Statistics for Experimenters

An Introduction to Design, Data Analysis, and Model Building

GEORGE E. P. BOX

WILLIAM G. HUNTER

J. STUART HUNTER

Statistics for Experimenters

An Introduction to Design, Data Analysis, and Model Building

GEORGE E. P. BOX

WILLIAM G. HUNTER

J. STUART HUNTER

John Wiley & Sons

New York • Chichester • Brisbane • Toronto

Library of Congress Cataloging in Publication Data

Box, George E. P.
Statistics for experimenters.

(Wiley series in probability and mathematical statistics)
Includes index.
1. Experimental design. 2. Analysis of variance. I. Hunter, William Gordon, 1937– joint author. II. Hunter, J. Stuart, 1923– joint author. III. Title.

QA279.B68 001.4'24 77-15087
ISBN 0-471-09315-7

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Preface

Collaboration with research workers has been a source of satisfaction to us as practicing statisticians. It has put us in touch with a wide variety of investigations and has stimulated us by intellectual contact with many investigators. We have certainly learned from them, and they have sometimes learned from us. This book owes much to such interaction. It is written for those who collect data and try to make sense of it, and it is an introduction to those ideas and techniques that we have found especially useful.

For a number of years we have used preliminary versions (in the form of notes) of this book in teaching engineers, chemists, biologists, statisticians, and other scientists at the University of Wisconsin and Princeton University. We have also used this material in courses for professional societies, industry, and government. It is neither a cookbook nor a textbook on mathematical statistics. It is an introduction to the philosophy of experimentation and the part that statistics plays in experimentation.

Actual problems are never so straightforward that they can safely be solved mechanically. Therefore we emphasize the necessity for thinking about the real nature of the scientific problem itself, for mulling over data plots and other graphical displays, as well as for understanding potentially useful statistical principles and their practical consequences. We point out possible difficulties caused by the violation of assumptions, such as the lack of independence of data. Furthermore we discuss ways by which such difficulties may be overcome. To illustrate principles we use both real and constructed data. For example, when the analysis of variance is first introduced, we simplify a real set of data so that all the values required in the analysis are whole numbers. Special emphasis is placed on the design of experiments because this is the most valuable aspect of statistical method. Frequently conclusions are easily drawn from a well-designed experiment, even when rather elementary methods of analysis are employed. Conversely, even the most sophisticated statistical analysis cannot salvage a badly designed experiment.

Statistical theory is introduced as it becomes necessary. Readers are assumed to have no previous knowledge of the subject; and although the book presents a properly sophisticated view of the philosophy of scientific investigation and the part played by statistical design and analysis, the mathematics needed is elementary. In particular, calculus does not appear in the main body of the text and is not needed to understand and use the book. Even more important than learning about statistical techniques is the development of what might be called a capability for *statistical thinking*. We hope this book will serve as a vehicle for this purpose. To help convey this facet of statistics, which is a challenge to the instructor, we have included several different types of questions and problems. In addition to exercises throughout the text, there are review questions at the end of each chapter, and problems at the end of each part of the book.

In addition to experimenters, others will find this book useful, for example, those who assess reported findings (Do the data really support these claims?), managers who direct research projects (What is the best way to approach this investigation?), and statisticians who work with experimenters (What is the best method of design and analysis for this problem?).

A typical one-semester college course might include fairly complete coverage of the first three parts of the book (Chapters 1–13) and a selection of topics from the fourth part (Chapters 14–18). Although this book is primarily intended for use as a text, we have also kept in mind the reader who will use it for self-instruction and/or as a reference.

Special Notes to the Reader

The questions at the end of each chapter reflect its contents and can be used in two ways: You can consider them for review *after* reading the chapter, or you can study them *before* reading the chapter to help identify key points. You may also find it helpful to use the problems listed at the end of each part of the book to guide your reading.

As you apply the ideas in this book, and especially if you meet with spectacular success or failure, we shall be interested to learn of your experience. We have tried to write a book that is useful and clear. If you have suggestions about how it can be improved, please write to us. As Chapter 1 suggests, we believe in iteration.

Acknowledgments

The manuscript was begun at Princeton University and continued at the University of Wisconsin, Madison, at both the Department of Statistics and the Mathematics Research Center. We record our thanks to Bovas Abraham, Steve Bailey, Gina Chen, Johannes Ledolter, Greta Ljung, and Lars Pallesen for helping with the calculations and for checking the final manuscript. We are also indebted to the hundreds of students who used the notes on which this book is based and their proposed improvements.

We are grateful to Professor E. S. Pearson and the Biometrika Trustees for permission to reprint condensed and adapted forms of various tables from *Biometrika Tables for Statisticians*, Vol. 1, 3rd ed., 1966. These are listed at the end of this book as Tables A, B1, C, D, F, and G.

We are grateful also to the Literary Executor of the late Sir Ronald A. Fisher, F.R.S., to Dr. Frank Yates, F.R.S., and to Longman Group Ltd., London, for permission to partially reprint Table III from their book *Statistical Tables for Biological, Agricultural and Medical Research*, 6th ed., 1974.

Our thanks are due to William G. Cochran and Gertrude M. Cox for allowing us to reproduce material on incomplete blocks designs from their classic book, *Experimental Designs*.

Finally, we are glad to acknowledge our debt to Mary Esser, Mary Arthur, and Doris A. Whitmore, who carefully typed the final manuscript.

G. E. P. Box
W. G. Hunter
J. S. Hunter

Madison, Wisconsin
Princeton, New Jersey
March 1978

Greek Alphabet

Α α	alpha	Ν ν	nu
Β β	beta	Ξ ξ	xi
Γ γ	gamma	Ο ο	omicron
Δ δ	delta	Π π	pi
Ε ϵ	epsilon	Ρ ρ	rho
Ζ ζ	zeta	Σ σ	sigma
Η η	eta	Τ τ	tau
Θ θ	theta	Υ υ	upsilon
Ι ι	iota	Φ ϕ	phi
Κ κ	kappa	Χ χ	chi
Λ λ	lambda	Ψ ψ	psi
Μ μ	mu	Ω ω	omega

Contents

PART II COMPARING MORE THAN TWO TREATMENTS

CHAPTER 1

Science and Statistics

1.1. THE LEARNING PROCESS

Scientific research is a process of guided learning. The object of statistical methods is to make that process as efficient as possible.

Learning is advanced by the iteration illustrated in Figure 1.1. An initial hypothesis leads by a process of *deduction* to certain necessary consequences that may be compared with data. When consequences and data fail to agree, the discrepancy can lead, by a process called *induction*, to modification of the hypothesis. A second cycle in the iteration is thus initiated. The consequences of the modified hypothesis are worked out and again compared with data (old or newly acquired*) that in turn can lead to further modification and gain of knowledge.

This process of learning can also be depicted as a feedback loop (Figure 1.2) in which the discrepancy between the data and the consequences of hypothesis H_1 leads to the modified hypothesis H_2, H_2 leads to H_3, and so on.

Nursery School Example

In a certain nursery school the teacher was pouring juice into cups for the children, who were sitting around a table. The teacher asked what would happen if, rather than stopping, she just kept pouring the juice into one of the cups. One child said the juice would fly *up* to the sky, and all the other children, except one, agreed. The one dissenting child (very likely drawing on her own past experience) said the juice would overflow and run *down* onto the table. The teacher took the opportunity to perform the experiment and demonstrate what actually happens. Learning took place that morning.

* In most of our applications the data acquiring process is scientific experimentation, but it could be a walk to the library to unearth already existing information or the conduct of a sample survey.

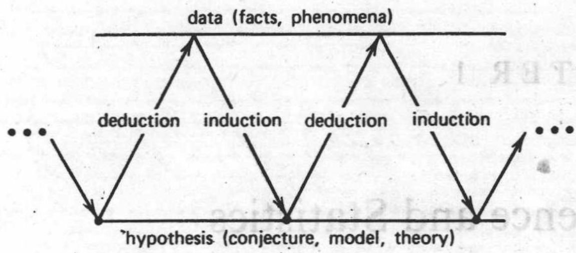

FIGURE 1.1. The iterative learning process.

Chemical Example

A chemist might have the following learning experience.

Hypothesis 1 and its deduced consequences	Because of certain properties of a newly discovered catalyst, its presence in a particular reaction mixture would probably cause chemical A to combine with chemical B to form, in high yield, a valuable product C.

The chemist has a tentative hypothesis and deduces its consequences, but he has no data on which to verify or deny its truth because, as far as he can tell from discussion with colleagues and careful examination of the literature, no one has ever performed the operation in question. He therefore decides to run some experiments.

Experimental design 1	A run is made at carefully selected reaction conditions. In particular, the chosen reaction temperature is 600°C.

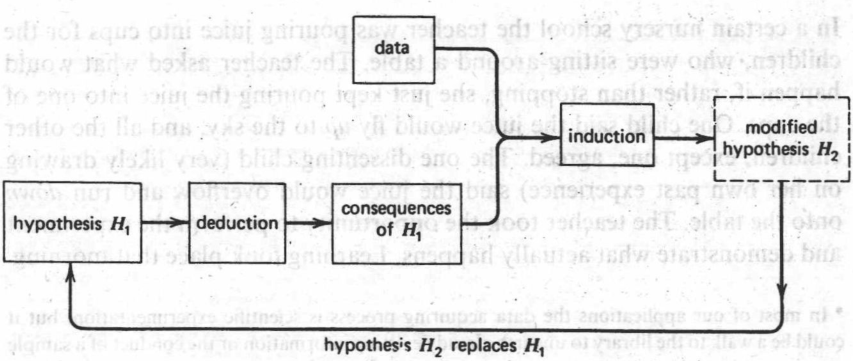

FIGURE 1.2. The learning process as a feedback loop.