


BASIC STATISTICS FOR MEDICAL AND HEALTH SCIENCES



**Noorah Mohamed Salleh
Zairina Abdul Rahman
Mohd Dzulkhairi Mohd Rani**

BASIC STATISTICS FOR MEDICAL AND HEALTH SCIENCES



BASIC STATISTICS FOR MEDICAL AND HEALTH SCIENCES

Nooriah Mohamed Salleh
Zairina Abdul Rahman
Mohd Dzulkhairi Mohd Rani

USIM Publisher
Universiti Sains Islam Malaysia
Bandar Baru Nilai
Negeri Sembilan
2013

First Published in 2013

© Universiti Sains Islam Malaysia

All rights reserved; no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission from USIM publisher.

Published in Malaysia by:
USIM Publisher
Universiti Sains Islam Malaysia
71800 Bandar Baru Nilai
Negeri Sembilan Darul Khusus
Tel: 06-798 8226/6081 Faks: 06-798 6083
www.usim.edu.my

Printed in Malaysia by:
Card Information Sdn. Bhd.
49, Jalan 31/10A,
Taman Perindustrian IKS, Mukim Batu,
68100 Kuala Lumpur, Malaysia.
Tel: 03-6185 4008 Faks: 03-6187 4008

USIM Publisher is a member of the
MALAYSIAN SCHOLARLY PUBLISHING COUNCIL (MAPIM)

Perpustakaan Negara Malaysia

Cataloguing in Publication Data

Basic statistics for medical and health sciences

Bibliography: p.

ISBN 978-967-0393-50-6

1. Medical statistics. 2. Statistics.

519.502461



CONTENTS

FOREWORD	7
PREFACE	9
ACKNOWLEDGEMENTS	11

Chapter	Topic	Page
1	Introduction To Biostatistics	13
2	Types of Variables, Data and Scales of Measurement	17
3	Describing Data and Data Presentation	23
4	Measures of Central Tendency and Dispersion	35
5	Normal Distribution and Standard Normal Curve	43
6	Hypothesis Testing	51
7	Estimation: Point Estimation and Confidence Interval	61
8	t-tests	71
9	ANOVA	89
10	Chi-square Test	95
11	McNemar's Test	107
12	Nonparametric Tests	113
13	Correlation	137
14	Simple Linear Regression	145
15	Population, Sampling Methods and Sample Size Calculation	161
16	Rates and Ratios	175
17	Standardization of Rates	181
REFERENCES		189
INDEX		191
AUTHORS		193



FOREWORD

The importance of understanding Biostatistics in the fields of Medicine, Dentistry and the Allied Health Sciences is increasing as research findings continue to unravel novel discoveries in the understanding and management of diseases. Hence an essential component of the undergraduate curriculum for health professionals includes Biostatistics which will equip them with the skills of collating and summarizing data, applying the appropriate statistical tests and interpreting the results of statistical analyses.

The authors of this book have tried to approach the subject to make it simple and relevant for the particular needs of these future health professionals. Thus it is hoped that this book will benefit and facilitate the students' learning process.

I would like to commend the authors for their effort and hope that they will continue with their good work that will benefit students of Medicine, Dentistry and the Allied Health Sciences. Thank you.

With best wishes,

Professor Dr. Musa Ahmad
Deputy Vice Chancellor (Academic)
Universiti Sains Islam Malaysia (USIM)



PREFACE

This book is written by the Community Health lecturers of the Faculty of Medicine and Health Sciences, Universiti Sains Islam Malaysia to supplement the lectures and notes delivered to the medical students. It is hoped that the simplistic approach used throughout this book will facilitate the students' understanding of Biostatistics because of the challenges they have to surmount in terms of time allocation in view of the multiple subjects they have to master during their undergraduate years.

It is inevitable that practitioners of Medicine and the Health Sciences have to comprehend a vast amount of knowledge because of rapid advances in these fields. Current updates of medical knowledge are rapidly being published in journals every day. Thus to prepare the medical students for the real world they have to be equipped with the basic medical sciences and clinical skills and they also need to comprehend basic statistical concepts to have a better understanding of current medical advances reported in medical journals. Moreover, with the emphasis on the practice of Evidence-based Medicine it becomes more imperative for doctors to have a better understanding of statistical reports of research findings in order to apply current practices in the day-to-day management of their patients.

Medical students are already burdened with a host of subjects, so it is not surprising that they often view Statistics as another barrier to the advancement of their career. Statistics provide the tools that are essential in facilitating health professionals to summarize and present data; select and apply the appropriate statistical tests and interpret the results of research findings. Hence it is the purpose of this book to do away with formulae and complex mathematics because with the availability of statistical software such as SPSS, the students need not be burdened with unnecessary memorization of non-vital information.

Hopefully, this book will be useful for medical students who have doubts about their ability to master statistics or for those who wonder about the relevance of statistics in Medicine.



ACKNOWLEDGEMENTS

We would like to thank Professor Dr. Musa B. Ahmad for his continuous support and encouragement in the writing of this book. We would also like to extend our gratitude to our medical students for their feedback during their lectures. Our acknowledgement goes to our spouses and children for their patience, understanding and inspiration. Thank you.

CHAPTER 1

INTRODUCTION TO BIOSTATISTICS

Contents:

1. What is *statistics*?
2. What is *biostatistics*?
3. Why should medical students study biostatistics?
4. Uses of biostatistics in Medicine and other allied health sciences
5. Two main branches of statistics:
 - Descriptive
 - Inferential

Learning Objectives:

After studying this chapter you should be able to:

1. Define statistics.
2. List reasons for studying statistics.
3. List uses of statistics in Medicine and other allied health sciences.
4. Distinguish between descriptive and inferential statistics.



REFERENCES

- Archives of Orofacial Sciences 2006; 1:9-14
- Coakes S J, Steed L, Ong C. SPSS version 16.0 for Windows: Analysis without Anguish. Queensland: John Wiley & Sons Australia, Ltd.; 2009.
- Daniel W. Biostatistics: Basic concepts and methodology for the health sciences. Ninth edition. International Student Edition. John Wiley & Sons; 2010.
- Daniel W. Biostatistics: A foundation for analysis in the health sciences. Seventh edition. John Wiley & Sons; 1999.
- Dawson B, Trapp RG, Trapp R. Basic & Clinical Biostatistics: Fourth Edition. McGraw-Hill; 2011.
- Gaur AS, Gaur SS. Statistical Methods for Practice and Research. A guide to data analysis using SPSS. New Delhi: Response Books. Sage Publications; 2008.
- Kirkwood B, Jonathan S. Essential Medical Statistics. 2nd Edition. Oxford: Blackwell Science; 2003.
- Kuzma JW, Bohnenblust SE, Bohnenblust S. Basic Statistics for the Health Sciences. Fifth edition. Singapore: McGraw Hill International Edition; 2005.
- L.Naing, T.Winn, B.N. Rusli. Practical Issues in Calculating the Size for Prevalence Studies.
- Leech NL, Barrett KC, Morgan GA. SPSS for Intermediate Statistics Use and Interpretation. 3rd Edition. New York: Lawrence Erlbaum Associates; 2008.
- Pallant J. SPSS Survival Manual. Fourth edition. Australia: Allen & Unwin; 2011.



INDEX

A

Alpha 51, 57, 59
Alternative hypothesis (H_a) 88
Analysis of variance (ANOVA) 90
ANOVA (see Analysis of variance) 73,
89, 90, 92, 93, 114, 131, 154
Area under the curve 45
Association 52, 68, 96, 103, 104, 108,
139, 144, 147, 159, 179

B

Bar chart 24-26
Bimodal 36, 38
Bonferroni's method 131-133
Box-and-whiskers plot (boxplot) 41,
157
Boxplots 33

C

Categorical data 23, 24, 27, 95-97,
108, 191
Central limit theorem 48
Chi-square 5, 48, 95-100, 103-105,
108
Coefficient of determination (r^2) 155,
158
Coefficient of variation (CV) 35, 39,
41
Confidence interval 5, 57, 61, 63-68,
70, 77, 79, 81, 83, 87, 93, 114, 172
Continuous variable 19, 64, 151, 157
Correlation 5, 56, 159
Correlation coefficient (r) 137, 139,
140-144
Crude relative risk 176, 182, 185
Cumulative frequency 176
Cumulative incidence 68, 179
Cumulative probabilities 59
CV 41

D

Data presentation 23, 24, 28
Degree of freedom 81, 96, 97, 98
Dependent variables 17, 19, 20, 139
Descriptive statistics 15, 23, 25, 27,
28, 31, 33, 81, 101, 104, 127, 128

E

Expected frequencies 96, 98
Error 41, 44, 51, 59, 64, 65, 68, 93,
131, 150

F

False negative 54, 56, 139
False positive 54, 56, 139
Fisher's exact test 108
Frequency polygon 28, 29-31
Frequency table 24, 25

G

Gaussian distribution 44
Graphs 30, 31

H

Hazard ratio 182
Histogram 28, 29, 47, 157
Homogeneity 71, 79, 90, 91, 97, 100,
151
Hosmer and Lameshow's goodness of
fit 97-99
Hypothesis testing 5, 15, 48, 51, 52,
55, 60, 70, 87, 108, 111, 142

I

Incidence 68, 104, 175, 179
Independent variable 17, 19, 20, 104,
139, 147, 157
Interquartile range 35, 42



L

Level of significance 55-57, 74, 76, 79, 83, 98
 Levene's test 71, 79, 90
 Linear regression 5, 145, 147, 148, 151-154, 156, 158, 159
 Logrank test 114, 121, 122, 124-126, 192

M

Mann-Whitney-Wilcoxon test 113
 Mantel-Haenzel chi square test 5, 48, 95, 96-105, 108
 Mc Nemar's test 107, 108
 Mean 14, 19, 35, 36, 38, 40, 41, 44-52, 54-58, 62-67, 69-79, 81-83, 85, 90, 91, 93, 94, 114, 148, 151, 170, 186
 Measures of central tendency 5, 35, 36, 38
 Measures of dispersion 35
 Median 35-38, 47-49, 114, 115, 120, 121, 123-127, 135
 Meta-analysis
 Mode 35-38, 47-49

N

Nominal scale 18
 Non parametric analysis 108
 Normal distribution 5, 43-46, 48, 66, 67, 74, 115, 156
 Null hypothesis 51, 53, 55, 56, 59, 72, 86, 88, 98
 Numerical data 23, 24, 28

O

Odds ratio 175, 178
 One-tailed test 54, 57, 85, 88
 Ordinal scale 18

P

Paired *t*-test 71, 82-84, 114
 Parameters 52, 68, 163
 Power of study 59
 Prevalence 172, 175, 179, 189
 Probability 15, 44, 47, 48, 59, 63, 65, 72, 73, 75, 86, 87, 108, 161, 162, 164, 170, 179

P value 68, 71-73, 81, 82, 85, 120, 123, 126, 131, 132-134, 142, 144, 158

Q

Quantitative variables 139, 147
 Quartile 39, 40-42, 118, 121, 125

R

Rate 20, 56, 69, 138, 175-179, 181-187
 Ratio 5, 19, 20, 24, 73, 79, 82, 142, 154, 175, 178, 185
 Regression 142, 145-148, 151-156, 158, 159
 Rejection region 75, 85, 86, 87, 88

S

Sample size 48, 59, 68, 69, 87, 97, 114, 115, 121, 124, 135, 158, 161, 162, 167, 169, 171-174
 Scatter plot 137, 139, 140, 152, 157
 Significance level 59, 90, 115, 121, 124, 127, 142
 Sign test 88
 Simple random sampling 164-170
 Skewness 38, 47, 48
 Slope 148, 152, 153, 155
 Standard deviation 35, 39-41, 45-48, 63-67, 74, 75, 157
 Stem-and-leaf 33

T

t-test 57, 71, 73-85, 90, 114, 121, 123

V

Variance 35, 39, 40, 52, 62, 71, 79, 90, 91, 150, 151, 154, 157

W

Wilcoxon test 113

Y

y-axis 139, 158





AUTHORS



Professor Dr. Nooriah Mohamed Salleh

Qualifications:

1974: MBBS (University Malaya)

1978: MPH (Tulane University, USA)

1985: DrPH (Tulane University, USA)

Area of Speciality: Public Health (Epidemiology & Biostatistics)

She taught Epidemiology to medical and biomedical undergraduates and Master of Public Health candidates at Faculty of Medicine, University Malaya (2001-2004 & 2007). Taught Epidemiology and Biostatistics to medical undergraduates at Universiti Kebangsaan Malaysia (1976-87), UniKL Royal College of Medicine Perak (2005-2006), Master of Public Health candidates at Faculty of Medicine & Allied Health Sciences, University Malaysia Sarawak (2008) and medical students at Faculty of Medicine and Health Sciences, University Science Islam Malaysia (2009 till now). She was the Head of Community Health Department at UKM (1986-1987) and Head of Epidemiology and Biostatistics Unit, Department of Social & Preventive Medicine, Faculty of Medicine at University Malaya (2001-2004 & 2007).



Dr. Zairina Binti A Rahman

Qualifications:

1996: MD (National University of Malaysia)

2002: MPH (University Malaya)

2005: MPH (Occupational) (University Malaya)

Area of Speciality: Public Health (Occupational Health)

After graduating with Master of Public Health in 2005, she worked in Ministry of Health Malaysia as Public Health Physician. She was involved in training health personnel on matters related to public health issues. She joined Universiti Sains Islam Malaysia (USIM) in 2008 as a Medical Lecturer and is currently holding position as Dean of Faculty of Medicine and Health Sciences. She teaches Biostatistics, Epidemiology, Occupational Health, Environmental Health and Research Methodology to students at Faculty of Medicine and Health Sciences and Faculty of Dentistry.



Dr. Mohd Dzulkhairi bin Mohd Rani

Qualifications:

- 2002 : MBBS (International Islamic University of Malaysia)
- 2006 : Cert. in Medical Law (International Islamic University of Malaysia)
- 2010 : M.Comm.Med (Occupational) Health (Universiti Sains Malaysia)

Area of Specialty: Public Health (Occupational Health)

He joined Universiti Sains Islam Malaysia (USIM) in 2010 as a Senior Medical Lecturer. He currently teaches Biostatistics, Epidemiology, Occupational Health, Environmental Health and Research Methodology to medical undergraduates students at Faculty of Medicine and Health Sciences and Faculty of Dentistry.



BASIC STATISTICS FOR MEDICAL AND HEALTH SCIENCES

This book is written by the Community Health lecturers of the Faculty of Medicine and Health Sciences, Universiti Sains Islam Malaysia. The uniqueness of this book is that it guides the readers through the stages of understanding biostatistics and performing statistical analysis using the SPSS software by aids of illustration. This book is simple, practical and easily understood by medical and health sciences students, social sciences students and medical practitioners.

