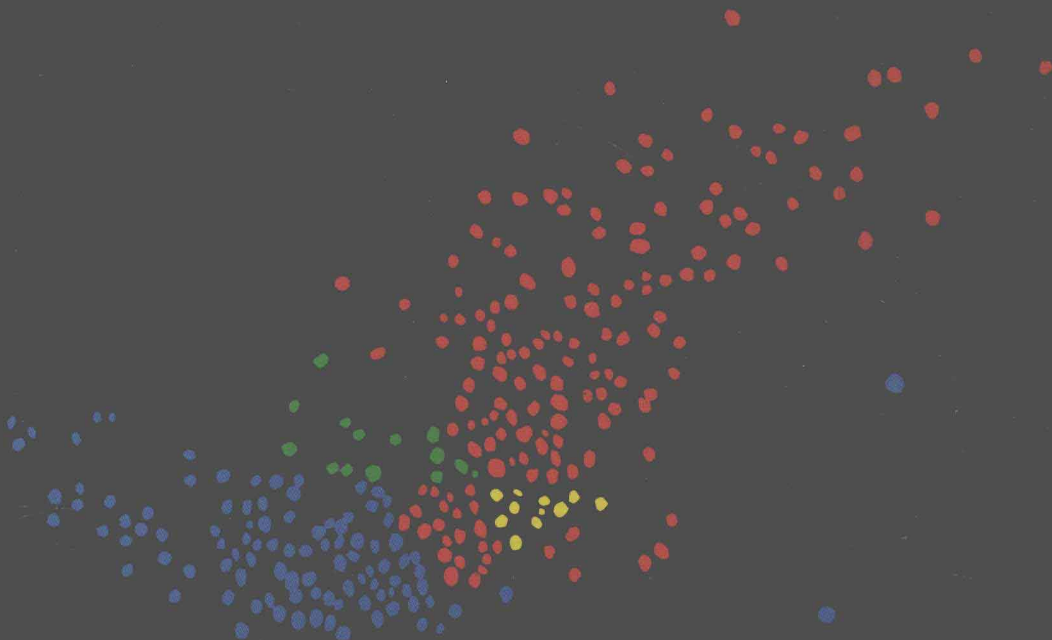


Think and Explain with Statistics

Lincoln E. Moses

Think and Explain with Statistics

Lincoln E. Moses, Stanford University

Addison-Wesley Publishing Company
Reading, Massachusetts • Menlo Park, California
Don Mills, Ontario • Wokingham, England • Amsterdam
Sydney • Singapore • Tokyo • Mexico City • Bogotá
Santiago • San Juan

Sponsoring Editor: Jeffrey Pepper
Illustrations: Capricorn Design
Production Supervisor: Marion E. Howe
Copy Editor: Helen Greenberg
Text Designer: Herbert Caswell
Manufacturing Supervisor: Ann Delacey
Cover Design: Mike Fender

Library of Congress Cataloging in Publication Data

Moses, Lincoln E.
Think and explain with statistics.

Includes index.
1. Statistics. I. Title.
QA276.12.M668 1986 519.5 85-6097
ISBN 0-201-15619-9

ABCDEFGHIJ-HA-89876

Preface

Understanding is the taproot of success in thinking and explaining with statistics. Thus the central aim of this book is to provide the reader with the basis for understanding the material in it. This aim accounts for many of the features of the book.

Examples are numerous and come from a variety of fields. They are typically used first to introduce a topic, and often a second time to illustrate the use of a technique.

Notation is kept as simple as possible in order not to hinder the less mathematical reader.

Seldom does the progress of the reasoning unfold by a sequence of algebraic steps.

Intuitive arguments are consistently supplied (and they are typically true in the sense that they underlie the rigorous omitted mathematical justification).

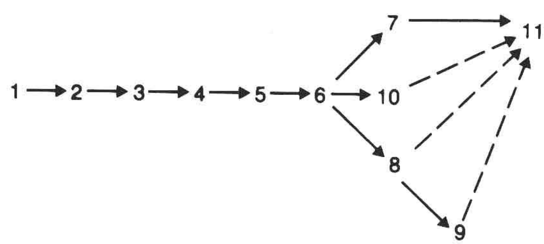
Key principles are often presented from several different points of view (e.g., the Mann–Whitney and Wilcoxon tests both appear — together, and with their equivalence laid bare). Alternative views deepen understanding.

Thinking is a private activity. Explaining is a public one. Statistical methods can be very useful in both processes. There are many informal techniques that can greatly aid a person in thinking about the data; such techniques have an important place here. (Spearman's rank correlation is a familiar example; the poor man's standard error is an unfamiliar one.) The more public activity, explaining, calls upon descriptive techniques; they, too, have a place here, perhaps more than is usual.

"Seeing is believing." The idea of random variation, especially in statistics computed from samples, is best acquired by seeing sampling experiments done. The book develops the central limit theorem on the basis of sampling experiments and turns to them again at other points.

Each chapter begins with an introduction and ends with (1) a summary of statements that recapitulate the content of the chapter and (2) a list of all the new key terms in the chapter. A reader who may have forgotten an earlier statistics course will find these end-of-chapter lists useful; if each statement in a chapter's summary has clear meaning, then that chapter does not need further review.

The logical dependencies among the chapters are as diagrammed. (Dashed lines indicate that certain portions of Chapter 11 are separately accessible using the indicated predecessor.) The diagram shows that there are several choices available for carrying the course beyond the first six chapters. The book has been used in general undergraduate courses and in courses for medical students, education majors, psychology majors, and biology majors. The text uses examples from all these fields and others.



To the Student

Expect the material in this book, after careful reading, *to make sense*. If it does not, you are probably misunderstanding something, because by and large, statistical methods provide reasonable extensions of common sense. Do the homework, and do not fall behind; you will expend less total effort and gain more total understanding by following this advice. Check the end-of-chapter summaries and key terms thoroughly to test your understanding. (This also provides a good way to prepare for examinations.) Be aware that one desirable outcome of studying statistics is for you to develop the ability to read graphs and tables more easily than you can read verbal descriptions of them; to this end, you should never shirk study of the graphs and tables in this book.

To the Teacher

Each chapter section has exercises at the end, some of which may be useful as class material. Any of them should be useful for homework. Often the students benefit more from doing such an exercise if it is talked over in advance, with an eye to answering the question, "Why is this problem useful to an understanding of the material?"

Sampling experiments in class can rivet the attention of the students, who literally watch random variation occur. Each student uses random numbers, draws a sample, computes a statistic, and reports its value when called on by the instructor, who adds that value to the tabulation at the blackboard. Advance preparation by the instructor is very worthwhile. It makes the lessons go more quickly and effectively. The following questions deserve advance thought. How can one be assured that different students are using different parts of the random number table? (Possible answer: Issue each student a personal computer-generated page.) How many observations will be necessary to illustrate the point at issue? Thus, how many observations are needed from each student? How much time will that take? What is the range of data points that will have to be tabulated? What scale should be used for graphing? What intervals are needed for tabulation? If each student reports two numbers, are they to be tabulated in two arrays, or one, or in x, y coordinates? What statistical summary of the data can be made effectively during class at the board? Will two colors of chalk help? Are the computations to be done by each student easy enough to be both quickly and accurately performed—by nearly every student? Are there useful checks you can cause students to apply? Most of these questions are easy enough to answer *before* class. Then the rewards of a well-conducted sampling experiment can be great. The students' statistical intuition is strengthened, their specific understanding of symbols and technical terms is hardened, and they vicariously experience statistical summarization of data that they are interested in.

The Appendix contains a data set of 300 observations, each with 18 different variables measured on each subject. Some variables are bodily dimensions like height, span, and leg length; some are ratios of such measurements, like forearm length divided by height; two are performance measures (spirometer and grip strength). These data may be useful in assigning exercises, in doing sampling experiments, and in illustrating descriptive and analytic techniques.

Acknowledgments

This book has been long in the making, and many people have helped it come to fruition. Again and again as I worked, I was aware of my great intellectual debt to John Tukey for much of the book's content and for much of my outlook. Fred Mosteller's generous, insightful, and numerous suggestions for change have improved its substance and teachability; I am very grateful. Mark Appelbaum did great numbers of computer simulations that now form a key part of the early chapters. Merrill Carlsmith generously read and commented often over the years. The book is better for this help, but I claim full credit for its weaknesses.

I have benefited from the comments of many who have read and others who have taught from the earlier versions. I thank them here: Peter Bacchetti, Stuart Bessler, Byron Brown, Debbie Burr, Lee Cronbach, Reese Cutler, Bradley Efron, David Freedman, Victor Fuchs, Arthur Goldberger, Gloria Guth, Joe Harrison, Allan Oaten, Jan Pederson, Martin Rein, David Rogosa, Lee Ross, Jack Schneider, Charles Sherman, and Elaine Ung.

Without the willing and accurate typing of Wanda Edminster, the book might never have been completed. And without the typing of Betsy Shopler, it might never have been started. Erin Nakashima deftly effected the final changes.

Doug Danforth, Marguerite Frick, Jerry Halpern, Carol Treanor, and Neil Willits provided computer assistance at various times.

I am greatly indebted to the Psychometric Laboratory of the University of North Carolina and its director, Lyle Jones, and to the Center for Advanced Study in the Behavioral Sciences for fellowships and other support that were essential to the work.

Finally, the transition from a virtually finished manuscript to a publisher-ready one might never have been accomplished without the willing and effective work of James N. Miller, whom I thank here most warmly.

Stanford, California
St. Swithin's Day, 1985

Lincoln E. Moses

Contents

CHAPTER One

Welcome to Statistics

1.A Two Meanings of the Word

Statistics (plural) calls to mind graphs, charts, numbers, tables of numbers, and the like. There is a second correct use of the word. *Statistics* (singular) denotes a body of methods for reaching conclusions from data.

So, in the plural, statistics means facts and figures, but in the singular it indicates a discipline that is partly a science and partly an art. Both meanings of the word are worth examining.

Statistics (Plural)

Everyday life makes the first usage familiar. Public opinion polls report the percentage of people who approve of the way the President is doing his job. The Bureau of Labor Statistics publishes changes in the Consumer Price Index. The registrar of a university announces that last year's undergraduate enrollment stood at 18,540. Each of these statistics is very public in its character; each exemplifies the idea that "any organization that seeks to understand itself and its surroundings . . . finds its thinking pervaded with statistical questions."[1]

Statistics in the sense of facts and figures also occur in other ways. Price determinations of bulk commodities may depend upon the statistical results of samples drawn from the lots that are being purchased. When the Surgeon General of the United States "determined that cigarette smoking may be harmful to your health," he did so on the basis of statistical studies showing that heavy cigarette smokers had a 10-fold increase in the risk of lung cancer

1. *Federal Statistics*, President's Commission on Federal Statistics, Washington, D.C., 1971, Vol. I, p. 17.

("10-fold"—another statistic). Scientists announce that their laboratory has bounced radar signals off the sun (or a laser beam off the moon) and have received the bounced-back signal; this announcement is based on a statistical analysis of squiggles in some curves.

Statistics as facts and figures occur in many ways.

Statistics (Singular)

The methods for reaching conclusions from data, which comprise statistics in the singular, involve collecting, organizing, interpreting, and reporting. The usual trappings of a learned discipline are all present: books, journals, professional societies, research on open questions, university departments. Research proceeds steadily on improving and extending the methods that are the substance of the field.

There are some clearly recognizable subfields of statistics. Econometrics, psychometrics, biometrics (or biostatistics), engineering statistics, and educational statistics are all important examples. Each of these has one or more professional organizations and at least one journal.

Despite specialization there is a large common core of statistical methodology that is best presented without closely relating it to one particular field of application. This book addresses that common core.

The Connection Between the Two Concepts of Statistics

On the wall above my desk is the following quotation:

> "The Government are very keen on amassing statistics. They collect them, add them, raise them to the *n*th power, take the cube root and prepare wonderful diagrams. But you must never forget that every one of those figures comes in the first instance from the village watchman, who just puts down what he damn pleases."
>
> (Sir Josiah Stamp)

The message of this vivid statement is:

> A statistic simply reports the outcome of some process. The *name* attached to the statistic, be it "the cost of living" or "the probability of an accident in a nuclear power plant," may or may not fairly describe what it purports to describe. The quality and credibility of a statistic, indeed its very meaning, depend upon the process that produced it.

The following example nicely illustrates how the meaning of a statistic depends on the process that produced it.

Example 1 Purse snatching in 1973 and 1974
Based on the Federal Bureau of Investigation's (FBI's) *Uniform Crime Reports,*[2] one would estimate that purse snatching in Pittsburgh increased by 23 percent between 1973 and 1974. But one would conclude from the Law Enforcement Assistance Administration's (LEAA's) National Crime Panel victimization survey[3] that instead there was an 18 percent *drop* between the two years. Two different processes here result in statistics on purse snatching: the first is based on routine records of crimes reported to the police; the second is based on responses obtained in a sample survey of a cross section of the American people about crimes they have experienced, whether or not the police were informed.

This example also shows that the connection between statistics as numbers and statistics as methods is intimate and inescapable.

1.B Some Examples of Uses of Statistics

Statistical methods are applied to enormously varied problems in a wide range of contexts. To introduce the field, it is helpful to look briefly at some applications.

1. *Estimating the value of standing timber in a forest tract.* Both the purchaser and the seller of standing timber are interested in the value of the trees. (So is the tax assessor.) How many trees are there? What are their sizes? If it is a mixed forest, these questions may need to be answered for several species. The estimation will be done by means of *samples,* which is essentially a statistical process.
2. *Developing high-yield, disease-resistant crops.* Agriculture experiment stations have developed better varieties of corn, wheat, turkeys, dairy cattle, and so on. The process of research and development depends on organizing and interpreting statistical data. In fact, the close connection between agricultural research and statistics has greatly enriched statistical theory, as well as serving the development of better varieties of crops and livestock and better agricultural practices.
3. *Measuring attitudes, aptitudes, and academic achievement.* Tests like the Scholastic Aptitude Test and the Medical College Admissions Test could

2. *Uniform Crime Reports: Crime in the U.S.* (annual report by the FBI), September 6, 1974, and November 17, 1975.
3. *Sourcebook of Criminal Justice Statistics.* (annual report by the LEAA), July 1976, pp. 414–424.

not have been developed, nor can they be wisely interpreted, without extensive recourse to statistical ideas.

4. *Measuring the cost of living*. The Consumer Price Index, which affects many wage scales under the provisions of labor contracts, is based on enormous numbers of facts continuously being collected. What kinds of purchases are typical? In what amounts? At what prices? These questions are answered on the basis of sample data and combined in carefully designed ways to provide this much-used statistic.
5. *Ascertaining health effects of air pollution*. This task requires the solution of several different problems, including (a) measuring health variables in populations, (b) measuring air pollution levels to which those populations are exposed, and (c) eliminating (after measuring them) the effects on health of other variables like weather and the economic well-being of the population. Each of these measurement problems is intrinsically statistical.
6. *Determining the effectiveness of cloud seeding*. Clouds are individual, and, when seeded with silver iodide, some will produce rain, while others will not. The same is true for clouds that are not seeded. Measuring the effect of cloud seeding requires us to take account of that variation. This demands the use of statistics not only in analyzing the data but, more importantly, in planning the studies from which the data come.
7. *Resolving a question of disputed authorship*. Several of the Federalist Papers, unsigned, have long been known to have been written by either Alexander Hamilton or James Madison, but the choice between them lay in doubt for decades. Sometimes such a problem can be settled by noting that one author writes long sentences and the other short ones. But in Federalist Papers of known authorship, Madison's sentences averaged 34.59 words, while Hamilton's averaged 34.55. More sophisticated analysis was necessary. In an important book, original statistical methods were developed and applied successfully to resolve the problem.

What causes statistics to be so prominently involved in so many different kinds of substantive questions? The answer lies in the following fact.

> If you try to solve a substantive problem by acquiring information about it from many people, or many places, or many times, or many experimental animals, then what began as a problem about timber or reading ability or air pollution will necessarily also become a statistical problem.

This fact carries important implications. First, it is wise to learn how to deal well with statistical issues; they are pervasive in problem solving. Second, it is wise to think *at the beginning* of a study about the statistical problems it will present before it is completed; this thinking may lead to improving the plan for doing the study.

1.C A Brief Sketch of the History of Statistics

Astronomy was studied by the ancients. Geometry was highly developed by the Greeks in pre-Christian times. By these standards statistics is a young discipline, with its main roots reaching back only to the middle of the seventeenth century. (It is true that population counts, censuses, were made in some places and times by Greeks, Romans, Persians, Egyptians, Hebrews, and ancient Japanese, but they were not responsible for the development of the field.

In 1662 an Englishman, John Graunt, published an important work entitled "Natural and Political Observations Made Upon the Bills of Mortality." By studying the information contained in death notices in London, he attempted to answer questions like "What is the population of London?" This was an early and important example of ***the analysis of social data,*** one of the two main streams of origin of the modern field of statistics. Only five years earlier, in 1657, a Dutch physicist and mathematician, Christiaan Huygens, had published the first printed work on games of chance—an early landmark in the development of ***probability theory,*** the second main stream of origin.

The analysis of social data grew steadily. A close associate of John Graunt, Sir William Petty, is called the father of econometrics. He was the author (in 1665) of the first known national income estimates. At about this time, censuses began to appear in more and more countries. For example, the English Board of Trade ordered 27 censuses in the North American colonies between 1635 and 1776. The Constitution of the United States required a decennial census; the first was in 1790. The U.S. Census is the oldest continuing periodic census.

The mathematical stream was also growing. Many of the great mathematicians worked on the theory of probability (Pascal, Fermat, James Bernoulli, DeMoivre, Laplace, Gauss). Some theoretical contributions were addressed to gambling, others to the theory of errors (of astronomical observations).

By the early nineteenth century, there was a considerable body of theory relevant to statistics and growing practical activity in the field. What we have described as two streams of origin were not entirely separate, of course. Graunt had made a serious effort to construct a life table; such tables are the basis of life insurance. His result was incorrect; but about 30 years later, in 1693, the astronomer Edmund Halley (for whom the famous comet is named) produced the first correct life table. Another typical contact between the two streams occurred in 1828, when a general statistics handbook for Belgium appeared; it was produced by Adolphe Quetelet—physicist, astronomer, and mathematician.

The nineteeenth century saw enormous growth in many branches of science. Statistics participated in this growth. John Snow in 1854—a generation before the acceptance of the germ theory—demonstrated, from data, that 500 cholera deaths in London were the result of contaminated water at a certain pump. (This same physician introduced ether anesthesia into England.) Between 1837 and 1880, William Farr, in the British General Register Office, developed the field of vital statistics, and its present structure still shows his hand.

The mathematics of evolution and heredity engaged the energies of Francis Galton and Karl Pearson in the late 1800's; in the early 1900's Sewell Wright's and R. A. Fisher's work in genetics and evolution further enriched statistics. Fisher's main contributions to statistics arose in the treatment of agricultural research, but he also made notable contributions arising from problems in other fields such as geology and botany.

Psychology as a modern discipline is about 100 years old. Its development has also been interwoven with the development of statistical theory. Karl Pearson and R. A. Fisher, but also C. E. Spearman, and much later Harold Hotelling and S. S. Wilks, found the source of much of their work in psychological inquiries.

World War II was a turning point after which the field of statistics burgeoned. Just before the war, it had received a vital infusion of seminal ideas from Jerzy Neyman and E. S. Pearson. During the war statistical methods were applied, with great success, on an unprecedentedly large scale to industrial processes, to planning of economic and social actions by the government, to experimentation in research and development work, and to military operations.

After the war, university departments of statistics began to appear in the United States. New journals were founded, and the size of the profession increased rapidly. The body of theory grew. Growth continues today, perhaps more rapidly than ever.

1.D Some Philosophical Observations About Statistics

Statistics and Physics

The physicist whose name is associated with supersonic speeds, Ernst Mach, said of his field, "Physics is experience arranged in economical order." To some extent, any science might be described in similar terms; the experiences would