparametric Tests: One Sample.
NPTESTS
/ONESAMPLE TEST (Coping_Level) WILCOXON(TESTVALUE=15)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Research Question #5

```
*Descriptive Analysis:
DESCRIPTIVES VARIABLES=Rater_score
/STATISTICS=MEAN STDDEV MIN MAX.
*Runs test with mean cutpoint:
*Nonparametric Tests: One Sample.
NPTESTS
/ONESAMPLE TEST (Rater_score) RUNS(GROUPCATEGORICAL=SAMPLE
GROUPCONTINUOUS=CUTPOINT(SAMPLEMEAN))
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
*Runs test with dichotomous variable:
*Nonparametric Tests: One Sample.
NPTESTS
/ONESAMPLE TEST (Rater_score Confidante)
RUNS(GROUPCATEGORICAL=SAMPLE
GROUPCONTINUOUS=CUTPOINT(SAMPLEMEAN))
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Research Question #6

```
*Histograms for Age and Self_esteem
FREQUENCIES VARIABLES=Age Self_esteem
/HISTOGRAM NORMAL
/ORDER=ANALYSIS.
*Filter cases (use only youths who have not come out to
  anyone)
USE ALL.
COMPUTE filter_$=(Coming_Out=1).
VARIABLE LABELS filter_$ 'Coming_Out=1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
*Filter cases (use only youths who have come out to one or
  more people)
USE ALL.
COMPUTE filter_$=(Coming_Out<>1).
VARIABLE LABELS filter_$ 'Coming_Out<>1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
```

```
FILTER BY filter _ $.
EXECUTE.
*Remove any active filter
FILTER OFF.
USE ALL.
EXECUTE.
*Correlation with all youths
NONPAR CORR
/VARIABLES=Age Self _ esteem
/PRINT=BOTH TWOTAIL NOSIG
/MISSING=PAIRWISE.
*Sort cases
SORT CASES BY Age(A).
*filter and correlation with only youths who have not come
  out to anyone
USE ALL.
COMPUTE filter _ $=(Coming _ Out=1).
VARIABLE LABELS filter _ $ 'Coming _ Out=1 (FILTER)'.
VALUE LABELS filter _ $ 0 'Not Selected' 1 'Selected'.
FORMATS filter _ $ (f1.0).
FILTER BY filter _ $.
EXECUTE.
NONPAR CORR
/VARIABLES=Age Self _ esteem
/PRINT=BOTH TWOTAIL NOSIG
/MISSING=PAIRWISE.
*filter and correlation with only youths who have come out
  to one or more people
USE ALL.
COMPUTE filter _ $=(Coming _ Out<>1).
VARIABLE LABELS filter _ $ 'Coming _ Out<>1 (FILTER)'.
VALUE LABELS filter _ $ 0 'Not Selected' 1 'Selected'.
FORMATS filter _ $ (f1.0).
FILTER BY filter _ $.
EXECUTE.
NONPAR CORR
/VARIABLES=Age Self _ esteem
/PRINT=BOTH TWOTAIL NOSIG
/MISSING=PAIRWISE.
*Remove any active filter
FILTER OFF.
USE ALL.
EXECUTE.
```

CHAPTER 3

Research Question #7

```
*Independent Group Comparison and CI
*Nonparametric Tests: Independent Samples.
NPTESTS
/INDEPENDENT TEST (Percent _ Pre) GROUP (Received _ MI)
MOSES(TRIMOUTLIERS=SAMPLE) HODGES _ LEHMANN
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Research Question #8

```
*Independent Group Analysis:
*Nonparametric Tests: Independent Samples.
NPTESTS
/INDEPENDENT TEST (Percent _ DIF) GROUP (Received _ MI)
KOLMOGOROV _ SMIRNOV
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
*Independent Group Analysis using three tests:
*Nonparametric Tests: Independent Samples.
NPTESTS
/INDEPENDENT TEST (Percent _ DIF) GROUP (Received _ MI)
MANN _ WHITNEY KOLMOGOROV _ SMIRNOV WALD _ WOLFOWITZ
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
*Frequency for MI=1 respondents
USE ALL.
COMPUTE filter _ $=(Received _ MI=1).
VARIABLE LABELS filter _ $ 'Received _ MI=1 (FILTER)'.
VALUE LABELS filter _ $ 0 'Not Selected' 1 'Selected'.
FORMATS filter _ $ (f1.0).
FILTER BY filter _ $.
EXECUTE.
FREQUENCIES VARIABLES=Percent _ DIF
/HISTOGRAM NORMAL
/ORDER=ANALYSIS.
*Frequency for MI=2 respondents
USE ALL.
COMPUTE filter _ $=(Received _ MI=2).
VARIABLE LABELS filter _ $ 'Received _ MI=2 (FILTER)'.
```

```
VALUE LABELS filter _ $ 0 'Not Selected' 1 'Selected'.
FORMATS filter _ $ (f1.0).
FILTER BY filter _ $.
EXECUTE.
FREQUENCIES VARIABLES=Percent _ DIF
/HISTOGRAM NORMAL
/ORDER=ANALYSIS.
```

Research Question #9

```
*Filter to get only respondents in the treatment group:
USE ALL.
COMPUTE filter _ $=(Received _ MI=1).
VARIABLE LABELS filter _ $ 'Received _ MI=1 (FILTER)'.
VALUE LABELS filter _ $ 0 'Not Selected' 1 'Selected'.
FORMATS filter _ $ (f1.0).
FILTER BY filter _ $.
EXECUTE.
*Nonparametric Tests: Independent Samples.
NPTESTS
/INDEPENDENT TEST (Percent _ Post) GROUP (Partner)
KRUSKAL _ WALLIS(COMPARE=PAIRWISE)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Research Question #10

```
*Filter to get only respondents in the treatment group:
USE ALL.
COMPUTE filter _ $=(Received _ MI=1).
VARIABLE LABELS filter _ $ 'Received _ MI=1 (FILTER)'.
VALUE LABELS filter _ $ 0 'Not Selected' 1 'Selected'.
FORMATS filter _ $ (f1.0).
FILTER BY filter _ $.
EXECUTE.
*Jonckheere _ Terpstra test:
*Nonparametric Tests: Independent Samples.
NPTESTS
/INDEPENDENT TEST (Percent _ DIF) GROUP (Stage _ of _ Change)
JONCKHEERE _ TERPSTRA(ORDER=ASCENDING COMPARE=PAIRWISE)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
*Nonparametric Tests: Independent Samples.
```

```
NPTESTS
/INDEPENDENT TEST (Percent _ DIF) GROUP (Stage _ of _ Change)
JONCKHEERE _ TERPSTRA(ORDER=ASCENDING COMPARE=PAIRWISE)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.1 CILEVEL=95.
```

Research Question #11

```
* set filter to only use respondents in treatment group
USE ALL.
COMPUTE filter _ $=(Received _ MI=1).
VARIABLE LABELS filter _ $ 'Received _ MI=1 (FILTER)'.
VALUE LABELS filter _ $ 0 'Not Selected' 1 'Selected'.
FORMATS filter _ $ (f1.0).
FILTER BY filter _ $.
EXECUTE.
*Nonparametric Tests: Independent Samples.
NPTESTS
/INDEPENDENT TEST (Self _ efficacy) GROUP (Condom _ Request)
MEDIAN(TESTVALUE=SAMPLE COMPARE=PAIRWISE)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

CHAPTER 4

Research Question #12

```
* Analysis using McNemar test:
*Nonparametric Tests: Related Samples.
NPTESTS
/RELATED TEST(Violent _ crime _ w1 Violent _ crime _ w2)
MCNEMAR(SUCCESS=LIST(1))
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Research Question #13

```
*Descriptive Analyses:
FREQUENCIES VARIABLES=Friends _ w2 Friends _ w4
/ORDER=ANALYSIS.
*Nonparametric Tests: Related Samples.
NPTESTS
```

```
/RELATED TEST(Friends _ w2 Friends _ w4)
MARGINAL _ HOMOGENEITY
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Research Question #14

```
* Analysis using Sign test:
*Nonparametric Tests: Related Samples.
NPTESTS
/RELATED TEST(Anxiety _ w2 Anxiety _ w4) SIGN
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Research Question #15

```
*Analysis using Wilcoxon Signed-Rank test and CI:
*Nonparametric Tests: Related Samples.
NPTESTS
/RELATED TEST(Anxiety _ w2 Anxiety _ w4)
WILCOXON HODGES _ LEHMAN
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.1 CILEVEL=95.
COMPUTE Median _ DIFF=Anxiety _ w2 - Anxiety _ w4.
EXECUTE.
```

Research Question #16

```
*Analysis using Cochran's Q test:
*Nonparametric Tests: Related Samples.
NPTESTS
/RELATED TEST(Assault _ w2 Assault _ w3 Assault _ w4)
COCHRAN(SUCCESS=LIST(1) COMPARE=PAIRWISE)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Research Question #17

```
*Analysis using Kendall's Coefficient of Concordance:
*Nonparametric Tests: Related Samples.
NPTESTS
/RELATED TEST(Support _ w1 Support _ w2 Support _ w3 Support _ w4)
```

```
KENDALL(COMPARE=PAIRWISE)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

Research Question #18

```
*Analysis using Friedman's Two-Way ANOVA test:
*Nonparametric Tests: Related Samples.
NPTESTS
/RELATED TEST(Support _ w1 Support _ w2 Support _ w3 Support _ w4)
FRIEDMAN(COMPARE=PAIRWISE)
/MISSING SCOPE=ANALYSIS USERMISSING=EXCLUDE
/CRITERIA ALPHA=0.05 CILEVEL=95.
```

CHAPTER 5

Research Question #19

n/a

Research Question #20

n/a

Research Question #21

```
*Factor Analysis:
CROSSTABS
/TABLES=Edu _ Level BY Gender
/FORMAT=AVALUE TABLES
/CELLS=COUNT
/COUNT ROUND CELL.
UNIANOVA LifeSatisfaction BY Gender Employment
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/POSTHOC=Gender(LSD BONFERRONI)
/PLOT=PROFILE(Gender*Employment Employment*Gender)
/EMMEANS=TABLES(OVERALL)
/PRINT=LOF HOMOGENEITY DESCRIPTIVE PARAMETER
/PLOT=RESIDUALS
/CRITERIA=ALPHA(.05)
```

```
/DESIGN=Gender Employment Gender*Employment.
UNIANOVA LifeSatisfaction BY Gender Employment
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/POSTHOC=Gender(LSD BONFERRONI)
/PLOT=PROFILE(Gender*Employment Employment*Gender)
/EMMEANS=TABLES(OVERALL)
/PRINT=LOF HOMOGENEITY DESCRIPTIVE PARAMETER
/PLOT=RESIDUALS
/CRITERIA=ALPHA(.05)
/DESIGN=Gender.
```

Appendix B
Missing Data

Most researchers will have to deal with missing data at some point in their careers. The challenge of missing data is that there are many reasons for the missingness, and probably even more ways in which to address missingness in a data set. As you may remember from your introduction to statistics course, missing data fall into three categories—missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). MCAR and MAR are, for the most part, manageable when it comes to dealing with the gaps in the information (i.e., the missing data).

However, MNAR can be problematic. For example, the gaps in information might be due to how a few respondents reacted to the data-collection procedures—say they interpreted a survey question differently from the majority of other respondents, or perhaps they elected to skip a response for personal reasons outside of the study issues. This systematic missingness undermines the accuracy of your findings, and in some cases can invalidate your study altogether, without your even recognizing it. For this reason, whenever missing data are present, you must conduct an extensive review of the reliability and validity of the study methods. MNAR can be caused by issues in the data collection methods, such as bias in sampling methods, uneven interpretation of survey questions, and unbalanced response options on a question—just to name a few. Graphing the variables, comparing groups of respondents, asking others to offer reasons why a respondent would skip questions, and looking to see if any variables in the data set can predict which respondents have missingness are just a few ways to examine a data set for reasons for potential MNAR. This examination is necessary to identify what type of missingness exists, to detect an explanation for the missingness, and to help you decide what to do next.

Unfortunately, exploring the reasons for the missing data is the easy part. Deciding what to do about the missing data can be much more challenging, in part because there are many ways to address missingness and thoughtful scholars occasionally disagree. The disagreement is primarily due to the varying levels in which researchers are comfortable using information from other respondents to fill in the gaps. With smaller data sets, a limited amount of information is available, and using this limited amount to estimate other values that are missing is ill advised. In a small data set, not enough information is available to effectively estimate the missing values. This leaves very few options when dealing with missing data in these small data sets.

When making these decisions, return to the primary responsibility of any researcher—to search for accurate answers using only the information (both data and theory) available. Therefore, in these situations it is recommended to use only listwise (also called casewise) deletion for dealing with missingness. This process deletes the cases with missing values for variables included in your analyses. By using listwise deletion, you are acknowledging that the data are in some fashion missing at random (i.e., not MNAR) and that replacing the missing data with values from a type of statistical procedure (e.g., imputation, partial imputation, mean replacement, etc.) would only create additional bias in your findings. (For more information on missing data, see Allison, 2001; Enders, 2010; Rubin & Roderick, 2002.)

Appendix C
Other Resources

Appendix C contains a list of resources for additional information about dealing with small samples, using nonparametric procedures, and statistics in general. The information below is in no particular order of importance or quality, and is in no way an exhaustive list of helpful options. The resources are grouped into three categories—books (and book chapters), articles, and websites. If you are looking for an extremely detailed and comprehensive statistics-driven reference book, we highly recommend *Nonparametric Statistical Methods* by Myles Hollander and Douglas A. Wolfe. However, if you are not ready for in-depth statistical explanations, we suggest *Nonparametric Statistics for Health Care Research,* Second Edition, by Marjorie A. Pett. The articles listed below are just a few resources that support ideas and concepts discussed in this book, but if you have time to read only one article, we strongly suggest you read David Kenny's online article titled *Measuring Model Fit*. The listed websites will take you to web locations that are rich with information and how-to examples for all different levels of statistical experience. Our favorite site happens to be the one managed by UCLA, which offers so much information you can spend years reading their content on statistical procedures.

References

CHAPTER 1

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods, 16*(2), 270–301. doi:10.1177/1094428112470848

Barnett, V., & Lewis, T. (1978). *Outliers in statistical data*. Chichester: Wiley.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2d ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159. doi:10.1037/0033-2909.112.1.155

Corder, G. W., & Foreman, D. I. (2009). *Nonparametric statistics for non-statisticians.* Hoboken, NJ: Wiley.

Davidson, A. C., & Hinkley, D. V. (1998). *Bootstrap methods and their application*. Cambridge, U.K.: Cambridge University Press.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics, 7,* 1–26. https://doi.org/10.1214/aos/1176344552

Elliott, A. C. (2015). *Determine equality of variance for two groups.* http://stattutorials.com/EXCEL/EXCEL-EQUAL-VARIANCE.html

Gaskin, J. (2015). *Plotting homoscedasticity in SPSS* [Video file]. Available from https://www.youtube.com/watch?v=WrQ1O_He63Q

Hojat, M., & Xu, G. (2004). A visitor's guide to effect sizes: Statistical significance versus practical (clinical) importance of research findings. *Advances in Health Sciences Education, 9,* 241–249. https://doi.org/10.1023/b:ahse.0000038173.00909.f6

Hoskin, T. (n.d.) *Parametric and nonparametric: Demystifying the terms.* Retrieved from http://www.mayo.edu/mayo-edu-docs/center-for-translational-science-activities-documents/berd-5-6.pdf

Hunter, M. A., & May, R. B. (1993). Some myths concerning parametric and nonparametric tests. *Canadian Psychology, 34*, 384–389. https://doi.org/10.1037/h0078860

Janssen, A., & Pauls, T. (2003). How do bootstrap and permutation tests work?. *Annals of Statistics, 31*(3), 768–806. https://doi.org/10.1214/aos/1056562462

Mathews, P. (2010). *Sample size calculations: Practical methods for engineers and scientists.* Fairport Harbor, OH: Mathews Malnar and Bailey, Inc.

McDonald, J. H. (2014). *Handbook of biological statistics* (3d ed.). Baltimore, MD: Sparky House Publishing.

Mumby, P. J. (2002). Statistical power of non-parametric tests: A quick guide for designing sampling strategies. *Marine Pollution Bulletin, 44*, 85–87. https://doi.org/10.1016/s0025-326x(01)00097-2

NIST/SEMATECH (2012). *e-Handbook of statistical methods.* Retrieved from http://www.itl.nist.gov/div898/handbook/pri/section2/pri24.htm

Ryan, T. P. (2013). *Sample size determination and power.* Hoboken, NJ: Wiley.

Sawilowsky, S. (2005). Misconceptions leading to choosing the *t*-test over the Wilcoxon Mann-Whitney test for shift in location parameter. *Journal of Applied Statistical Methods, 4,* 598–600. Available from http://digitalcommons.wayne.edu/coe_tbf/12

Stat Trek (2016). *Transformations to achieve linearity.* Retrieved from http://stattrek.com/regression/linear-transformation.aspx?Tutorial=AP

StatSoft (2000). *Multiple regression.* Retrieved from http://www.uta.edu/faculty/sawasthi/Statistics/stmulreg.html#aassumption

Taylor, C. (n.d.). *How do we determine what is an outlier?* Retrieved from http://statistics.about.com/od/Descriptive-Statistics/a/How-Do-We-Determine-What-Is-An-Outlier.htm

Wasserman, R. L., & Lazar, N. A. (2016). The ASA's statement on *p*-values: Context, process, and purpose. *The American Statistician, 70,* 129–133. http://dx.doi.org/10.1080/00031305.2016.1154108

Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *The Annals of Statistics, 14,* 1261–1295. https://doi.org/10.1214/aos/1176350142

CHAPTER 2

Cochran, W. G. (1952). The χ2 test of goodness of fit. *Annals of Mathematical Statistics, 25*, 315–345. https://doi.org/10.1214/aoms/1177729380

CHAPTER 3

Tsangari, H., & Akritas, M. G. (2004). Nonparametric ANCOVA with two and three covariates. *Journal of Multivariate Analysis, 88,* 298–319. https://doi.org/10.1016/s0047-259x(03)00098-8

CHAPTER 5

Chatterjee, S., & Hadi, A. S. (2012). *Regression analysis by example*. Hoboken, NJ: Wiley.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* **7**4, 368, 829–836. doi:10.2307/2286407

Cleveland, W. S. (1981). LOWESS: A program for smoothing scatterplots by robust locally weighted regression. *The American Statistician 35*, 54. https://doi.org/10.2307/2683591

Cleveland, W. S., & Devlin, S. J. (1988). Locally-weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association, 83,* 596–610. doi:10.2307/2289282

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). Internet, phone, mail, and mixed-mode surveys: The Tailored Design Method. Hoboken, NJ: Wiley.

Grande, T. L. (2015). Testing the assumption of independent errors with ZERESID, ZPRED, and Durbin-Watson using SPSS. Retrieved from https://www.youtube.com/watch?v=Lyk_S9T-oCw

Raudenbush, S.W. (1993). Hierarchical linear models and experimental design. In L. K. Edwards (Ed.), *Applied analysis of variance in behavioral science*. New York: Marcel Dekker.

APPENDIX B

Allison, P. D. (2001). *Missing data*. Thousand Oaks, CA: Sage Publications, Inc.

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.

Rubin, D. B., & Roderick J. A. L. (2002). *Statistical analysis with missing data*. New York: Wiley.

APPENDIX C

Books and Book Chapters

Corder, G. W., & Foreman, D. I. (2009). *Nonparametric statistics for non-statisticians: A step-by-step approach.* Hoboken, N.J: Wiley.

Deshpande, J. V., Gore, A. P., & Shanubhogue, A. (1995). *Statistical analysis of nonnormal data*. Hoboken, NJ: Wiley Interscience.

Gibbons, J. D. (1993). *Nonparametric statistics: An introduction*. Newbury Park, CA: Sage Publications.

Green, S. B., & Salkind, N. J. (2014). *Using SPSS for Windows and Macintosh: Analyzing and understanding data*. New York: Pearson.

Grotenhuis, M., & Matthijssen, A. (2016). *Basic SPSS tutorial.* Thousand Oaks, CA: Sage.

Hollander, M., & Wolfe, D. A. (1999). *Nonparametric statistical methods.* New York: Wiley.

Hoyle, R. H. (1999). *Statistical strategies for small sample research.* Thousand Oaks, CA: Sage Publications.

Huff, D. (1993). *How to lie with statistics.* New York: Norton.

Krauth, J. (1988). *Distribution-free statistics: An application-oriented approach.* Amsterdam: Elsevier.

Layder, D. (2013). *Doing excellent small-scale research.* Thousand Oaks, CA: Sage Publications.

Noether, G. E., & Dueker, M. (1991). *Introduction to statistics: The nonparametric way.* New York: Springer-Verlag.

Petersen, M. A. (2005). In K. Sijtsma & I. W. Molenaar (Eds.), *Introduction to nonparametric item response theory*, Sage Publications Inc., Thousand Oaks, CA, 2002, 168 pp. ISBN: 0-7619-0812-9.

Pett, M. A. (2016). *Nonparametric statistics for health care research: Statistics for small samples and unusual distributions.* Thousand Oaks, CA: Sage Publications.

Salkind, N. J. (2010). *Statistics for people who (think they) hate statistics: Excel 2007 edition*. Thousand Oaks, CA: Sage Publications.

Sprent, P., & Smeeton, N. C. (2001). *Applied nonparametric statistical methods.* Boca Raton, FL: Chapman & Hall/CRC.

Wilcox, R. R. (2009). *Basic statistics: Understanding conventional methods and modern insights.* New York: Oxford.

Wilcox, R. R. (2009). Nonparametric estimation. In D. A. Belsey & E. Kontoghiorghes (Eds.), *Handbook of computational econometrics* (pp. 153–182). West Sussex, U.K.: Wiley

Articles

Carroll, R. J., Ruppert, D., Crainiceanu, C. M., Tosteson, T. D., & Karagas, M. R. (2004). Nonlinear and nonparametric regression and instrumental variables. *Journal of the American Statistical Association, 99*, 736–750. http://dx.doi.org/10.1198/016214504000001088

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association, 32*, 675–701. http://dx.doi.org/10.1080/01621459.1937.10503522

Friedman, M. (1939). A correction: The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association, 34*, 109. http://dx.doi.org/10.2307/2279169

Geyer, C. J. (2005). *Fuzzy p-values and ties in nonparametric tests*. Retrieved from http://www.stat.umn.edu/geyer/fuzz/ties.pdf

Han, L. (2008). *Calculating the point estimate and confidence interval of Hodges-Lehmann's median using SAS software.* Proceedings from the Southeastern SAS Users' Group (SESUG). Retrieved November 20, 2016 from http://analytics.ncsu.edu/sesug/2008/ST-154.pdf

Kenny, D. (November 24, 2015). *Measuring model fit.* Retrieved from http://davidakenny.net/cm/fit.htm

Kraemer, H. C. (1976). The small sample nonnull properties of Kendall's coefficient of concordance for normal populations. *Journal of the American Statistical Association, 71*(355), 608–613. http://dx.doi.org/10.1080/01621459.1976.10481536

Park, H. M. (2003). *Estimating regression models for categorical dependent variables using SAS, Stata, LIMDEP, and SPSS*. Retrieved from https://www.ikhebeenvraag.be/mediastorage/FSDocument/101/LogitProbit.pdf

Sawilowsky, S. S. (1990). Nonparametric tests of interaction in experimental design. *Review of Educational Research, 60*(1), 91–126. http://dx.doi.org/10.3102/00346543060001091

Steele, J. M. (2005). Darrell Huff and fifty years of how to lie with statistics. *Statistical Science, 20*, 205–209. http://dx.doi.org/10.1214/088342305000000205

Warner, C. B., & Meehan, A. M. (2001). Microsoft Excel as a tool for teaching basic statistics. *Teaching of Psychology, 28*, 295–298. http://dx.doi.org/10.1207/S15328023TOP2804_11

Wilcox, R. (2009). Comparing Pearson correlations: Dealing with heteroscedasticity and non-normality. *Communications in Statistics—Simulation and Computation, 38*, 2220–2234. http://dx.doi.org/10.1080/03610910903289151

Wilcox, R. R. (2005). Depth and a multivariate generalization of the Wilcoxon-Mann-Whitney test. *American Journal of Mathematical and Management Sciences, 25*, 343–361. http://dx.doi.org/10.1080/01966324.2005.10737655

Websites

Data analysis examples by Institute for Digital Research and Education at UCLA http://www.ats.ucla.edu/stat/dae/

Downloadable books listed by Institute for Digital Research and Education at UCLA http://www.ats.ucla.edu/stat/books/downloadable_books.htm

Resources to help you learn and use SPSS by Institute for Digital Research and Education at UCLA http://www.ats.ucla.edu/stat/spss/

Nonparametric methods by StatsDirect http://www.statsdirect.com/help/default.htm#nonparametric_methods/nonmarametric_methods.htm%3FTocPath%3DNonparametric%2520methods%7C_____0

Nonparametric tests by IBM Knowledge Center http://www-01.ibm.com/support/knowledgecenter/SSLVMB_20.0.0/com.ibm.spss.statistics.help/nonparametric_tests_intro_legacy.htm

Index