

**Basic
Behavioral
Statistics**

Gehring

1
2
3
4
5
6
7
8
9

0213
G311

9061810

5-8

贈閱

Basic Behavioral Statistics

Robert E. Gehring
Indiana State University Evansville



The Foundation for People's Friendship

美国友好书刊基金会

赠阅书



E9061810

Houghton Mifflin Company Boston
Dallas Geneva, Illinois Hopewell, New Jersey
Palo Alto London

To the memory of my mother

Copyright © 1978 by Houghton Mifflin Company. All rights reserved.

No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or by any information storage or retrieval system, without permission in writing from the publisher.

Printed in the U.S.A.

Library of Congress Catalog Card Number: 77-078447

ISBN: 0-395-24684-9

Basic Behavioral Statistics

Preface

A New Twofold Approach to Introductory Statistics

Here is perhaps the first serious attempt to make introductory behavioral statistics as easy as possible to understand, yet at the same time challenging to more capable students. Compared with traditional approaches, which aim at either a relatively lower or a relatively higher level of student, this text deals with different motivational and ability levels by offering two books in one.

Every main topic gets double coverage. The initial treatment of each topic is an easy-to-understand presentation suitable for all students. Additional coverage is presented in the supplementary material at the end of each chapter. The supplementary material includes more detailed and specialized applications, plus mathematical relationships and derivations of formulas. This additional coverage offers extra precision, rigor, comprehensiveness, and thought-provoking stimulation for more ambitious students.

A Study Guide Is Available

A companion Study Guide for students gives complete step-by-step instructions for performing the most important computations. The Study Guide also contains many basic exercises in behavioral applications. It is strongly recommended that students test and perfect their learning by working homework problems in the text, and if they need additional help and practice, in the Study Guide also.

The Study Guide incorporates an additional feature. A number of tables and figures from the text are reproduced in the Study Guide. This enables students to keep each table or figure in front of them for repeated reference while they are reading on in the text and turning to new pages. They do not have to keep their fingers constantly in the text to mark the location of important tables and graphs.

Overcoming Mental Blocks and Substandard Math Backgrounds

I am particularly sensitive to students' problems in introductory statistics because I teach several courses based largely upon the introductory course. These include advanced parametric statistics, nonparametric sta-

tistics, and the type of experimental psychology laboratory course in which some half-dozen research projects must be statistically analyzed.

With this perspective, I am particularly concerned that introductory statistics students develop at least a certain practical mastery and general understanding of techniques. They should learn enough from the introductory course to function effectively at higher levels, whether or not they are “whizzes” in math.

In order to make the learning process as interesting and as easy as possible, I have employed the following principles:

Most examples are from data that students in the behavioral sciences have found most interesting or most relevant to their studies. This brings a bit of the excitement of current research into the discussion of less exciting methods.

In the first, or basic, portion of each chapter, there is only information that I believe every student should learn. This starts each student off with only essential and understandable information, devoid of unnecessary details, complications, and digressions. Optional or advanced information appears only in the supplementary material at the end of each chapter, in approximate order of usefulness. The following ordering of supplementary material is typical: (1) useful alternate formulas, (2) details and advanced remarks omitted from the basic section, (3) related topics, and (4) mathematical derivations.

Verbal and pictorial or graphic explanations of principles precede mathematical treatment and the introduction of statistical symbols. Thus, principles are initially explained in much the same understandable manner as in “content” courses in behavioral science.

Necessary mathematical passages are presented in a conversational, rather than in a forbiddingly formal, manner.

Distributions of data in examples and exercises often contain just a few numbers in order to make the repetitive element in calculations as simple as possible.

Tables and figures are carefully integrated with text explanations.

Various types of questions often asked by students—some relatively enlightened and some otherwise—are answered in the text (sometimes in parentheses). This adds an element of comprehensiveness and should minimize wasting valuable class time in answering questions that can be answered conveniently in the text.

In certain sections of all chapters, students are periodically asked by rhetorical questions to answer queries concerning the material being covered. Students are best advised to try to answer these ques-

tions to the best of their ability before proceeding with their reading. The questions provide spot-checks of progress and understanding. They are designed to remedy problems in understanding *as they arise*. The questions solidify understanding and help prevent more serious difficulties from occurring later, as students work on the extensive exercises at the end of both the basic and supplementary material in each chapter. While students should think out the answers to the questions themselves, correct answers are also given soon afterward, in the text material that follows.

At various places in the text, brief comments appear in parentheses. These parenthetical comments give extra explanation or clarification, which some students will find helpful to their understanding. When students find the extra parenthetical clarification unnecessary, they may simply skip over it with a quick glance.

Appendices include: (1) a review of basic math, (2) the algebra and geometry of straight lines, (3) statistical tables, (4) review material for a final exam, (5) a glossary of symbols, (6) references, and (7) answers to all text exercises.

Boxes are occasionally used in the basic sections to add useful, but not essential, information.

Coverage allows for the ever-increasing use of preprogrammed calculators owned by behavioral science departments or by students. These calculators automatically compute such statistics as correlation and t at the touch of a key. Their use is the reason for the emphasis on conceptual understanding of formulas rather than on routine mechanics of computation.

The first goal of each student should be to ensure a uniformly high level of mastery of all the basic material. The ambitious student is advised to understand and apply the basic material to the solution of the basic exercises and the Study Guide problems. Whenever it is practical, two or more different ways of solving each type of problem are taught. Thus, solutions arrived at by one process can be checked by a different process to ensure accuracy. Confident students should also test their competence over all the basic material in such a way as to prevent gaps and inconsistencies in their knowledge. Once mastery of the basic material has been indicated by means of successful solutions to the basic text exercises and Study Guide problems, students may in some cases proceed to the supplementary material.

Challenging the More Ambitious and Capable Students

A substantial amount of useful information appears at the end of each

chapter, in the supplementary material. This material can be used in various ways, or not used at the discretion of the instructor.

Instructors who wish to concentrate on only the most basic and most useful material may assign only readings in the first, or basic, portion of each chapter. Other instructors may assign to all students the basic material plus certain selections from the supplementary material that they believe to be especially important. Or, if they wish, instructors may choose to assign readings in some supplementary sections only to students seeking the best grades or to students planning to study or use statistics beyond the introductory level. Supplementary sections that are not assigned by the instructor may be interesting and informative to students and studied at the students' option. Most supplementary sections are written to be comprehensible independent of the other supplementary material, in order to maximize the flexibility of instructor and student choices.

Summary of Main Objectives

This text's approach should, above all, help to give interest, confidence, and mastery to students who are capable of verbal and graphic understanding, even though their mathematical skills may be in some respects deficient. The approach should also help prevent students from getting in over their heads on topics that they imperfectly understand. Finally, the approach should help minimize embarrassing gaps in basic knowledge and applications. Therefore, students capable of higher levels of mastery should be able to do uniformly superior work on all the main types of statistical applications.

Though the first statistical methodology course is plagued with certain mathematical obstacles not encountered in most so-called "content" courses, the obstacles need not be overwhelming to a student who is motivated and able to do well in the other types of courses. Indeed, it is quite satisfying to see statistics students with bad former experiences in mathematics, high anxiety, initial disinterest, and low self-confidence persevere to master introductory statistics as capably—perhaps even more capably—than they master their other courses. And it is equally satisfying to see the better students achieve uniformly high, insightful, and comprehensive levels of mastery of all types of statistical tests covered.

Acknowledgments

Special thanks are due to the following individuals for their reviews of the text: Professors Charles R. Kessler (Los Angeles Pierce College),

Gregory Kimble (University of Colorado, Boulder), Joe W. Darnall (Harden-Simmons University), Eric R. Brown (University of Houston), Dennis Byrnes (University of Massachusetts, Boston), William A. Frederickson (Central State University, Edmond, Oklahoma), and Dr. Roger Bakeman (Georgia State University). The author gratefully acknowledges indebtedness to his statistics teachers, including Fred Sheffield, John Digman, Edward Crothers, Gregory Kimble, and Walter Kintsch. Special help in the teaching of statistics has come from Dave Chiszar and Sue Donaldson. Research help was provided by Ruth Hahn, and secretarial assistance was given by Rosanne Parady and Ramona Gehring. I am grateful 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 reprint Tables III, IV, V, VII, and XXXIII from their book *Statistical Tables for Biological, Agricultural and Medical Research*, 6th edition, 1974. Finally, many thanks are due to students at the University of Hawaii, the University of Colorado, and Indiana State University Evansville, who encouraged me to formulate the basic statistical concepts with the student's frame of reference foremost in mind.

R. E. G.

Contents

Preface *xvii*



1 Why Study Statistics? 1

Descriptive Statistics	1
What Inferential Statistics Tell Us	2
The Two Main Types of Statistics: Descriptive and Inferential	3
Sample Statistics Versus Population Parameters	3
The Nature of Data	5
Variables Versus Constants	5
Discrete Versus Continuous Variables	6
Making Statistical Understanding as Easy as Possible	6
Summary	7
Review of Most Important Points	8
Exercises	8

Supplement 12

Types of Numeric Scales	12
Two Types of Inferential Statistics: Parametric and Nonparametric	12
Review of Most Important Points	13
Supplementary Exercises	13

2 Summarizing Data Distributions Statistically 15

Frequency Distributions and Frequency

Graphs	15
Rounding	15
Tallying a Frequency Distribution	17
Plotting Frequency Data Graphically	18
The Frequency Histogram	19
A Grouped Frequency Distribution	20
Real Upper and Lower Limits	21
The Frequency Polygon	23

Traditional Guidelines for Making a Grouped Frequency Distribution	24
Relative Frequency	24
Cumulative Frequency and Cumulative Relative Frequency	26
Percentile	28
Summary	30
Exercises	31

Determining the Central Tendency of Distributions 32

Summing the Data and Finding the Mean	33
Other Measures of Central Tendency: The Median and the Mode	37
Summary	41
Exercises	42

Determining the Variability of Distributions 43

The Range	44
Preferred Measures of Variability: Variance and Standard Deviation	45
Computational Formulas for Variance and Standard Deviation	50
Variability and Hyperactive Children: An Application	51
General Principles of Graph Interpretation: The Pursuit of Happiness	52
Summary	52
Exercises	54
Review of Most Important Points	57
Exercises	57

Supplement 58

An Easy Way of Obtaining the Mean	58
Computing Skew	59
Precise Calculation of Percentile	60
Precise Calculation of the Median	63
Quartiles and Deciles	63
Computing the Mean and Standard Deviation from a Frequency Distribution	63
Algebraic Summation Rules	65
Effect of Score Changes on the Mean	67
Effect of Score Changes on the Variance and Standard Deviation	68

Deriving the Computational Σd^2 Formula from the Definitional Formula	70
Review of Most Important Points	73
Supplementary Exercises	74

3 Standard (z) Scores and Their Relation to Normal Distributions 79

Standardizing Sample Data 79

Additional Properties of Standardized Distributions	82
The Relationship Between z Scores and Raw Scores	83
The Relationship Between Percentile and z Score	84
Percentiles and z Scores in Normal Distributions	86
Summary	86
Review of Most Important Points	87
Exercises	87

The Characteristics of a Normal Distribution 88

Parameters of the Normal Distribution	93
Relative Frequencies in Normal Distributions	94
Relating Standard (z) Scores, Percentages, and Percentiles	99
Converting z Scores to Percentiles and Vice Versa	100
IQ Scores	102
Summary	103
Review of Most Important Points	104
Exercises	105

Supplement 107

SAT and Other National College Tests	107
CLEP Tests and T Scores	107
Why Normality?	107
The Standard-Score Formula for the Normal Curve	108
Proving that the Standard Deviation and the Variance of a Distribution of z Scores Are Both 1.0	109
Review of Most Important Points	111
Supplementary Exercises	111

4 The Logic of Statistical Inference and z Testing 113

Random Sampling	113
Basic Probability	114
Independent Choices in Random Sampling	114
Proving Principles Beyond Reasonable Doubt	116
A Research Question: Can Biofeedback Training Help Prevent Headaches?	116
The Hypotheses Involved in Making Inference	118
The 5% Alpha (α) Level	120
Inference: z Testing	121
Using the Sample Mean (\bar{X}) to Estimate the Population Mean (μ)	123
Biased and Unbiased Estimates	124
The Accuracy of a Sample Mean \bar{X} as an Estimator of μ	125
A Sampling Distribution of Means of Samples	129
An Estimation Rule for Finding $\sigma_{\bar{X}}$	131
Probability as Relative Frequency	132
Using the Standard Error of the Mean to Make Inference	132
Common Significance (α) Levels and Corresponding z_{crit} Values	134
Two-Tailed or Nondirectional Significance Levels	135
The Probability of Accepting H_1 Erroneously	138
Type I Error: Accepting H_1 When H_0 Is True	138
Type II Error: Accepting H_0 When H_1 Is True	138
Assumptions Underlying z Testing	139
Setting up Statistical Criteria Before Running a Study	140
An Additional Research Example	140
Summary	141
Review of Most Important Points	143
Exercises	144

Supplement 151

Principles of Experimental Design and Interpretation	151
--	-----

Understanding a Sampling Distribution of Means	157
Other Useful Types of z Tests	160
Review of Most Important Points	162
Supplementary Exercises	162

5 t Testing a Single Mean 166

A Research Problem with t : Reducing Tension with Transcendental Meditation	167
Estimating the Population Standard Deviation (σ)	168
Using $s_{\bar{x}}$ to Estimate $\sigma_{\bar{x}}$	170
Using the t Formula	171
Picturing t Distributions Versus z Distributions	172
Using the t Table	174
t Testing a Single Mean Against a Calculated Theoretical Value	176
Assumptions Underlying t	179
Summary	179
Review of Most Important Points	181
Exercises	181

Supplement 185

Confidence Intervals	185
“Student’s” Contribution	188
Demonstrating that s Is an Unbiased Estimator of a Population σ	189
Review of Most Important Points	192
Supplementary Exercises	192

6 Comparing Two Sample Means with t 194

t Tests for Correlated and Independent Samples	194
Comparing the Means of Two Correlated Samples with t	196
A Preview of Correlation in Correlated Samples	198
Reducing Two Correlated Samples to One Set of Difference Scores	199
Performing the t Calculation	200
The Number of df with $t_{\text{correlated}}$	201

Assumptions Underlying t for Correlated Samples	202
Comparing Means of Two Independent Samples	202
A Research Problem: LSD and Mystical Experience	203
Assumptions Underlying t for Independent Samples	208
Summary	208
Review of Most Important Points	210
Exercises	210
Supplement	214
Confidence Intervals for Mean Differences	214
Testing for Homogeneity with F	216
The Heterogeneity Formula for t , Independent Samples	218
Nonparametric Alternatives to t Tests	218
Deriving the Formulas for t_{hetero} and t_{homo}	219
Review of Most Important Points	225
Supplementary Exercises	226

7 One-Factor Analysis of Variance (ANOVA), Independent Samples 229

The Purpose of Analysis of Variance	229
A Research Problem: Teaching Creativity	229
The Logic of ANOVA	231
The Three Types of Deviation in ANOVA	235
Defining and Computing ANOVA	237
Interpreting ANOVA Results	243
Assumptions Underlying ANOVA	243
ANOVA with Unequal n	244
Summary	245
Review of Most Important Points	246
Exercises	247

Supplement 251

Hartley's F_{max} Test for Homogeneity	251
The Scheffé Method of Contrasts	252
A Procedure for <i>a priori</i> or Planned Comparisons	254
The Relationship Between t and F	255
t^2 and F	258

The Equivalence of Theoretical and Computational Formulas for SS_{bet}	258
A Nonparametric Alternative to ANOVA	259
Review of Most Important Points	260
Supplementary Exercises	260

8 Two-Factor Analysis of Variance 263

Introduction to Interaction	267
Computing SS for the Main Effects as in Simple ANOVA	270
Cells for Experimental Combinations	271
Calculating the Interaction SS	272
Error Within Cells	272
Finding the Degrees of Freedom	273
Pooling All Experimental (Between-Conditions) Effects	274
Completing the ANOVA	275
Interpreting Interactions	277
Assumptions Underlying Two-Factor ANOVA	277
Summary	278
Review of Most Important Points	279
Exercises	280

Supplement 286

More Complex Interaction Interpretations	286
Correlated Versus Independent Samples	287
Review of Most Important Points	287
Supplementary Exercises	288

9 Chi Square 289

Parametric Versus Nonparametric Tests	289
Introduction to Chi Square	289
Examples of Simple Chi Square	290
Interpreting Significance with More than One Degree of Freedom	301
Assumptions Underlying Chi Square	301
Summary	301
Review of Most Important Points	302
Exercises	302

Supplement 309

An Alternative General Formula for Chi Square 309

Correction for Continuity with One Degree of Freedom 309

An Easy Computational Formula for a 2×2 Table 311

Chi Square Testing Goodness of Fit 312

The Alternative Chi Square Formula with Correction for Continuity 315

Chi Square and Σz^2 316

Deriving the Expectancy Rule for 2×2 (and Larger) Tables 317

Computing Cell Expectancies: Joint Probability Method 319

Measures of Strength of Association in Two-Way Classification Tables 320

Review of Most Important Points 320

Supplementary Exercises 320

10 Correlation 323

Positive Correlation 323

Negative Correlation 326

The Range of Correlation, -1.00 to $+1.00$ 328

Zero Correlation 329

Scatter Plots 329

A Research Problem: Major Life Changes and Illness 333

Computing Correlation: Simplest Definitional Formula 335

Comparing Strength (Magnitude) and Sign in Correlation 340

A Computational Formula for Correlation 340

Testing r for Statistical Significance 341

Assumptions Underlying Correlation 345

Basic Correlation Is Linear Correlation 345

Causation Versus Correlation 346

Summary 346

Review of Most Important Points 348

Exercises 349

Supplement 355

Linear and Nonlinear Relationships 355