

THIRD EDITION

STATISTICS

in Criminology
and Criminal Justice

Analysis and Interpretation

Jeffery T. Walker and Sean Maddan

Statistics in Criminology and Criminal Justice Analysis and Interpretation

Third Edition

Jeffery T. Walker, PhD

Professor

Department of Criminal Justice
University of Arkansas–Little Rock

Sean Maddan, PhD

Assistant Professor

Department of Criminology and Criminal Justice
University of Tampa

JONES AND BARTLETT PUBLISHERS

Sudbury, Massachusetts

BOSTON TORONTO LONDON SINGAPORE

World Headquarters

Jones and Bartlett Publishers
40 Tall Pine Drive
Sudbury, MA 01776
978-443-5000
info@jbpub.com
www.jbpub.com

Jones and Bartlett Publishers
Canada
6339 Ormindale Way
Mississauga, Ontario L5V 1J2
Canada

Jones and Bartlett Publishers
International
Barb House, Barb Mews
London W6 7PA
United Kingdom

Jones and Bartlett's books and products are available through most bookstores and online booksellers. To contact Jones and Bartlett Publishers directly, call 800-832-0034, fax 978-443-8000, or visit our website www.jbpub.com.

Substantial discounts on bulk quantities of Jones and Bartlett's publications are available to corporations, professional associations, and other qualified organizations. For details and specific discount information, contact the special sales department at Jones and Bartlett via the above contact information or send an email to specialsales@jbpub.com.

Copyright © 2009 by Jones and Bartlett Publishers, LLC

All rights reserved. No part of the material protected by this copyright may be reproduced or utilized in any form, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without written permission from the copyright owner.

This publication is designed to provide accurate and authoritative information in regard to the Subject Matter covered. It is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional service. If legal advice or other expert assistance is required, the service of a competent professional person should be sought.

Production Credits

Acquisitions Editor: Jeremy Spiegel
Editorial Assistant: Maro Asadoorian
Production Director: Amy Rose
Senior Production Editor: Renée Sekerak
Production Assistant: Julia Waugaman
Associate Marketing Manager: Lisa Gordon
Manufacturing and Inventory Control Supervisor: Amy Bacus
Cover Design: Kristin E. Ohlin
Composition: Northeast Compositors, Inc.
Cover Image: © Joseph/Shutterstock, Inc.
Chapter Opener Image: © Stephen Sweet/Shutterstock, Inc.
Printing and Binding: Malloy Incorporated
Cover Printing: Malloy Incorporated

Library of Congress Cataloging-in-Publication Data

Walker, Jeffery T.

Statistics in criminology and criminal justice / by Jeffery T. Walker and Sean Maddan. — 3rd ed.
p. cm.

Includes bibliographical references and index.

ISBN 978-0-7637-5548-5 (hardcover)

1. Criminal justice, Administration of—Statistical methods. 2. Criminal statistics—Research. I. Maddan, Sean. II. Title.

HV7415.W32 2009

364.01'5195—dc22

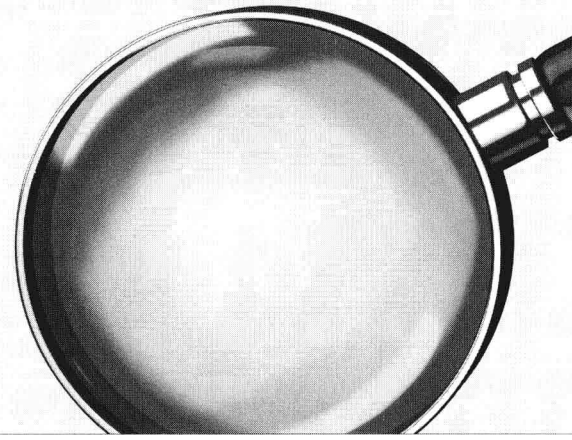
2008018116

6048

Printed in the United States of America

12 11 10 09 08 10 9 8 7 6 5 4 3 2 1

Contents



Use of SPSS and Data Sets with This Book xi

1 The Logic of Comparisons and Analysis. 2

Introduction: Why Analyze Data?	3
Some Statistical History	3
Uses of Statistics	4
Theory Construction at a Glance	5
What Is Theory?	5
Theory and Research	5
The Process of Scientific Inquiry	6
Observation and Inquisitiveness	6
Primary Questions	8
Research Questions	8
Research: Movement from Theory to Data and Back	8
Formulating Hypotheses	9
Constructing the Research Design	10
Conceptualization	10
Operationalization	11
Gathering the Data	12
Statistical Analysis: The Art of Making Comparisons	14
Foundations of Valid Comparisons	14
Comparing Appropriate Phenomena	15
Using Comparable Measures	15
Choosing Analysis Methods That Best Summarize the Data	16
Drawing Conclusions	16
Communicating the Results	17
Data and Purposes of This Book	17
Key Terms	19
Exercises	20
References	20
For Further Reading	21
Notes	21

2 Variables and Measurement 22

The Variable Defined	23
Transforming Characteristics into Data: The Process of Measurement	23
How Variables Can Differ	25
Levels of Measurement	26

Scale Continuity	36
Use in the Research Process	37
Conclusions	40
Key Terms	40
Exercises	40
References	44
For Further Reading	45
Notes	45
3 Understanding Data Through Organization	46
Frequency Distributions: A Chart of a Different Color	49
Conventions for Building Distributions	49
Frequency Distributions	52
Percentage Distributions	54
Combination Distributions	56
Graphical Representation of Frequencies	56
Pie Charts	56
Histograms and Bar Charts	57
Polygons and Area Charts	61
Analyzing Univariate Statistics	62
Analyzing Change	64
Line Charts	64
Ogives	64
Analyzing Bivariate and Multivariate Data	65
Scatter Plots	65
Normal Probability Plots	67
Path Diagrams	68
Analyzing Geographic Distributions	69
Pin, Spot, or Point Maps	69
Choropleth Maps	70
Conclusion	74
Key Terms	75
Exercises	75
References	77
Notes	77
4 Measures of Central Tendency	78
Univariate Descriptive Statistics	79
Measures of Central Tendency	79
Mode	80
Median	85
Mean	90
Selecting the Most Appropriate Measure of Central Tendency	92
Conclusion	94
Key Terms	95
Summary of Equations	95
Exercises	96
References	100
Notes	101

5	Measures of Dispersion	102
	Deviation and Dispersion	103
	Measures of Dispersion	105
	Range	105
	Index of Dispersion	108
	Mean Absolute Deviation	110
	Variance	111
	Standard Deviation	116
	Uses for the Variance and Standard Deviation	117
	Selecting the Most Appropriate Measure of Dispersion	117
	Conclusion	117
	Key Terms	118
	Summary of Equations	118
	Exercises	118
	References	123
	For Further Reading	123
	Note	123
6	The Form of a Distribution	124
	Moments of a Distribution	125
	Number of Modes	125
	Skewness	126
	Analysis of Skew	127
	Kurtosis	129
	Analysis of Kurtosis	129
	The Importance of Skew and Kurtosis	129
	Design of the Normal Curve	130
	Points to Remember About the Normal Curve	135
	Conclusion	136
	Key Terms	136
	Summary of Equations	136
	Exercises	136
	References	141
	For Further Reading	141
	Note	141
7	Introduction to Bivariate Descriptive Statistics	142
	Bivariate Tables and Analysis	143
	Statistical Tables versus Presentation Tables	145
	Constructing Bivariate Tables	147
	Ordinal Level Table Construction	148
	Nominal Level Table Construction	152
	Analysis of Bivariate Tables	152
	Conclusion	153
	Key Terms	153
	Exercises	153
	Notes	155

8	Measures of Existence and Statistical Significance	156
	Nominal Level Measures of Existence	157
	Tables, Percentages, and Differences	158
	Chi-Square	162
	Requirements for Using Chi-Square	170
	Limitations of Chi-Square	172
	Final Note on Chi-Square	173
	Tests of Existence for Ordinal and Interval Level Data	173
	Calculation and Interpretation for Ordinal Data	174
	Spearman's Rho and Pearson's r	174
	An Issue of Significance	179
	Conclusion	179
	Key Terms	180
	Summary of Equations	180
	Exercises	180
	References	188
	For Further Reading	188
	Notes	188
9	Measures of Strength of a Relationship	190
	What Is Association?	191
	Nominal Level Data	195
	Ordinal Level Data	199
	Tau	204
	Gamma	212
	Somers' d	214
	Spearman's Rho	216
	Interval Level Data	220
	Pearson's r	221
	Conclusion: Selecting the Most Appropriate Measure of Strength	228
	Key Terms	229
	Summary of Equations	229
	Exercises	230
	References	236
	Note	237
10	Measures of Direction and Nature of a Relationship	239
	Direction of the Association	239
	Establishing Direction for Ordinal Level Data	239
	Establishing Direction for Interval and Ratio Level Data	242
	Nature of the Association	244
	Establishing the Nature of the Distribution for Nominal and Ordinal Level Data	244
	Establishing the Nature of the Distribution for Interval and Ratio Level Data	247
	Conclusions	248
	Key Terms	249
	Exercises	249

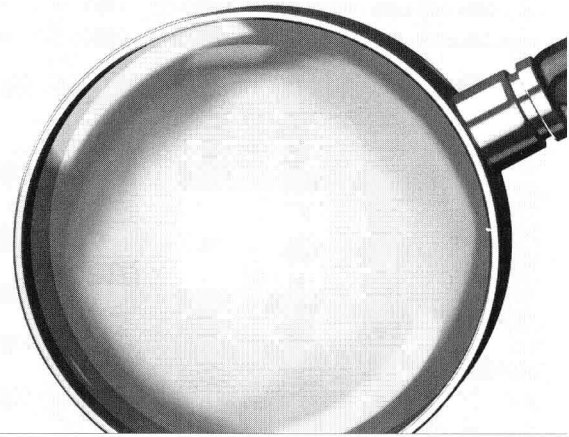
1 1	Introduction to Multivariate Statistics	254
	When Two Variables Just Aren't Enough	255
	Interaction Among Variables	255
	Causation	258
	Association	258
	Temporal Ordering	259
	Elimination of Confounding Variables	261
	Additional Concepts in Multivariate Analysis	262
	Robustness	262
	Error	263
	Parsimony	264
	Conclusion	265
	Key Terms	265
	Summary of Equations	265
	Exercises	265
	References	266
	Note	267
1 2	Multiple Regression I: Ordinary Least Squares Regression	269
	Regression	269
	Assumptions	271
	Analysis and Interpretation	274
	Steps in OLS Regression Analysis	278
	Other OLS Regression Information	283
	Limitations of OLS Regression	283
	Independent Variables with Lower Levels of Measurement and Nonlinear Relationships	283
	Dummy Variables	284
	Interaction Terms	285
	Nonlinear Relationships and Transformations	287
	Parabolic Functions	287
	Logarithmic Functions	291
	Multicollinearity	292
	Assessing Multicollinearity	293
	Adjusting for Multicollinearity	295
	Conclusion	295
	Key Terms	296
	Key Formulas	296
	Exercises	297
	References	298
	For Further Reading	299
	Notes	299
1 3	Multiple Regression II: Limited Dependent Variables	301
	Dealing with Limited Dependent Variables	301
	OLS Assumptions That Are Violated by Dichotomous Variables	302
	Logistic Regression	305
	Interpreting Logit Results	306
	Interactive Effects and Other Types of Logit	313
	Criticisms of Logistic Regression	315

Poisson and Negative Binomial Regression	316
A Note About Dispersion in Poisson and Negative Binomial Regression	317
Interpreting Poisson and Negative Binomial Regression	317
Criticisms of Poisson and Negative Binomial Models	320
Other Multiple Regression Techniques	320
Probit Regression	320
Tobit Regression	321
Multicollinearity and Alternative Regression Techniques	321
Conclusion	322
Key Terms	322
Exercises	322
References	323
14 Factor Analysis and Structural Equation Modeling	325
Introduction	325
Factor Analysis	325
Assumptions	327
Analysis and Interpretation	328
Structural Equation Modeling	341
Variables in Structural Equation Modeling	342
SEM Assumptions	342
Advantages of SEM	342
SEM Analysis	344
Conclusion	348
Key Terms	348
Key Equations	349
Questions and Exercises	349
References	349
15 Introduction to Inferential Analysis	352
Terminology and Assumptions	354
Normal Curve	355
Probability	357
Sampling	359
Probability Sampling	360
Nonprobability Sampling	363
Sampling Distributions	364
Central Limit Theorem	366
Confidence Intervals	367
Calculating Confidence Intervals	367
Interpreting Confidence Intervals	369
Conclusion	370
Key Words	370
Summary of Equations	371
Exercises	371
References	371

16 Hypothesis Testing	372
Null and Research Hypotheses	374
Steps in Hypothesis Testing	375
Type I and Type II Errors	380
Which Is Better, Type I or Type II Error?	382
Power of Tests	383
Conclusion	385
Key Terms	385
Summary of Equations	385
Exercises	385
References	386
For Further Reading	387
17 Hypothesis Tests	389
ZTest	389
Calculation and Example	390
Interpretation and Application: Known Probability of Error	392
One- versus Two-Sample Z Tests	396
t-test	396
Assumptions of a t-test	397
Calculation and Example	398
SPSS Analysis for Z tests and t-tests	400
Chi-square Test for Independence	406
Conclusion	407
Key Terms	407
Summary of Equations	407
Exercises	408
References	409
Note	409
18 Analysis of Variance (ANOVA)	411
ANOVA	411
Assumptions	412
Calculation and Interpretation	413
Post Hoc Tests	418
Conclusion	420
Key Terms	420
Summary of Equations	420
Exercises	420
References	421
For Further Reading	421
Notes	421
19 Putting It All Together	423
The Relationship Between Statistics, Methodology, and Theory	423
Describe It or Make Inferences	424
Abuses of Statistics	426

When You Are On Your Own	427
Conclusion	428
References	429
Notes	429
Appendix A	Math Review and Practice Test 431
Appendix B	Statistical Tables 435
Appendix C	The Greek Alphabet 441
Appendix D	Variables in Data Sets 443
Index	477

Use of SPSS and Data Sets with This Book



Although there are some formulas that can be worked by hand, this book is centered around determining the proper statistical analysis procedure to use with particular data and how to interpret the analyses. Throughout the book, SPSS output is displayed showing analyses. If you purchased a book bundled with SPSS, you may also follow the procedures in the book to conduct your own analyses. Sample data sets associated with the book are available at the Jones and Bartlett website associated with the book. The version of SPSS bundled with this book (or a copy of SPSS Grad Pack purchased separately) can be used to conduct most of the analyses in this book. A few of the multivariate analyses (Chapters 12, 13, and 14) require either a complete copy of SPSS or another program altogether, like SAS or Stata (for poisson/negative binomial regression).

THEORY

Observation
↓
Primary Question
↓
Research Questions

METHODS

Null Hypothesis
↓
Research Hypotheses
↓
Research Design

Concepts

Operationalization

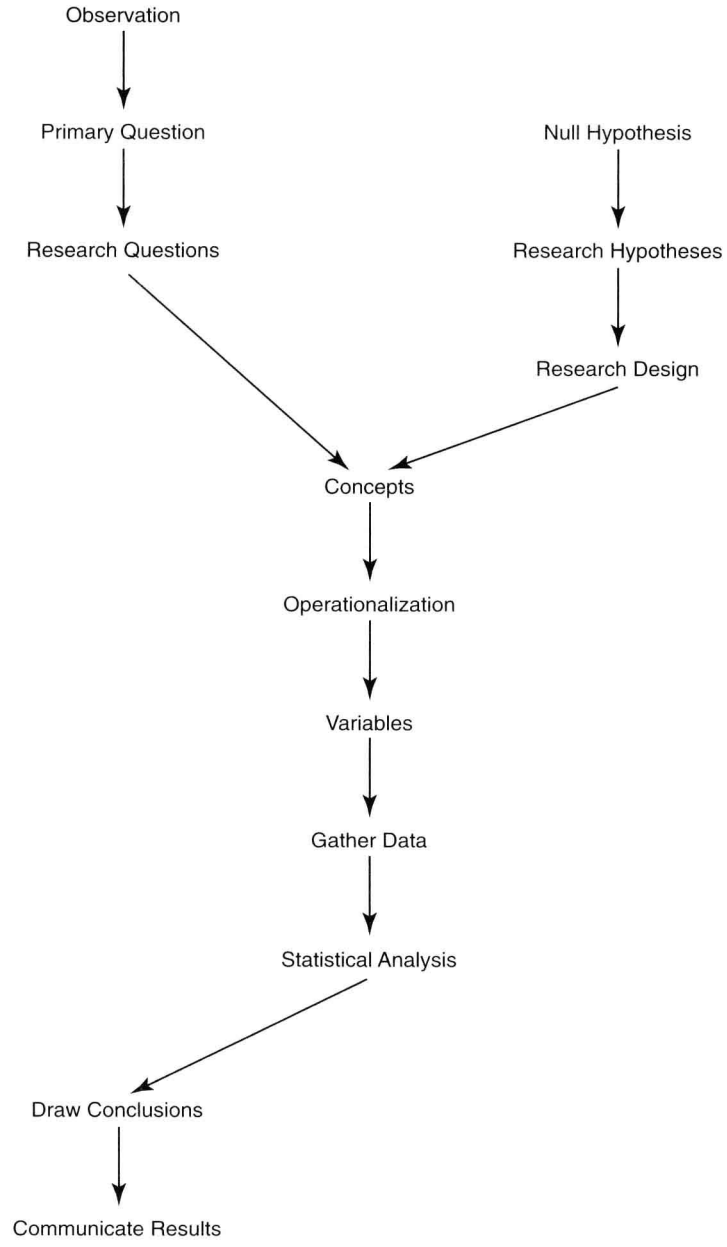
Variables

Gather Data

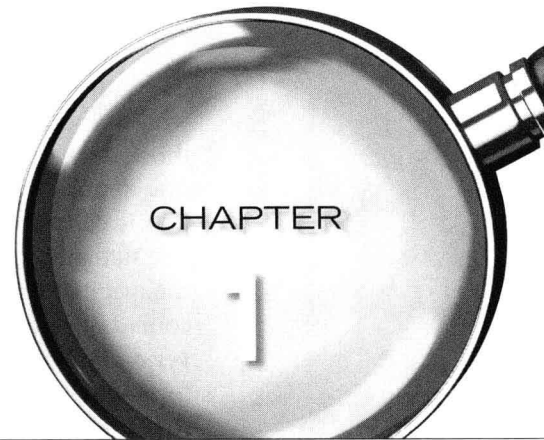
Statistical Analysis

Draw Conclusions

Communicate Results



The Logic of Comparisons and Analysis



Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.

—H. G. Wells

■ Introduction: Why Analyze Data?

Discovery and innovation may be the distinguishing characteristics between modern human activity and that of our ancestors. The Renaissance period brought forth an emphasis on learning and advancing our way of doing things that has prevailed to the present. Scientists, inventors, and others involved in the process of scientific inquiry have often been held in awe for their works. Galileo, Einstein, Madam Curie, and others are singled out in grade school books for their works and discoveries. As you will learn, so too should be people such as Pearson, Kendall, and Yule.

Statistical analysis is all about **discovery**. The process of **scientific inquiry** provides a method of examining things that interest us in a systematic manner. This process generally requires evidence to support an argument. One of the clearest methods of establishing evidence is by examining numbers associated with the objects being studied. That examination takes place through statistical analysis. As such, statistical analysis is the linchpin of discovery, and mastery of it draws us closer to Einstein and Galileo.

■ Some Statistical History

The earliest form of what is now considered **statistical analysis** was developed by Pythagoras in the 6th century BC. This was the forerunner of **descriptive statistics** (what would eventually be known as the *mean*, or what is commonly known as an *average*). The other type of statistical analysis (**inferential statistics**) is thought to have first developed in the Orient around 200 BC (Dudycha and Dudycha, 1972). This was a form of probability analysis used in assessing whether an expected child was likely to be male or female. Probability theory, as it would come to be known, continued in the form of gambling mathematics in the works of Blaise Pascal (1623–1662) and Lord Christianus Huygens (1629–1695) (David, 1962). Many of the other descriptive statistics were developed in the late 1800s and early 1900s by mathematicians and scientists such as Galton (1883) and Pearson (1895).

Statistics moved beyond gambling and purely mathematical concepts through what was called *political arithmetic*, a term coined because of its close association with those studying political topics, including economics. (This probably began the close association between political lying and statistical lying.) The first known use of this political arithmetic was by John Graunt (1662), who used what is now called descriptive statistics to study London's death rates. Although there is fierce debate concerning the original use of the term **statistics** (Yule, 1905), the greatest support is that it was coined by Eberhard August Wilhelm von Zimmerman in the preface of *A Political Survey of the Present State of Europe* (1787). Modern use of the term *statistics* (as opposed to *mathematics*) is often attributed to R. A. Fisher and his work *Statistical Methods for Research Workers* (1925), wherein he stated that "a statistic is a value calculated from an observed sample with a view to characterizing the population from which it is drawn." Since that time, statisticians have added to the techniques available to analyze data, many adding their names to the procedures; and the addition of statistical techniques continues today. Analysis procedures have been added to the statistical repertoire in the past few years that have greatly increased the ability of researchers in criminology, criminal justice, and other fields to examine the relationship between variables more accurately.

A single death is a tragedy, a million deaths is a statistic.

—Joseph Stalin

■ Uses of Statistics

The term *statistics* is often misunderstood because there are actually two practical applications of it. The first, reflecting the history of the term, is a collection of data—often expressed in summary form—that is collected and preserved. The best example of these are census statistics or mortality statistics, which depict the characteristics of the living or the causes of death, respectively. The second application is the subject of this book: a method of analyzing data. Statistics as you will come to know them are methods used to examine data collected in the process of scientific inquiry. These methods allow researchers to think logically about the data and to do one of two things: to come to some succinct and meaningful conclusions about the data (*descriptive statistics*), or to determine—or infer—characteristics of large groups based on the data collected on small parts (samples) of the group (*inferential statistics*). For example, data could be gathered on all correctional officers in Arkansas for a research project to determine the sex and race breakdown of the officers. This would be a descriptive analysis that could be used to examine the employment patterns for the Arkansas Department of Corrections. Alternatively, a sample of correctional officers from each state could be collected and the data from the sample used to make statements about all correctional officers in the nation. This would be drawing conclusions (inferences) about a large group based on information about a sample of the group.

Statistical analysis is the workhorse of discovery and knowledge. The scientific process, using research to test theory, requires that empirical evidence (data) drawn from the research subjects be examined systematically. The use of mathematics in general and statistical analysis in particular allows researchers to make these comparisons and to discover new information that will provide a better understanding of their subject.

In the scientific process, the purpose is usually to discover something that was previously unknown or to prove something true or false that was previously thought to be true but was never supported by hard evidence. The way to obtain that evidence is by gathering information (data) and subjecting it to statistical analysis.

■ Theory Construction at a Glance

Three elements in social science research, or any research for that matter, are essential to sound investigation: theory, research methods, and statistical analysis. Although these elements are intimately linked, there is debate—even among those most supportive of the research process—on their ordering, importance, and what should be included from each element in a textbook. It is not possible to cover all of these elements adequately in one course or in one textbook, so it becomes an issue of how much of each element should be included in a discussion of the other. In this book, theory is covered primarily in this chapter, research spans this chapter and several that follow, and statistical analysis prevails thereafter.

What Is Theory?

At the most basic level, **theory** consists of statements concerning the relationship or association among *social phenomena* such as events and characteristics of people or things. For example, in criminology, there are theories addressing how people learn to be criminal. In these theories, statements are constructed dealing with the role of peers in a person's learning criminal behavior, how the rewards from a crime can influence behavior, and what influence punishment can have on the decision to commit a crime.

The goal of these statements is to develop explanations of why things are as they appear and to try to explain their meaning. From an early age, humans have ideas about the causes of events and why things work the way that they do. The problem with these explanations, however, is that they are often too simplistic to be of any real value. Theory attempts to provide a stronger foundation for these ideas by asking questions about them, such as:

- What is the point of all of this?
- What does it mean?
- Why are things this way?

Without theory, there is often only conjecture and war stories. With theory, we may begin to develop statements or ideas that are based on sound observation and thought.

Theory and Research

Theory may be developed in several ways. Researchers may look at the world around them, find the **social phenomena** that pique their interest, and begin to develop statements concerning why these phenomena work the way they do. This is called **induction**. An example could be a researcher who follows crime trends in a city for a number of years. She may begin to see that the crimes follow a definite pattern of movement in the city, moving from east to west across the city. From this, she might set out to determine what the cause of this movement could be, ultimately developing a theory of crime movement in urban areas. This is a process of moving from data to theory and attempting to make sense of the data with the theory.

Alternatively, researchers may become curious about something and set out to develop statements and then to test them. This is called **deduction**. The process of deduction begins with an idea and an attempt to test the idea with data and analysis. For example, a researcher might believe that increased supervision of probationers would prevent them from becoming involved in subsequent crimes. This researcher might create an experiment where a random sample of probationers are put under intensive supervision while another random sample receives a normal amount of supervision. The results of this experiment could either support or refute the researcher's initial beliefs. This is a process of moving from theory to data, where the data tests the theory. It should be noted that Sherlock Holmes was not exactly correct in his understanding of the difference between induction and deduction. When Holmes made his famous statement, "brilliant deduction, Watson!", he should actually have been commending Watson on his inductive reasoning. Watson was drawing conclusions based on what he had observed, not testing previously developed conclusions, as discussed later.

Finally, and probably most often the case, a researcher may start with either induction or deduction, but by the time a project is finished, he or she has used both induction and deduction. This is called **retroduction**. With this process, the researcher investigating supervision of probationers might conduct the intensive supervision experiment as a deductive process. After examining the data, however, it might be obvious that the experiment could be done better or that there was something in the data that needed further explanation. For example, those probationers who received the most supervision were successful, whereas those who received intensive, but less than the most intensive, supervision were not successful. The researcher might then rethink part of the theory and set out to retest it. This process might continue until the theory was supported or disproven. This is a process of moving from theory to data to theory and so on; or data to theory to data and so on. The key here is that it is an alternating process between induction and deduction.

■ The Process of Scientific Inquiry

The process of scientific inquiry (using a deductive method) is shown in **Figure 1-1**. As shown in this diagram, theory is at the starting point of the process. Theory is driven by observations and leads researchers to initiate the research process through primary questions and research questions. It is from this process of theory building that researchers follow the process from developing a null hypothesis to communicating results. The process of scientific inquiry and its individual parts are discussed further in the remainder of the chapter.

Observation and Inquisitiveness

The first step in the process of scientific inquiry, and one of the most important, is often overlooked: **observation** and *inquisitiveness*. Many research projects are never begun because the researcher was not aware of his or her surroundings or did not recognize something as a topic worthy of research.

It is often theory that stimulates observation and scientific inquiry. As you go through school and read research and material you find interesting, you will sometimes think that you have a better way to do something, or what you read may stimulate you in other areas. By using a structured scientific process to evaluate your observations and formulate statements of why these phenomena are behaving the way they are, you are developing theory.