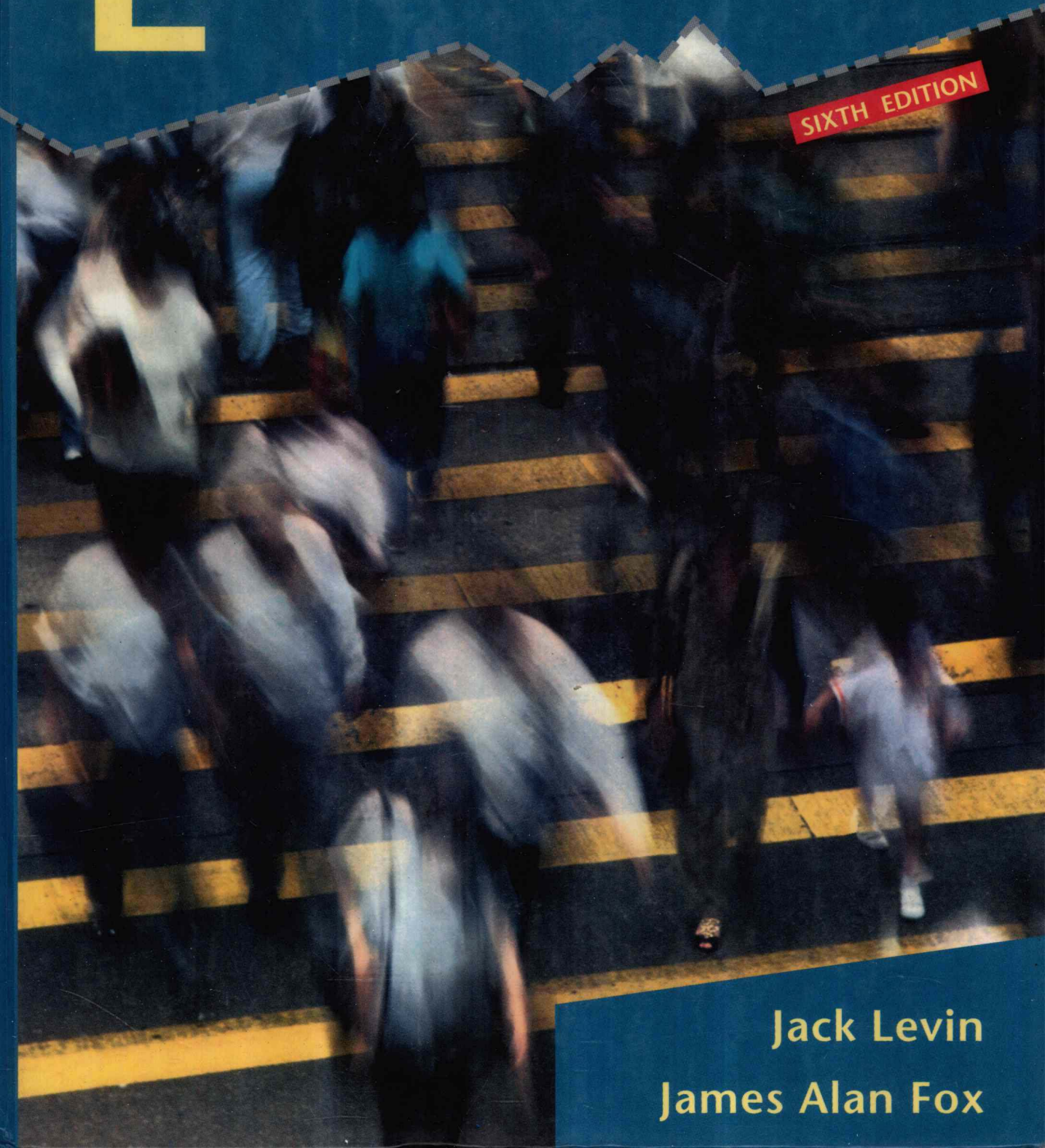
Elementary Statistics

in Social Research



SIXTH EDITION

Jack Levin

James Alan Fox

Elementary Statistics
in Social Research

SIXTH EDITION

Jack Levin
Northeastern University

James Alan Fox
Northeastern University

HarperCollins*CollegePublishers*

Acquisitions Editor: Alan McClare
Project Coordination, Text and Cover Design: Proof Positive/Farrowlyne Associates, Inc.
Cover Photo: Chuck O'Rear/Westlight
Production Manager: Kewal Sharma
Compositor: Black Dot Graphics
Printer and Binder: R. R. Donnelley & Sons Company
Cover Printer: The Lehigh Press, Inc.

Elementary Statistics in Social Research, Sixth Edition

Library of Congress Cataloging-in-Publication Data

Levin, Jack, 1941–
Elementary statistics in social research / Jack Levin, James Alan Fox.—6th ed.
p. cm.
Includes index.
ISBN 0-673-46958-1
1. Social sciences—Statistical methods. 2. Statistics. I. Fox, James Alan. II. Title.
HA29.L388 1993
519.5'0243—dc20

93-21091
CIP

95 96 9 8 7 6 5 4

We dedicate the sixth edition to our families:

Flea, Michael, Bonnie, and Andrea
Sue Ann, David, Jennifer, and Alexander

PREFACE

The sixth edition of *Elementary Statistics in Social Research* is designed to provide an introduction to statistics for students in sociology and related fields, including social work, political science, and criminal justice. This book does *not* claim to be 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 students who may not have a strong background in mathematics.

The book contains a number of pedagogical features. Most notably, detailed step-by-step illustrations of statistical procedures have been located at important points throughout the text. At the same time, we have attempted to provide clear and logical explanations for the rationale and use of statistical methods in social research. Additional pedagogical aids include computer output illustrating various statistical applications as well as numerous end-of-chapter problems, which are answered at the end of the book.

Following a detailed overview in Chapter 1, the text is divided into three parts. Part One (Chapters 2–4) 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 7–13) contains several well-known tests of significance, procedures for obtaining correlation coefficients, and an introduction to regression analysis. Part Three concludes with an important chapter in which students learn, through example, the conditions for applying statistical procedures to research problems.

A number of appendixes provides students with background material. Appendix A contains a campus profile data set on 100 colleges and universities, which is used in all the computer illustrations. Copies of the data set (with an SPSS control file) are available on diskette from the authors. Also included is an overview of computers as well as an introduction to SPSS. Additional appendixes contain a review of basic mathematics, statistical tables, a list of formulas, and a glossary of terms.

In this revision, we have enhanced the text in two major ways. First, we have clarified discussions of some of the concepts that students find most difficult to understand, including sampling distributions and tests of significance. Second, we have rewritten all end-of-chapter problem sets, placing the exercises in research contexts.

Instructors have told us that, for the most part, they like the coverage of the book as it is. As in the saying "if it's not broken, don't fix it," we have made only a few changes in the list of topics and their organization. Most notably, we have added one-way chi-square, largely as a useful pedagogical bridge to the chi-square test of independence, and percentile ranks for the normal curve. Additionally, we have incorporated the presentation of graphs into the frequency distribution chapter, and have reorganized the formulas for central tendency and variability in a more logical sequence.

Three new supplements are available with this edition. First, the student workbook has been replaced by a new Student Study Guide written by Maria I. Lasaga and Robert K. LeDuc. Also, an instructor's manual provides a test bank and handouts for classroom use. Finally, for those who choose to use computer-aided instruction, a tutorial program called *StatTutor*, containing step-by-step training, animation, exercises, and a statistics spreadsheet, is available in MS-DOS and Mac formats through the publisher.

We are grateful to Roger Jarjoura of the University of Indiana at Indianapolis for his careful reading of the manuscript and to Diana Roberts for her assistance. Special thanks are due to the following reviewers for their insightful comments and helpful suggestions: Harry Azuma, Southeastern Missouri State University; Mary Ann Coughlin, Springfield College; Kevin Fitzpatrick, The University of Alabama at Birmingham; Stephen Kulis, Arizona State University; Pam Rosenberg, Gettysburg College; and Barbara A. Zsembik, University of Florida.

We are also indebted to the Literary Executive of the late Sir Ron A. Fisher, F.R.S.; to Dr. Frank 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* (Sixth Edition, 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

— — — *Part Three Decision Making*

— — — *Appendixes*

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 sex of the person who gave birth to you is female. Therefore, in any group of individuals, sex 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 sex 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 might, therefore, 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, and participant observation. 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, or a participant observation of an extremist political group. 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, while 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.

The Survey

As we have seen, the experimenter actually has a direct hand in creating the effect that he or she seeks 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 is 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.

For example, a survey researcher who hypothesizes that frustration increases aggression might locate a number of severely aggressive individuals and interview them to identify the frustrating events in their lives, such as isolation, physical disabilities, poor grades in school, and poverty. Obviously, survey researchers cannot manipulate the variables by introducing these frustrating life events themselves; but they can attempt to discover and record them after they have occurred. To study the relationship between frustration and aggression, Stuart Palmer interviewed the mothers of 51 convicted murderers. He found many more frustrating circumstances in the early lives of these killers than in the lives of their control brothers who had not committed murder. Specifically,

the research showed that the murderers as children had been subjected to more serious illnesses, operations, accidents, beatings, physical defects, frightening experiences, and disapproval from their peers than had the control group.

Content Analysis

As an alternative to experiments and surveys, *content analysis* is a research method whereby a researcher seeks objectively to describe the content of previously produced messages. Researchers who conduct a content analysis have no need directly to observe behavior or to question a sample of respondents. Instead, they typically study the content of books, magazines, newspapers, films, radio broadcasts, photographs, cartoons, letters, verbal dyadic interaction, political propaganda, or music. In 1988, for example, Jack Levin, Arnold Arluke, and Amita Mody-Desbareau content analyzed celebrity and noncelebrity profiles printed in the four most widely circulated gossip tabloids: *The National Enquirer, The Star, The National Examiner,* and *The Globe.* Using appropriate coding sheets, they coded each of the 311 profiles published in a sample of issues during a six-month period from February through July 1983.

The researchers found that most of the celebrities they studied were featured because of some minor or mundane event (for example, a quarrel between spouses or a shopping spree). By contrast, the profiles of noncelebrities emphasized some extraordinary act of charity, strength, or heroism (for example, an 83-year-old woman who gives up her social security check to feed the homeless). The hidden message was laden with social control implications: Ordinary people ought to be content with their collective place in life. After all, they are capable of extraordinary accomplishments. In the tabloids, everyday life is exciting, even for the "little" people of the world.

Participant Observation

Another widely used research method is *participant observation,* whereby a researcher "participates in the daily life of the people under study, either openly in the role of researcher or covertly in some disguised role, observing things that happen, listening to what is said, and questioning people, over some length of time."[1]

The particular strength of participant observation is its ability to provide a *complete* form of information about a situation or a series of events. The

[1]Howard S. Becker and Blanche Geer, "Participant Observation and Interviewing: A Comparison," in William J. Filstead (ed.), *Qualitative Methodology* (Chicago: Markham, 1970), 133.

participant observer is able to determine the meaning of social situations, not from the viewpoint of outsiders, but as defined by the group members themselves.

An example of participant observation is found in John Dollard's classic study of the patterns of race relations in a small southern town. For a five-month period during 1936, Dollard operated in the role of researcher to observe what the townspeople said, did, and seemed to feel, without directly interrogating members of the community with predetermined questions. Based on his participant observation, Dollard discovered, among other things, that white, middle-class residents of this southern town gained some important economic, sexual, and prestige benefits by maintaining the traditional patterns of race relations between blacks and whites. He also found that race prejudice and discrimination generated forms of aggression among blacks, against whites, and against blacks.

— — — Why Test Hypotheses?

Social science is often referred to, quite unfairly, as the study of the obvious. However, it is desirable, if not necessary, to test hypotheses about the nature of social reality, even those that seem logical and self-evident. Our everyday commonsense observations are generally based on narrow, often biased preconceptions and personal experiences. These can lead us to accept without criticism invalid assumptions about the characteristics of social phenomena and behavior.

To demonstrate how we can be so easily misled by our preconceptions and stereotypes, consider what we "know" about mass murderers. How many of the following characteristics seem obvious to you? How many would not be worth studying because they are so obvious?

1. Mass murderers are almost always insane. (After all, a sane person would never shoot 14 people from the top of a tower as Charles Whitman did in 1966.)
2. Mass murderers are usually loners. (The neighbors always seem to say "He was quiet, he stayed pretty much to himself, we never really knew him.")
3. Mass murderers look different from the rest of us. (One generally imagines a glassy-eyed lunatic, as in the horror movies.)
4. Mass murderers are usually strangers to their victims, who are unlucky enough to be in the wrong place at the wrong time. (Cases like the McDonald's mass murder in a San Diego suburb in 1984 typically come to mind.)

That these conceptions about mass murderers seem so clear-cut and indisputable might explain why it took until 1985 for anyone to look systematically at the profile of the mass murderer. Compiling detailed information about 42