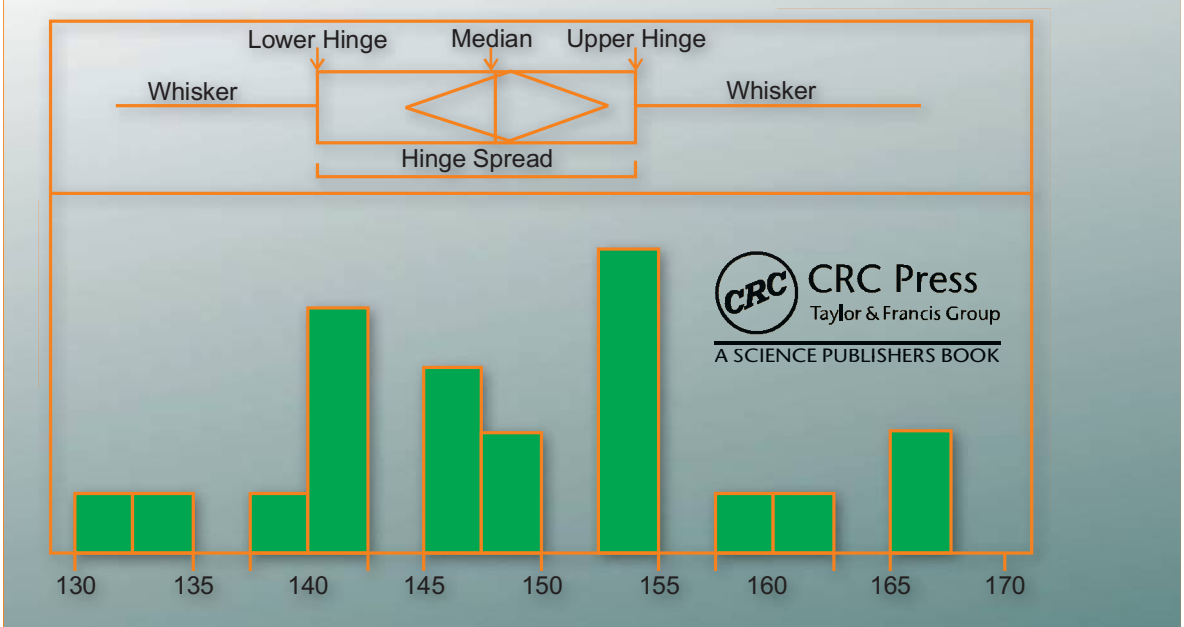
A Handbook of

Applied Statistics in Pharmacology

Katsumi Kobayashi
K. Sadasivan Pillai



CRC Press
Taylor & Francis Group
A SCIENCE PUBLISHERS BOOK

A Handbook of Applied Statistics in Pharmacology

A Handbook of Applied Statistics in Pharmacology

Katsumi Kobayashi
Safety Assessment Division, Chemical Management Center
National Institute of Technology and Evaluation (NITE)
Tokyo, Japan

K. Sadasivan Pillai
Frontier Life Science Services
(A Unit of Frontier Lifeline Hospitals)
Thiruvallur District
Chennai, India

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A SCIENCE PUBLISHERS BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2013 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Version Date: 2012919

International Standard Book Number-13: 978-1-4665-1540-6 (eBook - PDF)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Foreword

Life expectancy has significantly increased in the last century, thanks to the discovery and development of new drugs by pharmaceutical industries. Search for new therapeutics is the primary activity of the R&D of pharmaceutical industries and it involves complex network of tasks such as synthetic chemistry, *in vitro*/*in vivo* efficacy, safety, preclinical and clinical research. Statistical analysis has always been the foundation to establish the safety and efficacy of drugs. The decision to- or not to- advance preclinical drug candidates to very expensive clinical development heavily relies on statistical analysis and the resulting significance of preclinical data. Recent reports attributed failure of certain drugs in clinical stages of development to improper conduct of preclinical studies and inappropriate application of statistical tools. Applying appropriate statistical tools is sagacious to analysis of data from any research activity. Though scientists expect computerized statistical packages to perform analyses of the data, he/she should be familiar with the underlying principles to choose the appropriate statistical tool.

'A Handbook of Applied Statistics in Pharmacology' by Katsumi Kobayashi and K. Sadasivan Pillai is a very useful book for scientists working in R&D of pharmaceuticals and contract research organizations. Most of the routine statistical tools used in pharmacology and toxicology are covered perspicuously in the book. The examples worked out in the book are from actual studies, hence do not push a reader having less or no exposure to statistics outside his/her comfort zone.

Dr. K.M. Cherian
M.S., F.R.A.C.S., Ph.D., D.Sc. (Hon.), D.Sc. (CHC), D.Sc. (HC)
Chairman & CEO
Frontier Lifeline Hospitals
Chennai, India

Preface

Scientists involved in pharmacology have always felt that statistics is a difficult subject to tackle. Thus they heavily rely on statisticians to analyse their experimental data. No doubt, statisticians with some scientific knowledge can analyse the data, but their interpretation of results often perplexes the scientists.

Statistics play an important role in pharmacology and related subjects like, toxicology, and drug discovery and development. Improper statistical tool selection to analyze the data obtained from studies conducted in these subjects may result in erroneous interpretation of the performance- or safety- of drugs. There have been several incidents in pharmaceutical industries, where failure of drugs in clinical trials is attributed to improper statistical analysis of the preclinical data. In pharmaceutical Research & Development settings, where a large number of new drug entities are subjected to high-throughput *in vitro* and *in vivo* studies, use of appropriate statistical tools is quintessential.

It is not prudent for the research scientists to totally depend on statisticians to interpret the findings of their hard work. Factually, scientists with basic statistical knowledge and understanding of the underlying principles of statistical tools selected for analysing the data have an advantage over others, who shy away from statistics. Underlying principle of a statistical tool does not mean that one should learn all complicated mathematical jargons. Here, the underlying principle means only 'thinking logically' or applying 'common sense'.

The authors of this book, with decades of experience in contract research organizations and pharmaceutical industries, are fully cognizant of the extent of literacy in statistics that the research scientists working in pharmacology, toxicology, and drug discovery and development would be interested to learn. This book is written with an objective to communicate statistical tools in simple language. Utmost care has been taken to avoid complicated mathematical equations, which the readers may find difficult

to assimilate. The examples used in the book are similar to those that the scientists encounter regularly in their research. The authors have provided cognitive clues for selection of an appropriate statistical tool to analyse the data obtained from the studies and also how to interpret the result of the statistical analysis.

Contents

Foreword		**v**
Preface		**vii**
1.	**Probability**	**1**
	Probability and Possibility	1
	Probability—Examples	2
	Probability Distribution	3
	Cumulative Probability	3
	Probability and Randomization	4
2.	**Distribution**	**6**
	History	6
	Variable	6
	Stem-and-Leaf Plot	7
	Box-and-Whisker Plot	8
3.	**Mean, Mode, Median**	**11**
	Average and Mean	11
	Mean	11
	Geometric Mean	12
	Harmonic Mean	12
	Weighted Mean	13
	Mode	13
	Median	14
4.	**Variance, Standard Deviation, Standard Error, Coefficient of Variation**	**16**
	Variance	16
	Standard Deviation (SD)	18
	Standard Error (SE)	19
	Coefficient of Variation (CV)	19
	When to Use a Standard Deviation (SD)/Standard Error (SE)?	20

5. Analysis of Normality and Homogeneity of Variance **23**
Distribution of Data in Toxicology and Pharmacology Experiments 23
Analysis of Normality 23
Tests for Analyzing Normal Distribution 24
Shapiro-Wilk's *W* test 24
Power of Shapiro-Wilk's *W* test 27
Parametric and Non-parametric Analyses 31
Analysis of Homogeneity of Variance 31
Bartlett's homogeneity test 31
Levene's homogeneity test 32
Power of Bartlett's and Levene's homogeneity tests 32
Do We Need to Examine the Data for Both Normality and Homogeneity? 33
Which Test to be Used for Examining Homogeneity of Variance? 33

6. Transformation of Data and Outliers **37**
Transformation of Data 37
Outliers 38
Masuyama's Rejection Limit Test 40
Thompson's Rejection Test 41
Smirnov-Grubbs' Rejection Test 42
A Cautionary Note 42

7. Tests for Significant Differences **47**
Null Hypothesis 47
Significant Level, Type I and Type II Errors 48
Why at 5% Significant Level? 48
How to Express *P*? 50
One-sided and Two-sided Tests 50
Which Test to Use: One-sided or Two-sided? 51

8. *t*-Tests **56**
Student's *t*-Test—History 56
t-Test for One Group 56
t-Test for Two Groups 57
Student's *t*-test 58
Aspin-Welch's *t*-test 60
Cochran-Cox's *t*-test 62
Paired *t*-Test 64
A Note of Caution 65

9. Correlation Analysis **67**
Correlation and Association 67
Pearson's Product Moment Correlation Coefficient 68
Significance of *r* 69
Confidence Interval of Correlation Coefficient 70
Coefficient of Determination 71
Rank Correlation 71
Spearman's Rank Correlation 71
Canonical Correlation 72
Misuse of Correlation Analysis 72

10. Regression Analysis **74**
History 74
Linear Regression Analysis 74
Confidence Limits for Slope 78
Comparison of Two Regression Coefficients 79
R^2 80
Multiple Linear Regression Analysis 80
Polynomial Regression 81
Misuse of Regression Analysis 82

11. Multivariate Analysis **84**
Analysis of More than Two Groups 84
One-way ANOVA 85
post hoc Comparison 87
Dunnett's multiple comparison test 87
Tukey's multiple range test 89
Williams's test 90
Duncan's multiple range test 95
Scheffé's multiple comparison test 98
Two-way ANOVA 100
Dunnett's Multiple Comparison Test and Student's *t* Test—A Comparison 103

12. Non-Parametric Tests **106**
Non-parametric and Parametric Tests—Assumptions 106
Sign Tests 106
Calculation procedure of sign test for small sample size 107
Calculation procedure of sign test for large sample size 107

Signed Rank Sum Tests 109
Wilcoxon rank-sum test 109
Fisher's exact test 111
Mann-Whitney's *U* test 113
Kruskal-Wallis Nonparametric ANOVA by Ranks 117
Comparison of Group Means 119
Dunn's multiple comparison test for more than three groups 120
Steel's multiple comparison test for more than three groups 121
Rank Sum Tests—Some Points 123

13. Cluster Analysis 127

What is Cluster Analysis? 127
Hierarchical cluster analysis 127
Ward's method of cluster analysis 128
k-means cluster analysis 129

14. Trend Tests 136

Introduction 136
Jonckheere's trend test 136
The Cochran-Armitage test 139

15. Survival Analysis 143

Introduction 143
Hazard Rate 144
Kaplan-Meier Method 144
Kaplan-Meier product-limit estimator 144

16. Dose Response Relationships 150

Dose and Dosage 150
Margin of Exposure, NOAEL, NOEL 150
Determining NOEL and NOAEL 152
Benchmark Dose 153
Probit Analysis 154
IC_{50} and EC_{50} Determination 157
Hormesis 158

17. Analysis of Pathology Data 163

Pathology in Toxicology 163
Analysis of Pathology Data of Carcinogenicity Studies 164
Peto test 165
Decision rules 166

Poly-*k* Type test 167
Analysis of Tumour Incidence—Comparison with Historical Control Data 168
Analysis of Incidence of Tumour Using X^2 Test 170
Comparison of Incidence of Tumours in Human, Rats, Mice and Dogs 170
Analysis of Organ Weight Data 172
Interpretation of Pathology Observations 173

18. Designing An Animal Experiment in Pharmacology and Toxicology—Randomization, Determining Sample Size 178

Designing Animal Experiments 178
Acclimation 179
Randomization 179
Determining Sample Size 183
Animal Experimental Designs 185

19. How to Select An Appropriate Statistical Tool? 188

Good Statistical Design 188
Decision Trees 188
Statistical Procedures Used by National Toxicology Program (NTP), USA 191
Decision Tree Produced by OECD 193
Incongruence in Selection of a Statistical Tool 194
Selection of a Statistical Tool—Suggested Decision Trees or Flow Charts 195
Statistical Tools Suggested for the Analysis of Toxicology Data 196
Use of Statistics in Toxicology-Limitations 196

Appendices 201

Appendix 1. Coefficient for Shapiro-Wilk *W* Test 202
Appendix 2. Quantiles of the Shapiro-Wilk Test Statistic 206
Appendix 3. *Z* Score for Normal Distribution 208

***Index* 211**