

THE CAMBRIDGE DICTIONARY OF STATISTICS FOURTH EDITION

If you work with data and need easy access to clear, reliable definitions and explanations of modern statistical and statistics-related concepts, then look no further than this dictionary. Nearly 4000 terms are defined, covering medical, survey, theoretical and applied statistics, including computational and graphical aspects. Entries are provided for standard and specialized statistical software. In addition, short biographies of over 100 important statisticians are given. Definitions provide enough mathematical detail to clarify concepts and give standard formulae when these are helpful. The majority of definitions then give a reference to a book or article where the user can seek further or more specialized information, and many are accompanied by graphical material to aid understanding.

B. S. EVERITT is Professor Emeritus of King's College London. He is the author of almost 60 books on statistics and computing, including *Medical Statistics from A to Z*, also from Cambridge University Press.

A. SKRONDAL is Senior Statistician in the Division of Epidemiology, Norwegian Institute of Public Health and Professor of Biostatistics in the Department of Mathematics, University of Oslo. Previous positions include Professor of Statistics and Director of The Methodology Institute at the London School of Economics.



THE CAMBRIDGE DICTIONARY OF

Statistics

Fourth Edition

B. S. EVERITT

Institute of Psychiatry, King's College London

A. SKRONDAL

Norwegian Institute of Public Health Department of Mathematics, University of Oslo





CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521766999

First, second and third editions © Cambridge University Press 1998, 2002, 2006 Fourth edition © B. S. Everitt and A. Skrondal 2010

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 1998 Reprinted 1999 Second edition published 2002 Reprinted with corrections 2003, 2005 Third edition published 2006 Fourth edition published 2010

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-76699-9 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.



To the memory of my dear sister Iris

B. S. E.

To my children Astrid and Inge

A.S.



Preface to fourth edition

In the fourth edition of this dictionary many new entries have been added reflecting, in particular, the expanding interest in Bayesian statistics, causality and machine learning. There has also been a comprehensive review and, where thought necessary, subsequent revision of existing entries. The number of biographies of important statisticians has been increased by including many from outside the UK and the USA and by the inclusion of entries for those who have died since the publication of the third edition. But perhaps the most significant addition to this edition is that of a co-author, namely Professor Anders Skrondal.

Preface to third edition

In this third edition of the Cambridge Dictionary of Statistics I have added many new entries and taken the opportunity to correct and clarify a number of the previous entries. I have also added biographies of important statisticians whom I overlooked in the first and second editions and, sadly, I have had to include a number of new biographies of statisticians who have died since the publication of the second edition in 2002.

B. S. Everitt, 2005

Preface to first edition

The Cambridge Dictionary of Statistics aims to provide students of statistics, working statisticians and researchers in many disciplines who are users of statistics with relatively concise definitions of statistical terms. All areas of statistics are covered, theoretical, applied, medical, etc., although, as in any dictionary, the choice of which terms to include and which to exclude is likely to reflect some aspects of the compiler's main areas of interest, and I have no illusions that this dictionary is any different. My hope is that the dictionary will provide a useful source of reference for both specialists and non-specialists alike. Many definitions necessarily contain some mathematical formulae and/or nomenclature, others contain none. But the difference in mathematical content and level among the definitions will, with luck, largely reflect the type of reader likely to turn to a particular definition. The non-specialist looking up, for example, **Student's t-tests** will hopefully find the simple formulae and associated written material more than adequate to satisfy their curiosity, while the specialist

vii



seeking a quick reminder about **spline functions** will find the more extensive technical material just what they need.

The dictionary contains approximately 3000 headwords and short biographies of more than 100 important statisticians (fellow statisticians who regard themselves as 'important' but who are *not* included here should note the single common characteristic of those who are). Several forms of cross-referencing are used. Terms in *slanted roman* in an entry appear as separate headwords, although headwords defining relatively commonly occurring terms such as **random variable**, **probability**, **distribution**, **population**, **sample**, etc., are *not* referred to in this way. Some entries simply refer readers to another entry. This may indicate that the terms are synonyms or, alternatively, that the term is more conveniently discussed under another entry. In the latter case the term is printed in *italics* in the main entry.

Entries are in alphabetical order using the letter-by-letter rather than the word-by-word convention. In terms containing numbers or Greek letters, the numbers or corresponding English word are spelt out and alphabetized accordingly. So, for example, 2×2 table is found under **two-by-two table**, and α -trimmed mean, under **alpha-trimmed mean**. Only headings corresponding to names are inverted, so the entry for William Gosset is found under **Gosset**, **William** but there is an entry under **Box–Müller** transformation *not* under **Transformation**, **Box–Müller**.

For those readers seeking more detailed information about a topic, many entries contain either a reference to one or other of the texts listed later, or a more specific reference to a relevant book or journal article. (Entries for software contain the appropriate address.) Additional material is also available in many cases in either the *Encyclopedia of Statistical Sciences*, edited by Kotz and Johnson, or the *Encyclopedia of Biostatistics*, edited by Armitage and Colton, both published by Wiley. Extended biographies of many of the people included in this dictionary can also be found in these two encyclopedias and also in *Leading Personalities in Statistical Sciences* by Johnson and Kotz published in 1997 again by Wiley.

Lastly and paraphrasing Oscar Wilde 'writing one dictionary is suspect, writing two borders on the pathological'. But before readers jump to an obvious conclusion I would like to make it very clear that an anorak has never featured in my wardrobe.

B. S. Everitt, 1998

Acknowledgements

Firstly I would like to thank the many authors who have, unwittingly, provided the basis of a large number of the definitions included in this dictionary through their books and papers. Next thanks are due to many members of the 'allstat' mailing list who helped with references to particular terms. I am also extremely grateful to my colleagues, Dr Sophia Rabe-Hesketh and Dr Sabine Landau, for their careful reading of the text and their numerous helpful suggestions. Lastly I have to thank my secretary, Mrs Harriet Meteyard, for maintaining and typing the many files that contained the material for the dictionary and for her constant reassurance that nothing was lost!



Notation

The transpose of a matrix A is denoted by A'.

Sources

- Altman, D. G. (1991) *Practical Statistics for Medical Research*, Chapman and Hall, London. (SMR)
- Chatfield, C. (2003) *The Analysis of Time Series: An Introduction*, 6th edition, Chapman and Hall, London. (TMS)
- Evans, M., Hastings, N. and Peacock, B. (2000) *Statistical Distributions*, 3rd edition, Wiley, New York. (STD)
- Krzanowski, W. J. and Marriot, F. H. C. (1994) Multivariate Analysis, Part 1, Edward Arnold, London. (MV1)
- Krzanowski, W. J. and Marriot, F. H. C. (1995) Multivariate Analysis, Part 2, Edward Arnold, London. (MV2)
- McCullagh, P. M. and Nelder, J. A. (1989) *Generalized Linear Models*, 2nd edition, Chapman and Hall, London. (GLM)
- Rawlings, J. O., Sastry, G. P. and Dickey, D. A. (2001) *Applied Regression Analysis: A Research Tool*, Springer-Verlag, New York. (ARA)
- Stuart, A. and Ord, K. (1994) *Kendall's Advanced Theory of Statistics, Volume 1, 6th edition*, Edward Arnold, London. (KA1)
- Stuart, A. and Ord, K. (1991) *Kendall's Advanced Theory of Statistics, Volume 2, 5th edition*, Edward Arnold, London. (KA2)