Eighth Edition

Elementary Statistics in Social Research

Jack Levin • James Alan Fox

EIGHTH EDITION

Elementary Statistics in Social Research

Jack Levin

James Alan Fox

Northeastern University

Allyn and Bacon

Boston ■ London ■ Toronto ■ Sydney ■ Tokyo ■ Singapore

Series Editor: *Sarah L. Kelbaugh*
Editor in Chief, Social Sciences: *Karen Hanson*
Editorial Assistant: *Jennifer DiDomenico*
Marketing Manager: *Brooke Stoner*
Editorial Production Service: *Chestnut Hill Enterprises, Inc.*
Manufacturing Buyer: *Julie McNeill*
Cover Administrator: *Linda Knowles*

A Pearson Education Company
160 Gould Street
Needham Heights, MA 02494

Internet: www.abacon.com

Between the time web site information is gathered and published, some sites may have closed. Also, the transcription of URLs can result in typographical errors. The publisher would appreciate notification where these occur so that they may be corrected in subsequent editions.

Library of Congress Cataloging-in-Publication Data

Levin, Jack, 1941–
Elementary statistics in social research / Jack Levin, James Alan Fox. — 8th ed.
p. cm.
Includes index.
ISBN 0-321-04460-6
1. Social sciences—Statistical methods. 2. Statistics. I. Fox, James Alan. II. Title.
HA29.L388 2000
519.5'024'3—dc21 99-22898
CIP

Printed in the United States of America

10 9 8 7 6 5 4 3 2 RRD-VA 04 03 02 01 00

PREFACE

The eighth edition of *Elementary Statistics in Social Research* provides an introduction to statistics for students in sociology and related fields, including criminal justice, political science, and social work. This book is not a comprehensive reference work on statistical methods. On the contrary, our first and foremost objective is to be understandable to a broad range of students, particularly to those who may not have a strong background in mathematics.

Like its predecessors, the eighth edition contains a number of pedagogical features. Most notably, detailed step-by-step illustrations of statistical procedures continue to be located at important points throughout the text. We have again provided clear and logical explanations for the rationale and use of statistical methods in social research. And we have again included a large number of end-of-chapter problems, almost all of which are answered at the end of the book. In addition, we have added discussions of applying the weighted mean and analyzing archival data.

Students sometimes get lost in the "trees" of statistics, without seeing the "forest." For the first time in this edition, therefore, we have ended each part of the text with a new section entitled "Looking at the Larger Picture," which carries the student through the entire research process based on a hypothetical survey of smoking and drinking among high school students.

Following a detailed overview in Chapter 1, the text is divided into five parts. Part One (Chapters 2 to 4) introduces the student to the most common methods for describing and comparing data. Part Two (Chapters 5 and 6) serves a transitional purpose. Beginning with a discussion of the basic concepts of probability, it leads the student from the topic of the normal curve as an important descriptive device to the use of the normal curve as a basis for generalizing from samples to populations. Continuing with this decision-making focus, Part Three (Chapters 7 to 9) contains several well-known tests of significance. Part Four (Chapters 10 to 12) includes procedures for obtaining correlation coefficients, and an introduction to regression analysis. Finally, Part Five consists of an important chapter in which students learn, through example, the conditions for applying statistical procedures to research problems.

The text provides students with background material for the study of statistics. A review of basic mathematics, statistical tables, a list of formulas, and a glossary of terms are located in appendixes at the end of the book. For this edition, we have also included, in Appendix A, a list of Internet websites in which important social science data can be found.

Three supplements are available with this edition. First, there is a revised student workbook. Also, a revised and expanded Instructor's Manual, which includes a test bank. The test bank is also available in ESATest, in DOS, Windows and Mac formats.

We are grateful to the following reviewers of the eighth edition for their insightful and helpful suggestions: Robin Franck, Southwestern College; Gary Hampe, University of Wyoming; William Kelly, University of Texas at Austin; Paul Murray, Siena College; and Ira Wasserman, Eastern Michigan University.

We are also indebted to the Literary Executive of the late Sir Ron A. Fisher, F.R.S.; to Dr. Rank Yates, F.R.S.; and to Longman Group, Ltd., London, for permission to reprint Tables III, IV, V, and VI from their book *Statistical Tables for Biological, Agricultural and Medical Research,* 6th ed., 1974. Finally, we acknowledge the important role of our personal computers, without "whose" assistance this revision would not have been possible.

Jack Levin
James Alan Fox

CONTENTS

CHAPTER

1 Why the Social Researcher Uses Statistics

A little of the social scientist can be found in all of us. Almost daily, we take educated guesses concerning the future events in our lives in order to plan for new situations or experiences. As these situations occur, we are sometimes able to confirm or support our ideas; other times, however, we are not so lucky and must face the sometimes unpleasant consequences.

To take some familiar examples: We might invest in the stock market, vote for a political candidate who promises to solve domestic problems, play the horses, take medicine to reduce the discomfort of a cold, throw dice in a gambling casino, try to psych out our instructors regarding a midterm, or accept a blind date on the word of a friend.

Sometimes we win; sometimes we lose. Thus, we might make a sound investment in the stock market, but be sorry about our voting decision; win money at the craps table, but discover we have taken the wrong medicine for our illness; do well on a midterm, but have a miserable blind date; and so on. It is unfortunately true that not all of our everyday predictions will be supported by experience.

The Nature of Social Research

Similar to our everyday approach to the world, social scientists attempt to explain and predict human behavior. They also take "educated guesses" about the nature of social reality, although in a far more precise and structured manner. In the process, social scientists examine characteristics of human behavior called *variables*—characteristics that differ or vary from one individual to another (for example, age, social class, and attitude) or from one point in time to another (for example, unemployment, crime rate, and population).

Not all human characteristics vary. It is a fact of life, for example, that the gender of the person who gave birth to you is female. Therefore, in any group of individuals, gender of mother is the *constant* "female." A biology text would spend considerable time discussing why only females give birth and the conditions under which birth is possible, but a social scientist would consider the mother's gender a given, one that is not worthy of study because it never varies. It could not be used to explain differences in the mental health of children because all of their mothers are females. In contrast, mother's age, race, and mental health are variables: In any group of individuals, they will differ from person to person and can be the key to a greater understanding of the development of the child. A researcher therefore might study differences in the mental health of children depending on the age, race, and mental health of their mothers.

In addition to specifying variables, the social researcher must also determine the *unit of observation* for the research. Usually, social scientists collect data on individual persons. For example, a researcher might conduct interviews to determine if the elderly are victimized by crime more often than younger respondents. In this case, an individual respondent is the unit to be observed by the social scientist.

However, researchers sometimes focus their research on *aggregates*—that is, on the way in which measures vary across entire collections of people. For example, a researcher might study the relationship between the average age of the population and the crime rate in various metropolitan areas. In this study, the units of observation are metropolitan areas rather than individuals.

Whether focusing on individuals or aggregates, the ideas that social scientists have concerning the nature of social reality are called *hypotheses.* These hypotheses are frequently expressed in a statement of the relationship between two or more variables: at minimum, an *independent variable* (or presumed cause) and a *dependent variable* (or presumed effect). For example, a researcher might hypothesize that socially isolated children watch more television than children who are well-integrated into their peer groups, and he or she might conduct a survey in which both socially isolated and well-integrated children are asked questions regarding the time they spend watching television (social isolation would be the independent variable; TV-viewing behavior would be the dependent variable). Or a researcher might hypothesize that the one-parent family structure generates greater delinquency than the two-parent family structure and might proceed to interview samples of delinquents and nondelinquents to determine whether one or both parents were present in their family backgrounds (family structure would be the independent variable; delinquency would be the dependent variable).

Thus, not unlike their counterparts in the physical sciences, social researchers often conduct research to increase their understanding of the problems and issues in their field. Social research takes many forms and can be used to investigate a wide range of problems. Among the most useful research methods employed by social researchers for testing their hypotheses are the experiment, the survey, content analysis, participant observation, and secondary analysis. For example, a researcher may conduct an experiment to determine if arresting a wife batterer will deter this behavior in the future, a sample survey to investigate political opinions, a content analysis of values in youth magazines, a participant observation of an extremist political group, or a secondary analysis of government statis-

tics on unemployment. Each of these research strategies is described and illustrated in this chapter.

The Experiment

Unlike everyday observation (or, for that matter, any other research approach), the *experiment* is distinguished by the degree of *control* a researcher is able to apply to the research situation. In an experiment, researchers actually manipulate one or more of the independent variables to which their subjects are exposed. The manipulation occurs when an experimenter assigns the independent variable to one group of people (called an *experimental group*) but withholds it from another group of people (called a *control group*). Ideally, all other initial differences between the experimental and control groups are eliminated by assigning subjects on a random basis to the experimental and control conditions.

For example, a researcher who hypothesizes that frustration increases aggression might assign a number of subjects to the experimental and control groups at random by flipping a coin ("heads" you're in the experimental group; "tails" you're in the control group), so that the groups do not differ initially in any major way. The researcher might then manipulate frustration (the independent variable) by asking the members of the experimental group to solve a difficult (frustrating) puzzle, whereas the members of the control group are asked to solve a much easier (nonfrustrating) version of the same puzzle. After all subjects have been given a period of time to complete their puzzle, the researcher might obtain a measure of aggression by asking them to administer "a mild electrical shock" to another subject (actually, the other subject is a confederate of the researcher who never really gets shocked, but the subjects presumably do not know this). If the willingness of subjects to administer an electrical shock is greater in the experimental group than in the control group, this difference would be attributed to the effect of the independent variable, frustration. The conclusion would be that frustration does indeed tend to increase aggressive behavior.

In 1995, three University of Wisconsin researchers—Joanne Cantor, Kristen Harrison, and Marina Krcmar—conducted an experiment to study the effect of Motion Picture Association of America ratings (G, PG, PG-13, R) on children's decisions to watch a particular movie. The researchers manipulated the independent variable, motion picture ratings, by asking a sample of boys—aged 10–14—to select the one film they wished to watch from a list of three movies that had previously been judged to be equally appealing. In all cases, two of the movies were rated "PG"; only the rating of the third film, *The Moon-Spinners,* was varied. On a random basis, one-quarter of the boys were told it had been rated "G," one-quarter "PG," one-quarter "PG-13," and one-quarter "R." If the film's rating had no effect on their preferences, about 33% of the boys should have chosen to watch *The Moon-Spinners,* regardless of the rating it had been assigned.

The results showed something else: When the film *The Moon-Spinners* was rated "G," none of the boys chose it. When the film got a "PG" rating, 38.9% of the boys selected it over the other two films. But when *The Moon-Spinners* was rated "PG-13" or "R," at least 50% wanted to see it rather than the other motion pictures on the list. Apparently, at least for boys aged 10–14, declaring a film "off limits" or "restricted" makes the movie more appealing. Conversely, rating a film "G" only serves to reduce its popularity.

One final result deserves to be mentioned. Using an identical experimental procedure, the researchers determined that a sample of girls in the same age group were apparently not effected by the media equivalent of "forbidden fruit." Only 11% of the girls selected the "R" rated version of *The Moon-Spinners;* more than 29% chose the film when it had a "G" rating.

The Survey

As we have seen, experimenters actually have a direct hand in creating the effect that they seek to achieve. By contrast, *survey* research is *retrospective*—the effects of independent variables on dependent variables are *recorded* after—and sometimes long after—they have occurred. Survey researchers typically seek to reconstruct these influences and consequences by means of verbal reports from their respondents in self-administered questionnaires, face-to-face interviews, or telephone interviews.

Surveys lack the tight controls of experiments: Variables are not manipulated and subjects are not assigned to groups at random. As a consequence, it is much more difficult to establish cause and effect. Suppose, for instance, in a survey measuring fear of crime, that a researcher finds that respondents who had been victims of crime tend to be more fearful of walking alone in their neighborhoods than those who had not been victimized. Because the variable *victimization* was not manipulated, we cannot make the logical conclusion that victimization *causes* increased fear. An alternative explanation that the condition of their neighborhoods (poverty, for example) produces both fear among residents and crime in the streets is just as plausible.

Surveys also have advantages precisely because they do not involve an experimental manipulation. As compared with experiments, survey research can investigate a much larger number of important independent variables in relation to any dependent variable. Because they are not confined to a laboratory setting in which an independent variable can be manipulated, surveys can also be more *representative*—their results can be generalized to a broader range of people.

In September 1998, for example, a *New York Times* telephone survey of a national sample of U.S. women discovered differences by age in opinions about Monica Lewinsky—the White House intern whose alleged affair with President Clinton led to his impeachment. According to these results, younger women were somewhat more tolerant of Ms. Lewinsky than were their mothers or grandmothers. For example, when asked whether they approved or disapproved of the White House intern, nearly 75% of women over 65 but only slightly more than half of women under 30 said that they "disapproved" of her. Even differences by age could not, however, conceal an overall conclusion of the study: The majority of American women may not have denounced Monica Lewinsky, but they also did not hold her in high regard.

By contrast, a Gallup survey of 1,055 adults conducted at the end of 1998, just after the President had been impeached by the House of Representatives, determined that Clinton was the man most admired by Americans in 1998—far ahead of the second place winner, Pope John Paul II. In fact, according to these survey results, Clinton's rating was four