STATISTICS IN CRIMINAL JUSTICE

ANALYSIS AND INTERPRETATION

Jeffery T. Walker

disk included

AN ASPEN PUBLICATION

STATISTICS IN CRIMINAL JUSTICE

Analysis and Interpretation

Jeffery T. Walker, PhD
Professor
Department of Criminal Justice
University of Arkansas, Little Rock
Little Rock, Arkansas

AN ASPEN PUBLICATION®
Aspen Publishers, Inc.
Gaithersburg, Maryland
1999

Library of Congress Cataloging-in-Publication Data

Walker, Jeffery T.
Statistics in criminal justice: analysis and interpretation /
Jeffery T. Walker.
p. cm
Includes bibliographical references and index.
ISBN 0-8342-1086-X (alk. paper)
1. Criminal justice, Administation of—Statistical methods.
2. Criminal statistics—Research. I. Title.
HV7415.W33 1999
364'.01'5195—dc21
99-29104
CIP

Orders: (800) 638-8437
Customer Service: (800) 234-1660

Editorial Services: Nora Fitzpatrick
Library of Congress Catalog Card Number: 99-29104
ISBN: 0-8342-1086-X

Printed in the United States of America

1 2 3 4 5

This book is dedicated to
Orval and Ann Walker.

Table of Contents

PART V—INFERENTIAL STATISTICS

CHAPTER 1

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 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 like 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 gathering 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 sixth century B.C. This was the forerunner of descriptive statistics (what would eventually be known as the mean, or 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 B.C. (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 economic. (This probably began the close association between political lying and statistical lying.) The first known use of this political arithmetic was by John Graunt, who used what is now called descriptive statistics to study London's death rates (1662). Although there is fierce debate concerning the original use of the term *statistics* (Yule, 1905), the greatest support argues that it was coined by Eberhard August Wilhelm von Zimmerman in the preface of *A Political Survey of the Present State of Europe* (1787). The 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 criminology, criminal justice, and other researchers to more accurately examine the relationship between variables.

> *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 text: a method of analyzing data. Statistics as you will come to know them are methods used to examine the 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 Nebraska 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 Nebraska Department of Correction. 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 advancement. 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 hard 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. These are 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 any particular textbook. It is not possible to adequately cover all of these elements 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 text, theory will be primarily covered in this chapter, research will span this chapter and several that follow, and statistical analysis will prevail 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 that address how people learn to be criminal. In these theories, statements are constructed that deal with the role of peers in a person's learning criminal behavior, how the rewards from a crime can influence behavior, and what effect 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 that they do. This is called *induction*. An example might 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. 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 have actually been commending Watson on his inductive reasoning. Watson was actually drawing conclusions based on what he had observed, not testing previously developed conclusions, as will be discussed below.

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 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.

Finally, and probably most often the case, a researcher may start with either induction or deduction, but, by the time a project is finished, the researcher 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; while those who received intensive, but less than the most intensive, supervision were not successful. The researcher, then, might 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 (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 re-

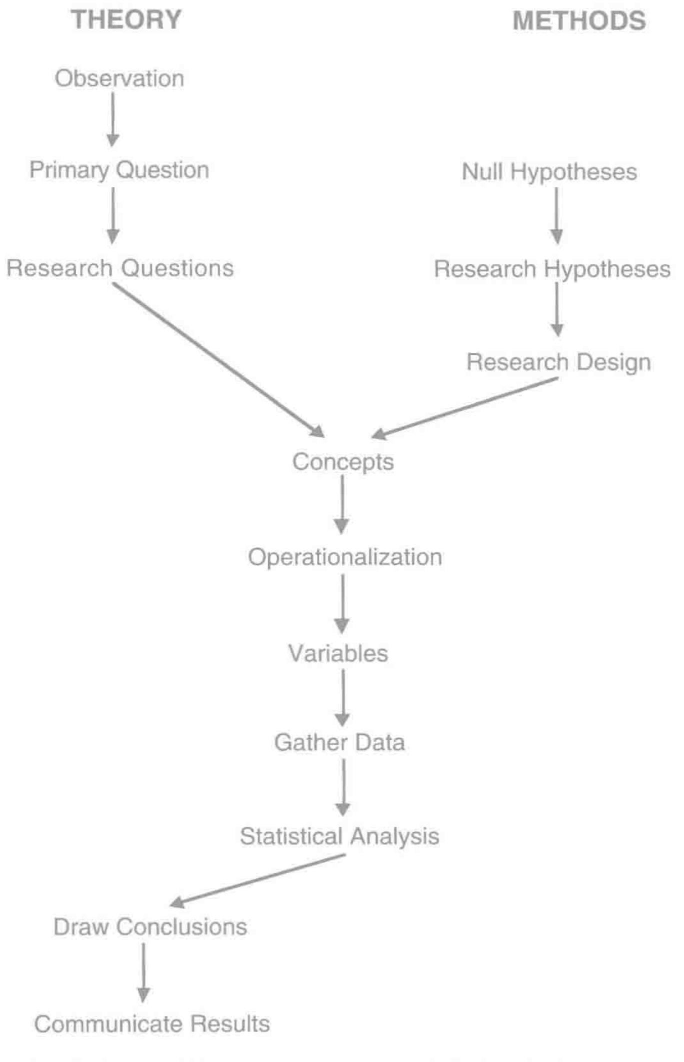

Figure 1–1 Process of scientific inquiry: Theory, research, and statistical analysis.

searchers 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 further discussed in the remainder of this chapter.

Observation and Inquisitiveness

The first step in the process of scientific inquiry, and one of the more 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 that you find interesting, sometimes you will think 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 that they are, you are developing theory.

An example of inductive theory development can be shown in Robert Burgess' Zonal hypothesis. Students at the University of Chicago were making maps of Chicago showing different characteristics of neighborhoods, such as welfare, infant mortality, and housing. Burgess observed that these maps followed very similar patterns throughout the city. His observations led him to develop a theory about how cities grow and change. The theory he developed from these maps proposed that cities grow in rings similar to when a rock is thrown into the water. In this configuration, the rings closest to the center of the city will be the most run down, have the highest level of infant mortality, and be characterized by other social ills that are not present in the outer rings. All of this was developed simply by examining the students' maps developed and by using inductive theory building.

Primary Questions

A primary question is the one, driving thought behind a research project. It should represent the whole reason for the study. Primary questions are important because how well a researcher meets the goals of the primary question will often be the criteria by which the research will be evaluated. The primary question should be a carefully worded, specific phrase that states exactly the focus of the study.

For example, in research on police use of deadly force, a possible primary question might be, What factors most influence police use of deadly force?[1] This question is very broad, and somewhat vague, but it can easily represent the goal of a research project.

Research Questions

Often the primary question will be theoretical, vague, and quite possibly not directly addressable through research. Research questions break down the primary question into subproblems that are more manageable, and make the primary question testable through research. If the primary question establishes the goal of the research, the research questions suggest ways of achieving that goal.

In the example concerning police use of deadly force used above, some possible research questions might include the following:

- What is the relationship between an officer's shift and the likelihood of the officer using deadly force?
- What is the relationship between the violent crime rate of an area and police use of deadly force?

---

[1] This example will be used throughout this discussion to show the research process using a single example in each stage.

- What is the relationship between an officer's level of education and the likelihood of the officer using deadly force?

These research questions break down the primary question into smaller parts that can be more easily examined. The answers to these questions are derived from the research process and statistical analysis and allow the researcher to answer the primary question.

RESEARCH: MOVEMENT FROM THEORY TO DATA AND BACK

Theory cannot stand alone. Nor can research or statistics. Theory without research and statistical analyses to back it up is little more than fable. Research without theory is like building a house without plans; and research without statistical analysis is like building a house without nails: it is possible and it is done, but it would be more effective with them. Statistical analysis without theory and research methodology to guide the research is like having your own intercontinental ballistic missile—nice to have and may impress your neighbors, but it is not really useful.

Whatever the technique of developing theory, research is the method used to test and validate it. Research, in its purest form is a scientific, systematic study to discover new information or to test the validity of theories previously developed. The prime goal of research is discovery. Depending on whether an inductive or deductive process is used, research is a systematic way of turning statistical observation and analysis into theory (induction) or testing theory through statistical analysis, through deduction.

Although there are no exact, cookbook steps that must be followed in conducting a research project, there are some general guidelines that should be followed, either consciously or unconsciously, to ensure that nothing is left out of the study. These steps are included in this chapter and provide a kind of map of where research and statistics fits into the overall process of scientific inquiry.

Formulating Hypotheses

Once the research questions have been developed, you must decide what the research is attempting to determine. Hypotheses are questions or statements whose answers will support or refute the theoretical propositions of the research. Hypotheses are generally broken down into *research hypotheses* and *null hypotheses*. Although these will be covered in more detail in chapter 13, a brief definition is provided here.

A research hypothesis is generally a statement, similar to a research question, that states the expected outcome of a part of a research project. If a research question in a project asked "What is the relationship between an officer's shift and the likelihood of using deadly force?" the research hypothesis might be that there is a statistically significant correlation between an officer's shift and the likelihood of using deadly force. In using research hypotheses, the researcher turns the relatively abstract wording of theory development into a more concrete, testable form appropriate for statistical analysis.

One of the often difficult to understand but vital elements of statistical analysis and hypothesis testing is that research alone cannot prove anything. Even when researchers find a great deal of support for an association between two variables, it may be that these results are occurring because information is missing or the model is somehow flawed. Other researchers may very well be able to disprove these findings by conducting additional research. If research cannot prove anything, then what can it do? It can be used to disprove something or eliminate alternatives. For example, even though research cannot prove that officers on night shifts use deadly force more than their daytime counterparts, it may disprove that there is no relationship between shift and use of deadly force. This is accomplished with a null hypothesis. A null hypothesis generally states one of the following:

- There is no statistically significant difference between the groups being compared.