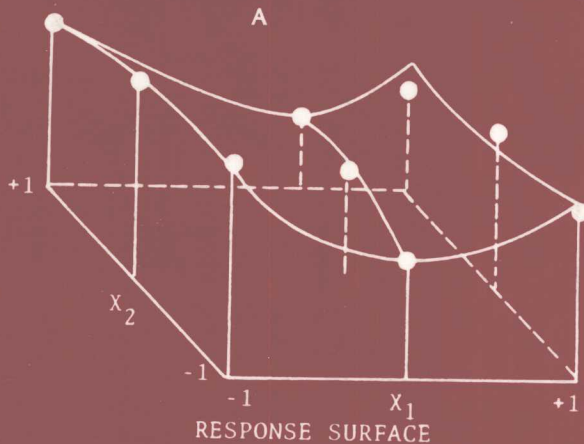


Pharmaceutical Statistics

Practical and Clinical Applications

Third Edition



Sanford Bolton

Pharmaceutical Statistics

Practical and Clinical Applications

Third Edition

Sanford Bolton

*Visiting Professor
University of Arizona
Tucson, Arizona*



MARCEL DEKKER, INC.

NEW YORK • BASEL • HONG KONG

Library of Congress Cataloging-in-Publication Data

Bolton, Sanford

Pharmaceutical statistics: practical and clinical applications /
Sanford Bolton. — 3rd ed.

p. cm. — (Drugs and the pharmaceutical sciences ; 80)

Includes bibliographical references and index.

ISBN 0-8247-9812-0 (alk. paper)

1. Pharmacy—Statistical methods. I. Title. II. Series: Drugs
and the pharmaceutical sciences ; v. 80.

[DNLM: 1. Technology, Pharmaceutical—methods. 2. Statistics.

W1 DR893B v.180 1997 / QV 25 B694p 1997]

RS57.B65 1997

615'.1'072—dc21

DNLM/DLC

for Library of Congress

96-49907

CIP

The publisher offers discounts on this book when ordered in bulk quantities.
For more information, write to Special Sales/Professional Marketing at the
address below.

This book is printed on acid-free paper.

Copyright © 1997 by Marcel Dekker, Inc. All Rights Reserved.

Neither this book nor any part may be reproduced or transmitted in any
form or by any means, electronic or mechanical, including photocopying,
microfilming, and recording, or by any information storage and retrieval
system, without permission in writing from the publisher.

Marcel Dekker, Inc.

270 Madison Avenue, New York, New York 10016

Current printing (last digit):

10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

The first edition of this book was published in 1984. Its modest aim—to help publicize the importance of statistics and expand its use among pharmaceutical scientists—has been met, as attested by its wide use and popularity. I can think of few more satisfying and heartwarming feelings than those imparted by my pharmaceutical colleagues when I am told how they have read and used the book to advantage. Most of these stories have a common theme that confirms the fulfillment of my original objective. Use of a simple, practical approach to current statistical problems that are commonly encountered in pharmaceutical research is an effective way of communicating statistics to the practicing pharmaceutical scientist. One statistician, the director of a statistical group at a major pharmaceutical company, told me that he gives a copy of the book to new statisticians so that they can get a general overview of the types of problems encountered in the pharmaceutical industry. Another scientist told me that he keeps three copies, one for himself and two to lend to friends and colleagues. This news is encouraging and gratifying, considering my aims and hopes for this book.

Therefore, the theme of the book remains the same: to present an uncomplicated explanation of the approaches and solutions to commonly encountered statistical problems, with examples that are relevant to scientists involved in pharmaceutical and related research. Because real examples often have different twists, one should be cautioned to take each example in context, and always seek expert advice, until experience gives sufficient confidence.

This third edition follows the approach of the previous editions. It includes more examples and techniques that reflect current practice and problems, and the continued emphasis and reliance on statistical techniques in drug and pharmaceutical research. Statistical tables and concepts have been expanded. In particular, analysis of bioequivalence studies, clinical studies, quality control, and concepts in GMP validation topics have been expanded. The “Barr decision,” resulting from a trial in which Barr Laboratories challenged an FDA injunction, has had a profound effect on the entire pharmaceutical industry. Some of the decisions resulting from the trial have been controversial, and my evaluation is presented in this new edition.

Among the new and expanded topics in this edition are recent developments in stability data analysis, concepts of statistical outliers, more detailed discussion of bioequivalence studies including some nonparametric applications, problems in sampling and devising limits for product release, covariance analysis, tolerance intervals, multiple endpoints, more nonparametric tests, and additional topics in clinical data analysis.

Although details of the statistical computations are presented in the book, current-day statistical analyses cannot be performed economically without computers. Otherwise, the structure remains similar to the previous editions. The initial chapters are general and introductory, whereas the later and expanded chapters are more specific and applied. No previous statistical education or higher mathematics is needed to understand the concepts of this book.

The exercises are an integral part of the book and can be used to test the reader’s comprehension of the text. More difficult or advanced problems are designated by double asterisks.

There are many people who deserve credit and thanks for this endeavor. In particular, I want to thank the graduate students with whom I have had the privilege to work. They are part of my family and have always inspired me. A very special expression of gratitude is extended to Jenny Chen, who volunteered to redo the figures using computer software. Also, many thanks are due to the PMA and generic pharmaceutical companies, both scientists and management, who have given me the opportunity of working with real-life stimulating problems. Once again, I want to acknowledge the encouragement and inspiration instilled in me by the late Dr. John Fertig of Columbia University and the late Dr. Takeru Higuchi, my mentors.

Finally, my sincere wish is that this book will continue to be useful to my fellow scientists and statisticians.

Sanford Bolton

CONTENTS

Preface	v
1. Basic Definitions and Concepts	1
1.1 Variables and Variation	1
1.2 Frequency Distributions and Cumulative Frequency Distributions	5
1.3 Sample and Population	12
1.4 Measures Describing the Center of Data Distributions	14
1.5 Measurement of the Spread of Data	20
1.6 Coding	28
1.7 Precision, Accuracy, and Bias	30
1.8 The Question of Significant Figures	33
Key Terms	35
Exercises	35
References	37
2. Data Graphics	39
2.1 Introduction	40
2.2 The Histogram	41
2.3 Construction and Labeling of Graphs	41
2.4 Scatter Plots (Correlation Diagrams)	51
	vii

2.5	Semilogarithmic Plots	51
2.6	Other Descriptive Figures	54
	Key Terms	58
	Exercises	58
	References	61
3.	Introduction to Probability: The Binomial and Normal Probability Distributions	62
3.1	Introduction	62
3.2	Some Basic Probability	63
3.3	Probability Distributions—The Binomial Distribution	68
3.4	Continuous Data Distributions	78
3.5	Other Common Probability Distributions	94
	Key Terms	98
	Exercises	98
	References	101
4.	Choosing Samples	102
4.1	Introduction	102
4.2	Random Sampling	104
4.3	Other Sampling Procedures: Stratified, Systematic, and Cluster Sampling	110
4.4	Sampling in Quality Control	113
	Key Terms	116
	Exercises	116
	References	119
5.	Statistical Inference: Estimation and Hypothesis Testing	120
5.1	Statistical Estimation (Confidence Intervals)	121
5.2	Statistical Hypothesis Testing	131
5.3	Comparison of Variances in Independent Samples	174
5.4	Test of Equality of More than Two Variances	178
5.5	Confidence Limits for a Variance	179
5.6	Tolerance Intervals	181
	Key Terms	183
	Exercises	183
	References	187
6.	Sample Size and Power	189
6.1	Introduction	190

6.2	Determination of Sample Size for Simple Comparative Experiments for Normally Distributed Variables	192
6.3	Determination of Sample Size for Binomial Tests	199
6.4	Determination of Sample Size to Obtain a Confidence Interval of Specified Width	201
6.5	Power	203
6.6	Sample Size and Power for More than Two Treatments	207
6.7	Sample Size for Bioequivalence Studies	210
	Key Terms	211
	Exercises	213
	References	214
7.	Linear Regression and Correlation	216
7.1	Introduction	217
7.2	Analysis of Standard Curves in Drug Analysis: Application of Linear Regression	222
7.3	Assumptions in Tests of Hypotheses in Linear Regression	224
7.4	Estimate of the Variance: Variance of Sample Estimates of the Parameters	225
7.5	A Drug Stability Study: A Second Example of the Application of Linear Regression	229
7.6	Confidence Intervals in Regression Analysis	234
7.7	Weighted Regression	239
7.8	Analysis of Residuals	241
7.9	Nonlinear Regression	243
7.10	Correlation	248
7.11	Comparison of Variances in Related Samples	257
	Key Terms	259
	Exercises	259
	References	263
8.	Analysis of Variance	265
8.1	One-Way Analysis of Variance	265
8.2	Planned Versus a Posteriori (Unplanned) Comparisons in ANOVA	274
8.3	Another Example of One-Way Analysis of Variance: Unequal Sample Sizes and the Fixed and Random Models	286
8.4	Two-Way Analysis of Variance (Randomized Blocks)	289

8.5	Statistical Models**	306
8.6	Analysis of Covariance**	308
8.7	ANOVA for Pooling Regression Lines as Related to Stability Data	315
	Key Terms	320
	Exercises	321
	References	325
9.	Factorial Designs**	326
9.1	Definitions (Vocabulary)	326
9.2	Two Simple Hypothetical Experiments to Illustrate the Advantages of Factorial Designs	331
9.3	Performing Factorial Experiments: Recommendations and Notation	335
9.4	A Worked Example of a Factorial Experiment	338
9.5	Fractional Factorial Designs	345
9.6	Some General Comments	351
	Key Terms	351
	Exercises	351
	References	354
10.	Transformations and Outliers	355
10.1	Transformations	356
10.2	Outliers	369
	Key Terms	382
	Exercises	382
	References	383
11.	Experimental Design in Clinical Trials	384
11.1	Introduction	384
11.2	Some Principles of Experimental Design and Analysis	385
11.3	Parallel Design	391
11.4	Crossover Designs and Bioavailability/Bioequivalence Studies	397
11.5	Repeated Measures (Split-Plot) Designs**	428
11.6	Multiclinic Studies	435
11.7	Interim Analyses	437

**A more advanced topic.

Key Terms	440
Exercises	440
References	442
12. Quality Control	444
12.1 Introduction	444
12.2 Control Charts	445
12.3 Acceptance Sampling and Operating Characteristic Curves	463
12.4 Statistical Procedures in Assay Development	467
12.5 Establishing In-House Limits	478
12.6 Some Statistical Aspects of Quality and the “Barr Decision”	482
Key Terms	487
Exercises	487
References	489
13. Validation	490
13.1 Process Validation	491
13.2 Assay Validation	505
13.3 Concluding Remarks	513
Key Terms	513
Exercises	513
References	514
14. Consumer Testing	515
14.1 Introduction	515
14.2 Product Comparisons	516
14.3 Discrimination Tests: The Triangle Test	529
Key Terms	535
Exercises	535
References	536
15. Nonparametric Methods	537
15.1 Data Characteristics and an Introduction to Nonparametric Procedures	538
15.2 Sign Test	541
15.3 Wilcoxon Signed Rank Test	544
15.4 Wilcoxon Rank Sum Test (Test for Differences Between Two Independent Groups)	549

15.5	Kruskal–Wallis Test (One-Way ANOVA)	557
15.6	Friedman Test (Two-Way Analysis of Variance)	560
15.7	Nonparametric Analysis of Covariance	565
15.8	Runs Test for Randomness	567
15.9	Contingency Tables	570
	Key Terms	584
	Exercises	584
	References	588
16.	Optimization Techniques**	590
16.1	Introduction	590
16.2	Optimization Using Factorial Designs	591
16.3	Composite Designs to Estimate Curvature	604
16.4	The Simplex Lattice	610
16.5	Sequential Optimization	619
	Key Terms	624
	Exercises	624
	References	625
Appendix I:	Some Properties of the Variance	627
I.1	Pooling Variances	627
I.2	Components of Variance	628
I.3	Variance of Linear Combinations of Independent Variables	629
	Reference	630
Appendix II:	Comparison of Slopes and Testing of Linearity: Determination of Relative Potency	631
	Reference	637
Appendix III:	Multiple Regression	638
	References	644
Appendix IV:	Tables	645
Table IV.1	Random Numbers	645
Table IV.2	Cumulative Normal Distribution: Cumulative Area Under the Normal Distribution (Less Than or Equal to Z)	646

**A more advanced topic.

Table IV.3	Individual Terms of the Binomial Distribution for $N = 2$ to 10 and $P = 0.2, 0.5,$ and 0.7	647
Table IV.4	t Distributions	650
Table IV.5	Short Table of Chi-Square Distributions	651
Table IV.6A	Upper 5% Values of the F Distribution	652
Table IV.6B	Short Table of Upper 1% Values of the F Distribution	653
Table IV.7A	Upper 5% Points in the Studentized Range	654
Table IV.7B	Values of t' for Dunnett's Comparison of Several Treatments and a Control ($\alpha = 0.05$)	655
Table IV.8	Dixon's Criteria for Rejecting Outliers	656
Table IV.9	Critical Values of T for a Two-Sided Test at the 5% Level of Significance (Test for Outliers)	657
Table IV.10	Factors for Determining Upper and Lower 3σ Limits for Mean (\bar{X}) and Range (R) Charts, and for Estimating σ from \bar{R}	658
Table IV.11	Number of Correct Guesses Needed for Significance in the Triangle Test	659
Table IV.12	Number of Positive or Negative Signs Needed for Significance for the Sign Test	660
Table IV.13	Values Leading to Significance for the Wilcoxon Signed Rank Test (Two-Sided Test)	661
Table IV.14	Critical Values for Number of Runs at the 5% Level of Significance	662
Table IV.15	Probability of Getting at Least One Run of Given Size for N Samples	662
Table IV.16	Critical Values for Wilcoxon Rank Sum Test ($\alpha = 0.05$)	663
Table IV.17	Critical Differences for Significance ($\alpha = 0.05$) Comparing All Possible Pairs of Treatments for Nonparametric One-Way ANOVA	664
Table IV.18	Critical Differences for Significance ($\alpha = 0.05$) Comparing All Possible Pairs of Treatments for Nonparametric Two-Way ANOVA	665

Table IV.19	Factors for Two-Sided Tolerance Limits for Normal Distributions	666
Table IV.20	Test for Outliers	672
Appendix V:	Outlier Tests and Chemical Assays	675
V.1	Introduction	675
V.2	Can Outlier Tests Be Justified?	676
V.3	Why Is There Not a USP Test for Outliers for Chemical Assays?	678
V.4	Some Comments on the Nature of Outliers and Outlier Tests, and Other Inconsistencies in the Decision That Outlier Tests Be Used for Biological Assays but Not for Chemical Assays	679
V.5	What Is the Purpose of Performing Replicate Assays and When is Averaging Appropriate?	680
V.6	In What Situations Might Outlier Tests Be Applicable? References	681 683
Appendix VI:	Should a Single Unexplained Failing Assay Be Reason to Reject a Batch?	685
	Conclusion	695
	References	696
Answers to Exercises		697
Index		725

BASIC DEFINITIONS AND CONCEPTS

Statistics has its own vocabulary. Many of the terms that comprise statistical nomenclature are familiar: some commonly used in everyday language, with perhaps, somewhat different connotations. Precise definitions are given in this chapter so that no ambiguity will exist when the words are used in subsequent chapters. Specifically, such terms as *discrete* and *continuous variables*, *frequency distribution*, *population*, *sample*, *mean*, *median*, *standard deviation*, *variance*, *coefficient of variation (CV)*, *range*, *accuracy*, and *precision* are introduced and defined. The methods of calculation of different kinds of means, the median, standard deviation, and range are also presented. When studying any discipline, the initial efforts are most important. The first chapters of this book are important in this regard. Although most of the early concepts are relatively simple, a firm grasp of this material is essential for understanding the more difficult material to follow.

1.1 VARIABLES AND VARIATION

Variables are the measurements, the values, which are characteristic of the data collected in experiments. These are the data that will usually be displayed, analyzed, and interpreted in a research report or publication. In statistical terms, these observations are more correctly known as *random variables*. Random variables take on values, or numbers, according to some corresponding probability function. Although we will wait until Chapter 3 to discuss the concept of probability, for the present we can think of a

random variable as the typical experimental observation that we, as scientists, deal with on a daily basis. Because these measurements may take on different values, repeat measurements observed under apparently identical conditions do not, in general, give the identical results (i.e., they are usually not exactly reproducible). Duplicate determinations of serum concentration of a drug 1 hr after an injection will not be identical no matter if the duplicates come from (a) the same blood sample or (b) from separate samples from two different persons or (c) from the same person on two different occasions. Variation is an inherent characteristic of experimental observations. To isolate and to identify particular causes of variability requires special experimental designs and analysis. Variation in observations is due to a number of causes. For example, an assay will vary depending on:

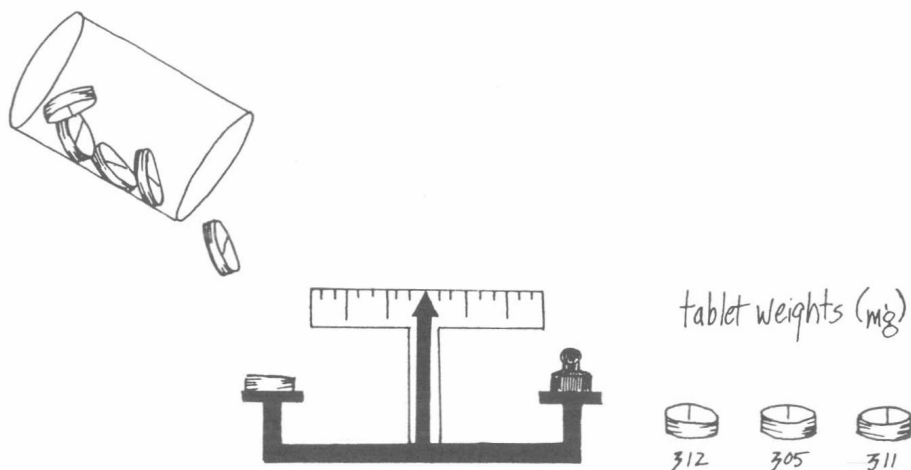
1. The instrument used for the analysis
2. The analyst performing the assay
3. The particular sample chosen
4. Unidentified, uncontrollable background error, commonly known as “noise”

This inherent variability in observation and measurement is a principal reason for the need of statistical methodology in experimental design and data analysis. In the absence of variability, scientific experiments would be short and simple: interpretation of experimental results from well-designed experiments would be unambiguous. In fact, without variability, single observations would often be sufficient to define the properties of an object or a system. Since few, if any, processes can be considered absolutely invariant, statistical treatment is often essential for summarizing and defining the nature of data, and for making decisions or inferences based on these variable experimental observations.

1.1.1 Continuous Variables

Experimental data come in many forms*. Probably the most commonly encountered variables are known as *continuous variables*. A continuous variable is one that can take on *any* value within some range or interval (i.e., within a specified lower and upper limit). The limiting factor for the total number of possible observations or results is the sensitivity of the measuring instrument. When weighing tablets or making blood pressure measurements,

*For a further discussion of different kinds of variables, see Sec. 15.1 in Chapter 15, Nonparametric Methods.



Tablet weights: an example of a variable measurement (a random variable).

there are an infinite number of possible values that can be observed if the measurement could be made to an unlimited number of decimal places. However, if the balance, for example, is sensitive only to the nearest milligram, the data will appear as discrete values. For tablets targeted at 1 g and weighed to the nearest milligram, the tablet weights might range from 900 to 1100 mg, a total of 201 possible integral values (900, 901, 902, 903, . . . , 1098, 1099, 1100). For the same tablet weighed on a more sensitive balance, to the nearest 0.1 mg, values from 899.5 to 1100.4 might be possible, a total of 2010 possible values, and so on.

Often, continuous variables cannot be easily measured but can be ranked in order of magnitude. In the assessment of pain in a clinical study of analgesics, a patient can have a continuum of pain. To measure pain on a continuous numerical scale would be difficult. On the other hand, a patient may be able to differentiate slight pain from moderate pain, moderate pain from severe pain, and so on. In analgesic studies, scores are commonly assigned to pain severity, such as no pain = 0, slight pain = 1, moderate pain = 2, and severe pain = 3. Although the scores cannot be thought of as an exact characterization of pain, the value 3 does represent more intense pain than the values 0, 1, or 2. The scoring system above is a representation of a continuous variable by discrete "scores" which can be rationally ordered or ranked from low to high. This is commonly known as a rating scale, and the ranked data are on an ordinal scale. The rating scale is an effort to quantify a continuous, but subjective, variable.

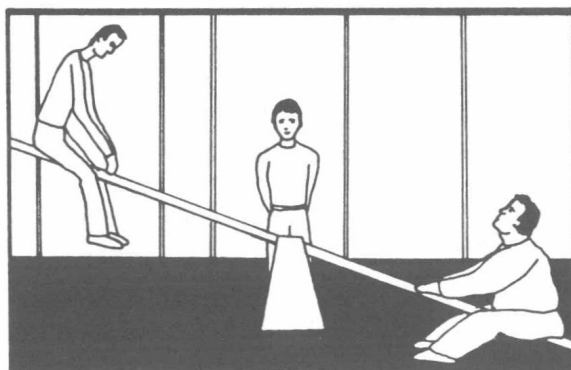
1.1.2 Discrete Variables

In contrast to continuous variables, *discrete variables* can take on a countable number of values. These kinds of variables are commonly observed in biological and pharmaceutical experiments and are exemplified by measurements such as the number of anginal episodes in 1 week or the number of side effects of different kinds after drug treatment. Although not continuous, discrete data often have values associated with them which can be numerically ordered according to their magnitude, as in the examples given earlier of a rating scale for pain and the number of anginal episodes per week.

Discrete data that can be named (nominal), categorized into two or more classes, and counted are called categorical variables, or *attributes*; for example, the attributes may be different side effects resulting from different drug treatments or the presence or absence of a defect in a finished product. These kinds of data are frequently observed in clinical and pharmaceutical experiments and processes. A finished tablet classified in quality control as “defective” or “not defective” is an example of a categorical or attribute type of variable. In clinical studies, the categorization of a patient by sex (male or female) or race is a classification according to attributes. When calculating ED_{50} or LD_{50} , animals are categorized as “responders” or “non-responders” to various levels of a therapeutic agent, a categorical response. These examples describe variables that cannot be ordered. A male is not associated with a higher or lower numerical value than a female.

Continuous variables can always be classified into discrete classes where the classes are ordered. For example, patients can be categorized as “un-

Classification by attributes: patients categorized by weight.



Underweight Normal weight Overweight