Nonparametric Statistics

A Step-by-Step Approach

Second Edition

Gregory W. Corder • Dale I. Foreman

WWW. LINK AVAILABLE

WILEY

NONPARAMETRIC STATISTICS

SECOND EDITION

NONPARAMETRIC STATISTICS

A Step-by-Step Approach

GREGORY W. CORDER

DALE I. FOREMAN

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Corder, Gregory W., 1972– author.
Nonparametric statistics : a step-by-step approach / Gregory W. Corder. Dale I. Foreman – Second edition.
pages cm
Includes bibliographical references and index.
ISBN 978-1-118-84031-3 (hardback)
1. Nonparametric statistics. I. Foreman, Dale I., author II. Title.
QA278.8.C67 2014
519.5′4–dc23

2013042040

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

CONTENTS

PREFACE

The social, behavioral, and health sciences have a need for the ability to use nonparametric statistics in research. Many studies in these areas involve data that are classified in the nominal or ordinal scale. At times, interval data from these fields lack parameters for classification as normal. Nonparametric statistical tests are useful tools for analyzing such data.

Purpose of This Book

This book is intended to provide a conceptual and procedural approach for nonparametric statistics. It is written so that someone who does not have an extensive mathematical background may work through the process necessary to conduct the given statistical tests presented. In addition, the outcome includes a discussion of the final decision for each statistical test. Each chapter takes the reader through an example from the beginning hypotheses, through the statistical calculations, to the final decision as compared with the hypothesis. The examples are then followed by a detailed, step-by-step analysis using the computer program SPSS®. Finally, research literature is identified which uses the respective nonparametric statistical tests.

Intended Audience

While not limited to such, this book is written for graduate and undergraduate students in social science programs. As stated earlier, it is targeted toward the student who does not have an especially strong mathematical background, but can be used effectively with a mixed group of students that includes students who have both strong and weak mathematical background.

Special Features of This Book

There are currently few books available that provide a practical and applied approach to teaching nonparametric statistics. Many books take a more theoretical approach to the instructional process that can leave students disconnected and frustrated, in need of supplementary material to give them the ability to apply the statistics taught.

It is our hope and expectation that this book provides students with a concrete approach to performing the nonparametric statistical procedures, along with their application and interpretation. We chose these particular nonparametric procedures since they represent a breadth of the typical types of analyses found in social science research. It is our hope that students will confidently learn the content presented with the promise of future successful applications.

In addition, each statistical test includes a section that explains how to use the computer program SPSS. However, the organization of the book provides effective instruction of the nonparametric statistical procedures for those individuals with or without the software. Therefore, instructors (and students) can focus on learning the tests with a calculator, SPSS, or both.

A Note to the Student

We have written this book with you in mind. Each of us has had a great deal of experience working with students just like you. Over the course of that time, it has been our experience that most people outside of the fields of mathematics or hard sciences struggle with and are intimidated by statistics. Moreover, we have found that when statistical procedures are explicitly communicated in a step-by-step manner, almost anyone can use them.

This book begins with a brief introduction (Chapter 1) and is followed with an explanation of how to perform the crucial step of checking your data for normality (Chapter 2). The chapters that follow (Chapters 3–9) highlight several nonparametric statistical procedures. Each of those chapters focuses on a particular type of variable and/or sample condition.

Chapters 3–9 each have a similar organization. They each explain the statistical methods included in their respective chapters. At least one sample problem is included for each test using a step-by-step approach. (In some cases, we provide additional sample problems when procedures differ between large and small samples.) Then, those same sample problems are demonstrated using the statistical software package SPSS. Whether or not your instructor incorporates SPSS, this section will give you the opportunity to learn how to use the program. Toward the end of each chapter, we identify examples of the tests in published research. Finally, we present sample problems with solutions.

As you seek to learn nonparametric statistics, we strongly encourage you to work through the sample problems. Then, using the sample problems as a reference, work through the problems at the end of the chapters and additional data sets provided.

New to the Second Edition

Given an opportunity to write a second edition of this book, we revised and expanded several portions. Our changes are based on feedback from users and reviewers.

We asked several undergraduate and graduate students for feedback on Chapters 1 and 2. Based on their suggestions, we made several minor changes to Chapter 1 with a goal to improve understanding. In Chapter 2, we expanded the section that describes and demonstrates the Kolmogorov–Smirnov (K-S) one-sample test.

After examining current statistics textbooks and emerging research paper, we decided to include two additional tests. We added the sign test to Chapter 3 and the Kolmogorov–Smirnov (K-S) two-sample test to Chapter 4. We also added a discussion on statistical power to Chapter 3 as requested by instructors who had adopted our book for their courses.

Since our book's first edition, SPSS has undergone several version updates. Our new edition of the book also has updated directions and screen captures for images of SPSS. Specifically, these changes reflect SPSS version 21.

We have included web-based tools to support our book's new edition. If you visit the publisher's book support website, you will find a link to a Youtube channel that includes narrated screen casts. The screen casts demonstrate how to use SPSS to perform the tests included in this book. The publisher's book support website also includes a link to a decision tree that helps the user determine an appropriate type of statistical test. The decision tree is organized using Prezi. The branches terminate with links to the screen casts on YouTube.

Gregory W. Corder
Dale I. Foreman

LIST OF VARIABLES

English Symbols

C	number of columns in a contingency table; number of categories
C_F	tie correction for the Friedman test
C_H	tie correction for the Kruskal–Wallis H-test
D, $\tilde{D}$	divergence between values from cumulative frequency distributions
D_i	difference between a ranked pair
df	degrees of freedom
f_e	expected frequency
f_o	observed frequency
$\hat{f}_r$	empirical frequency value
F_r	Friedman test statistic
g	number of tied groups in a variable
h	correction for continuity
H	Kruskal–Wallis test statistic
H_A	alternate hypothesis
H_O	null hypothesis
k	number of groups
K	kurtosis
M	midpoint of a sample
n	sample size
N	total number of values in a contingency table
p	probability
P_i	a category's proportion with respect to all categories
r_b	biserial correlation coefficient for a sample
r_{pb}	point-biserial correlation coefficient for a sample
r_s	Spearman rank-order correlation coefficient for a sample
R	number of runs; number of rows in a contingency table
R_i	sum of the ranks from a particular sample
s	standard deviation of a sample
S_k	skewness
SE	standard error
t	t statistic
t_i	number of tied values in a tie group
T	Wilcoxon signed rank test statistic
U	Mann–Whitney test statistic
$\bar{x}$	sample mean
y	height of the unit normal curve ordinate at the point dividing two proportions

z	the number of standard deviations away from the mean
Z	Kolmogorov–Smirnov test statistic

Greek Symbols

α	alpha, probability of making a type I error
α_B	adjusted level of risk using the Bonferroni procedure
β	beta, probability of making a type II error
θ	theta, median of a population
μ	mu, mean value for a population
ρ	rho, correlation coefficient for a population
σ	sigma, standard deviation of a population
Σ	sigma, summation
χ^2	chi-square test statistic

CHAPTER 1

NONPARAMETRIC STATISTICS: AN INTRODUCTION

1.1 OBJECTIVES

In this chapter, you will learn the following items:

- The difference between parametric and nonparametric statistics.
- How to rank data.
- How to determine counts of observations.

1.2 INTRODUCTION

If you are using this book, it is possible that you have taken some type of introductory statistics class in the past. Most likely, your class began with a discussion about probability and later focused on particular methods of dealing with populations and samples. Correlations, z-scores, and t-tests were just some of the tools you might have used to describe populations and/or make inferences about a population using a simple random sample.

Many of the tests in a traditional, introductory statistics text are based on samples that follow certain assumptions called parameters. Such tests are called *parametric tests*. Specifically, parametric assumptions include samples that

- are randomly drawn from a normally distributed population,
- consist of independent observations, except for paired values,
- consist of values on an interval or ratio measurement scale,
- have respective populations of approximately equal variances,
- are adequately large,* and
- approximately resemble a normal distribution.

*The minimum sample size for using a parametric statistical test varies among texts. For example, Pett (1997) and Salkind (2004) noted that most researchers suggest $n > 30$. Warner (2008) encouraged considering $n > 20$ as a minimum and $n > 10$ per group as an absolute minimum.

Nonparametric Statistics: A Step-by-Step Approach, Second Edition. Gregory W. Corder and Dale I. Foreman.

If any of your samples breaks one of these rules, you violate the assumptions of a parametric test. You do have some options, however.

You might change the nature of your study so that your data meet the needed parameters. For instance, if you are using an ordinal or nominal measurement scale, you might redesign your study to use an interval or ratio scale. (See Box 1.1 for a description of measurement scales.) Also, you might seek additional participants to enlarge your sample sizes. Unfortunately, there are times when one or neither of these changes is appropriate or even possible.

BOX 1.1

MEASUREMENT SCALES.

We can measure and convey variables in several ways. *Nominal* data, also called categorical data, are represented by counting the number of times a particular event or condition occurs. For example, you might categorize the political alignment of a group of voters. Group members could either be labeled democratic, republican, independent, undecided, or other. No single person should fall into more than one category.

A *dichotomous* variable is a special classification of nominal data; it is simply a measure of two conditions. A dichotomous variable is either discrete or continuous. A *discrete dichotomous* variable has no particular order and might include such examples as gender (male vs. female) or a coin toss (heads vs. tails). A *continuous dichotomous* variable has some type of order to the two conditions and might include measurements such as pass/fail or young/old.

Ordinal scale data describe values that occur in some order of rank. However, distance between any two ordinal values holds no particular meaning. For example, imagine lining up a group of people according to height. It would be very unlikely that the individual heights would increase evenly. Another example of an ordinal scale is a Likert-type scale. This scale asks the respondent to make a judgment using a scale of three, five, or seven items. The range of such a scale might use a 1 to represent *strongly disagree* while a 5 might represent *strongly agree*. This type of scale can be considered an ordinal measurement since any two respondents will vary in their interpretation of scale values.

An *interval* scale is a measure in which the relative distances between any two sequential values are the same. To borrow an example from the physical sciences, we consider the Celsius scale for measuring temperature. An increase from −8 to −7°C degrees is identical to an increase from 55 to 56°C.

A *ratio* scale is slightly different from an interval scale. Unlike an interval scale, a ratio scale has an absolute zero value. In such a case, the zero value indicates a measurement limit or a complete absence of a particular condition. To borrow another example from the physical sciences, it would be appropriate to measure light intensity with a ratio scale. Total darkness is a complete absence of light and would receive a value of zero.

On a general note, we have presented a classification of measurement scales similar to those used in many introductory statistics texts. To the best of our knowledge, this hierarchy of scales was first made popular by Stevens (1946). While Stevens has received agreement (Stake, 1960; Townsend & Ashby, 1984) and criticism (Anderson, 1961; Gaito, 1980; Velleman & Wilkinson, 1993), we believe the scale classification we present suits the nature and organization of this book. We direct anyone seeking additional information on this subject to the preceding citations.

If your samples do not resemble a normal distribution, you might have learned a strategy that modifies your data for use with a parametric test. First, if you can justify your reasons, you might remove extreme values from your samples called outliers. For example, imagine that you test a group of children and you wish to generalize the findings to typical children in a normal state of mind. After you collect the test results, most children earn scores around 80% with some scoring above and below the average. Suppose, however, that one child scored a 5%. If you find that this child speaks no English because he arrived in your country just yesterday, it would be reasonable to exclude his score from your analysis. Unfortunately, outlier removal is rarely this straightforward and deserves a much more lengthy discussion than we offer here.* Second, you might utilize a parametric test by applying a mathematical transformation to the sample values. For example, you might square every value in a sample. However, some researchers argue that transformations are a form of data tampering or can distort the results. In addition, transformations do not always work, such as circumstances when data sets have particularly long tails. Third, there are more complicated methods for analyzing data that are beyond the scope of most introductory statistics texts. In such a case, you would be referred to a statistician.

Fortunately, there is a family of statistical tests that do not demand all the parameters, or rules, that we listed earlier. They are called *nonparametric tests*, and this book will focus on several such tests.

1.3 THE NONPARAMETRIC STATISTICAL PROCEDURES PRESENTED IN THIS BOOK

This book describes several popular nonparametric statistical procedures used in research today. Table 1.1 identifies an overview of the types of tests presented in this book and their parametric counterparts.

TABLE 1.1

Type of analysis	Nonparametric test	Parametric equivalent
Comparing two related samples	Wilcoxon signed ranks test and sign test	*t*-Test for dependent samples
Comparing two unrelated samples	Mann–Whitney *U*-test and Kolmogorov–Smirnov two-sample test	*t*-Test for independent samples
Comparing three or more related samples	Friedman test	Repeated measures, analysis of variance (ANOVA)
Comparing three or more unrelated samples	Kruskal–Wallis *H*-test	One-way ANOVA

(*Continued*)

*Malthouse (2001) and Osborne and Overbay (2004) presented discussions about the removal of outliers.

TABLE 1.1 (*Continued*)

Type of analysis	Nonparametric test	Parametric equivalent
Comparing categorical data	Chi square (χ^2) tests and Fisher exact test	None
Comparing two rank-ordered variables	Spearman rank-order correlation	Pearson product–moment correlation
Comparing two variables when one variable is discrete dichotomous	Point-biserial correlation	Pearson product–moment correlation
Comparing two variables when one variable is continuous dichotomous	Biserial correlation	Pearson product–moment correlation
Examining a sample for randomness	Runs test	None

When demonstrating each nonparametric procedure, we will use a particular step-by-step method.

1.3.1 State the Null and Research Hypotheses

First, we state the hypotheses for performing the test. The two types of hypotheses are null and alternate. The *null hypothesis* (H_O) is a statement that indicates no difference exists between conditions, groups, or variables. The *alternate hypothesis* (H_A), also called a research hypothesis, is the statement that predicts a difference or relationship between conditions, groups, or variables.

The alternate hypothesis may be directional or nondirectional, depending on the context of the research. A directional, or one-tailed, hypothesis predicts a statistically significant change in a particular direction. For example, a treatment that predicts an improvement would be directional. A nondirectional, or two-tailed, hypothesis predicts a statistically significant change, but in no particular direction. For example, a researcher may compare two new conditions and predict a difference between them. However, he or she would not predict which condition would show the largest result.

1.3.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis

When we perform a particular statistical test, there is always a chance that our result is due to chance instead of any real difference. For example, we might find that two samples are significantly different. Imagine, however, that no real difference exists. Our results would have led us to reject the null hypothesis when it was actually true. In this situation, we made a type I error. Therefore, statistical tests assume some level of risk that we call alpha, or α.

There is also a chance that our statistical results would lead us to not reject the null hypothesis. However, if a real difference actually does exist, then we made a type II error. We use the Greek letter beta, β, to represent a type II error. See Table 1.2 for a summary of type I and type II errors.

TABLE 1.2

	We do not reject the null hypothesis	We reject the null hypothesis
The null hypothesis is actually true	No error	Type-I error, α
The null hypothesis is actually false	Type-II error, β	No error

After the hypotheses are stated, we choose the level of risk (or the level of significance) associated with the null hypothesis. We use the commonly accepted value of $\alpha = 0.05$. By using this value, there is a 95% chance that our statistical findings are real and not due to chance.

1.3.3 Choose the Appropriate Test Statistic

We choose a particular type of test statistic based on characteristics of the data. For example, the number of samples or groups should be considered. Some tests are appropriate for two samples, while other tests are appropriate for three or more samples.

Measurement scale also plays an important role in choosing an appropriate test statistic. We might select one set of tests for nominal data and a different set for ordinal variables. A common ordinal measure used in social and behavioral science research is the Likert scale. Nanna and Sawilowsky (1998) suggested that nonparametric tests are more appropriate for analyses involving Likert scales.

1.3.4 Compute the Test Statistic

The test statistic, or obtained value, is a computed value based on the particular test you need. Moreover, the method for determining the obtained value is described in each chapter and varies from test to test. For small samples, we use a procedure specific to a particular statistical test. For large samples, we approximate our data to a normal distribution and calculate a z-score for our data.

1.3.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic

For small samples, we reference a table of critical values located in Appendix B. Each table provides a critical value to which we compare a computed test statistic. Finding a critical value using a table may require you to use such data characteristics