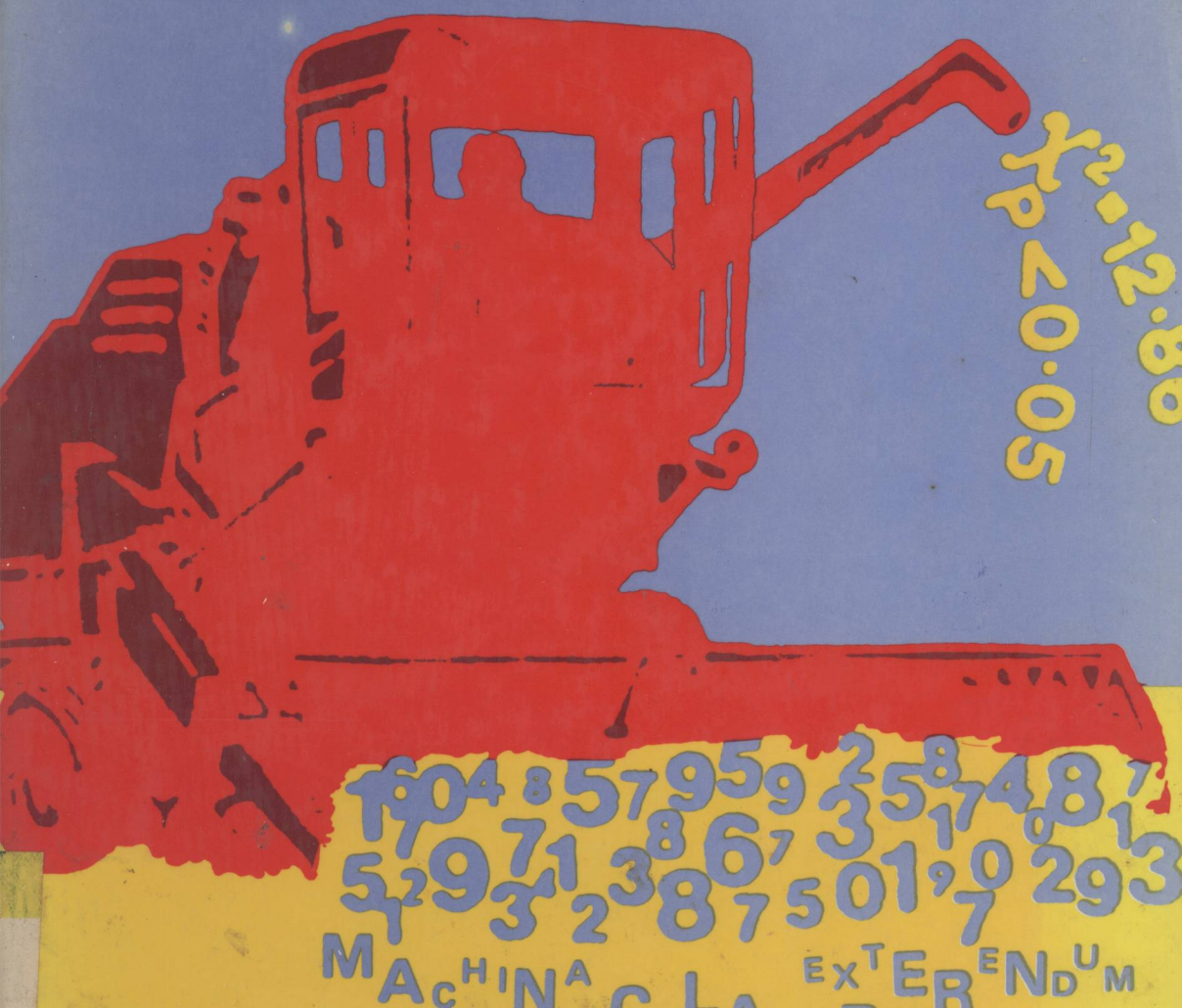
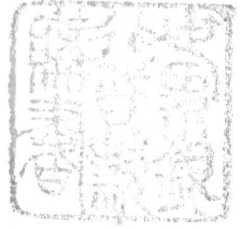


Ellis Horwood Series
MATHEMATICS AND ITS APPLICATIONS

APPLIED STATISTICS ALGORITHMS

editors
P. GRIFFITHS and I.D. HILL





APPLIED STATISTICS ALGORITHMS



ELLIS HORWOOD SERIES IN MATHEMATICS AND ITS APPLICATIONS

Series Editor: Professor G. M. BELL, Chelsea College, University of London

Statistics and Operational Research

Editor: B. W. CONOLLY, Professor of Operational Research, Queen Mary College, University of London

Beaumont, G. P.	Introductory Applied Probability
Beaumont, G. P.	Basic Probability and Random Variables*
Conolly, B. W.	Techniques in Operational Research: Vol. 1, Queueing Systems
Conolly, B. W.	Techniques in Operational Research: Vol. 2, Models, Search, Randomization
French, S.	Sequencing and Scheduling: Mathematics of the Job Shop
French, S.	Decision Theory
Griffiths, P. & Hill, I. D.	Applied Statistics Algorithms
Hartley R.	Linear Methods of Mathematical Programming
Jones, A. J.	Game Theory
Kemp, K. W.	Dice, Data and Decisions: Introductory Statistics
Oliveira-Pinto, F.	Simulation Concepts in Mathematical Modelling
Oliveira-Pinto, F. & Conolly, B. W.	Applicable Mathematics of Non-physical Phenomena
Schendel, U.	Introduction to Numerical Methods for Parallel Computers
Stoodley, K. D. C.	Applied and Computational Statistics A First Course
Stoodley, K. D. C., Lewis, T. & Stainton, C. L. S.	Applied Statistical Techniques
Thomas, L. C.	Games, Theory and Applications
Whitehead, J. R.	The Design and Analysis of Sequential Clinical Trials

*In preparation

0212

G3

8663793



APPLIED STATISTICS ALGORITHMS

Editors:

P. GRIFFITHS, B.Sc., Ph.D.
Oxford University Computing Service

and

I. D. HILL, D.Sc.
Division of Computing and Statistics
Clinical Research Centre of the Medical Research Council
Harrow, Middlesex



E8663793

Published by
ELLIS HORWOOD LIMITED
Publishers · Chichester



for
THE ROYAL STATISTICAL SOCIETY
London



First published in 1985 by

ELLIS HORWOOD LIMITED

Market Cross House, Cooper Street, Chichester, West Sussex, PO19 1EB, England

The publisher's colophon is reproduced from James Gillison's drawing of the ancient Market Cross, Chichester.

Distributors:

Australia, New Zealand, South-east Asia:

Jacaranda-Wiley Ltd., Jacaranda Press,

JOHN WILEY & SONS INC.,

G.P.O. Box 859, Brisbane, Queensland 40001, Australia

Canada:

JOHN WILEY & SONS CANADA LIMITED

22 Worcester Road, Rexdale, Ontario, Canada.

Europe, Africa:

JOHN WILEY & SONS LIMITED

Baffins Lane, Chichester, West Sussex, England.

North and South America and the rest of the world:

Halsted Press: a division of

JOHN WILEY & SONS

605 Third Avenue, New York, N.Y. 10016, U.S.A.

© 1985 Royal Statistical Society/Ellis Horwood Limited

British Library Cataloguing in Publication Data

Griffiths, Paul

Applied statistics algorithms. —

(Ellis Horwood series in mathematics and its applications)

1. Mathematical statistics 2. Algorithms

I. Title II. Hill, I.D. (Ian David), 1926—

001.4'22 QA276

Library of Congress Card No. 85-891

ISBN 0-85312-772-7 (Ellis Horwood Limited)

ISBN 0-470-20184-3 (Halsted Press)

Typeset by Ellis Horwood Limited.

Printed in Great Britain by Unwin Brothers of Woking.

COPYRIGHT NOTICE —

All Rights Reserved

No part of this book may be reproduced or transmitted by any means without the prior permission of the publisher. The only exceptions are for the purposes of review, or as provided for by the Copyright (Photocopying) Act or in order to enter the algorithms herein into a computer program for non-profit-making purposes.

8663793



Table of Contents

Foreword.	11
-------------------	----

Preface	13
-------------------	----

Chapter 1 – Introduction

1.1 The Royal Statistical Society and its Journals	17
1.2 Statistical computing	18
1.3 The publication of algorithms	19
1.4 The Chilton meeting, and the Working Party on Statistical Computing.	20
1.5 Algorithms in <i>Applied Statistics</i>	20
1.6 Languages and standards.	21
1.7 Remarks and corrigenda	22
1.8 Planning this book.	23
References	23

Chapter 2 – The construction and description of algorithms

revised by P. Griffiths, I. D. Hill, J. P. Royston and J. B. Webb

2.0 Introduction.	25
2.1 Submission.	25
2.2 Refereeing	25
2.3 Languages	26
2.4 Introductory text	26
2.5 Layout	28
2.6 Comments	29

2.7	Input/output	29
2.8	Constants and variables.	29
2.9	Precision	29
2.10	Machine-dependence	30
2.11	Error diagnostics	30
2.12	Parentheses	31
2.13	Labels, jumps and ifs	31
2.14	Construction of loops.	32
2.15	Parameters in Algol 60	32
2.16	Common storage and global identifiers	33
2.17	Use of work-space.	33
2.18	Multidimensional arrays	33
2.19	Standard Fortran.	33
2.20	Character handling.	34
2.21	Copyright and disclaimer	34
	References	35

The Algorithms

Algorithm AS 3	— The integral of Student's t -distribution by B. E. Cooper	38
Algorithm AS 5	— The integral of the non-central t -distribution by B. E. Cooper	40
Algorithm AS 6	— Triangular decomposition of a symmetric matrix by M. J. R. Healy	43
Algorithm AS 7	— Inversion of a positive semi-definite symmetric matrix by M. J. R. Healy	46
Algorithm AS 13	— Minimum spanning tree by G. J. S. Ross	49
Algorithm AS 14	— Printing the minimum spanning tree by G. J. S. Ross	53
Algorithm AS 15	— Single linkage cluster analysis by G. J. S. Ross.	56
Algorithm AS 30	— Half-normal plotting by D. N. Sparks	65
Algorithm AS 41	— Updating the sample mean and dispersion matrix by M. R. B. Clarke.	70
Algorithm AS 45	— Histogram plotting by D. N. Sparks.	74
Algorithm AS 47	— Function minimization using a simplex procedure by R. O'Neill	79
Algorithm AS 51	— Log-linear fit for contingency tables by S. J. Haberman	88
Algorithm AS 52	— Calculation of power sums of deviations about the mean by C. C. Spicer	98
Algorithm AS 57	— Printing multidimensional tables by S. J. Haberman	101
Algorithm AS 60	— Latent roots and vectors of a symmetric matrix by D. N. Sparks and A. D. Todd.	110
Algorithm AS 63	— The incomplete beta integral by K. L. Majumder and G. P. Bhattacharjee	117
Algorithm AS 64/AS 109	— Inverse of the incomplete beta function ratio by K. L. Majumder and G. P. Bhattacharjee	121

Algorithm AS 66 – The normal integral by I. D. Hill	126
Algorithm AS 75 – Basic procedures for large, sparse or weighted linear least squares problems by W. Morven Gentleman	130
Algorithm AS 76 – An integral useful in calculating non-central t and bivariate normal probabilities by J. C. Young and Ch. E. Minder	145
Algorithm AS 83 – Complex discrete fast Fourier transform by Donald M. Monro	149
Algorithm AS 91 – The percentage points of the χ^2 distribution by D. J. Best and D. E. Roberts	157
Algorithm AS 97 – Real discrete fast Fourier transform by Donald M. Monro	162
Algorithm AS 99 – Fitting Johnson curves by moments by I. D. Hill, R. Hill and R. L. Holder	171
Algorithm AS 100 – Normal–Johnson and Johnson–normal transformations by I. D. Hill	183
Algorithm AS 111 – The percentage points of the normal distribution by J. D. Beasley and S. G. Springer	188
Algorithm AS 136 – A K -means clustering algorithm by J. A. Hartigan and M. A. Wong	192
Algorithm AS 147 – A simple series for the incomplete gamma integral by Chi-Leung (Andy) Lau	203
Algorithm AS 154 – An algorithm for exact maximum likelihood estimation of autoregressive–moving average models by means of Kalman filtering by G. Gardner, A. C. Harvey and G. D. A. Phillips	206
Algorithm AS 168 – Scale selection and formatting by W. Douglas Stirling	222
Algorithm AS 169 – An improved algorithm for scatter plots by W. Douglas Stirling	230
Algorithm AS 183 – An efficient and portable pseudo-random number generator by B. A. Wichmann and I. D. Hill	238
Algorithm ACM 291 – Logarithm of gamma function by M. C. Pike and I. D. Hill	243
Index of statistical algorithms	247
List of routine names	261
Keyword index	267
Index to Chapters 1 and 2	305



To the memory of
Abu Ja'far Mohammed ben Musa
from whose description as
al-Khwārazmī (the native of Khwārazm)
the word Algorithm is derived

A'LGORISM, } *n. f.* Arabick words, which are used to imply
A'LGORITHM: } the six operations of arithmetick, or the sci-
ence of numbers. *Diſt.*

Dr Samuel Johnson's definition of algorithm has been reproduced in facsimile from his
Dictionary of the English Language, (1755)

Foreword

It gives me great satisfaction to introduce this collection of algorithms from *Applied Statistics*. It comes with a scholarly introduction describing the history of algorithms, which stretches back to a time long before a group of us began the algorithm section of *Applied Statistics* in 1968. We began tentatively, and we made inevitable mistakes, some of which are tactfully alluded to in the book; what I suspect none of us realized at the beginning were the difficulties that beset the production of a good algorithm. It is hard to get the code right, to ensure that the algorithm is efficient and well structured, that it adheres to the language standard, that all kinds of potential misuse are trapped, and that the description is clear and accurate. The whole exercise is iterative and, as the editors show, not even the scrutiny of past editors and referees (to whom all of us, users and authors alike, owe so much) has avoided further scrutiny and amendment in preparing the algorithms for reproduction in book form. (Doubtless some energetic individuals will soon be doing their best to prove that all is still not yet perfect – I predict that they will have a hard job.)

Statisticians have become major users of packages, and packages depend heavily on reliable and tested algorithms. The writer of a one-off program for a particular analysis is equally in need of such modules as the algorithms represent; they may be black boxes but the contents must be of top quality. To the extent that increasing numbers of people are now learning statistics from running programs rather than from reading books, the executable algorithm has become part of a new and powerful form of literature, with its own style and perhaps eventually its own masterpieces. The editors have made their careful selection – can you see which may eventually become classics?

J. A. Nelder
President (1985/86)
The Royal Statistical Society
London

The algorithms contained in this book are available in machine-readable form from the publisher. Please address all enquiries to:

Ellis Horwood Limited, Market Cross House, Cooper Street,
Chichester, England, PO19 1EB

Preface

A police electronics engineer admitted in the Johannesburg Magistrates' Court yesterday that, by applying unerring logic, a winning number could be determined in an electronic computerised card game played in a Hillbrow casino. . . . Col R.D. Hull said the electronic computerised machines have an algo rhythm, which is a series of rhythmic manipulations to determine the outcome of a game. The algo rhythm is a "mathematical recipe for numbers" which is the key to the game. The programmer knows the key.

Johannesburg Star

(Acknowledgement to *Computer Weekly* where we first saw this quoted.)

Applied Statistics, one of the journals published by the Royal Statistical Society, has included an Algorithms Section for the last 16 years, and the total number of algorithms published exceeds 200. The time has come for an anthology of them, which we now present.

Algorithms are the building blocks of computer programs, each designed to perform a specific purpose, and to be slotted into a program at will. At its best an algorithm should be so well tested, known to be efficient and always to give the correct results, that it can be used simply as a 'black box' whose inner workings can be ignored, and need not be understood, provided that one thoroughly understands its purpose, how to feed it with input and how to handle its output. But even the blackest of boxes needs its construction to be fully specified so that further copies of it can be made, and so that experts can suggest improvements. It is the aim of both the journal and this book that the algorithms presented should fulfil all such requirements.

As more fully described in Chapter 1, the book does not simply reproduce what has previously appeared in the journal. Corrections and improvements have been made, some previously published as separate entities but now fully incorporated, others new to the book, if further study has shown them to be desirable.

All who write computing instructions know the fascination of pitting one's wits against the pure logic of a machine that does what you told it, not what you meant to tell it. It is humbling as well as fascinating, and we have been humbled often enough to be aware that our aim, for which we have tried so hard, of a book that is free from 'bugs' has almost certainly not been achieved. We ask readers to let us know of any discovered.

ALGORISMS AND ALGORITHMS

Dr R A Griffiths and Mr W R Good (Radcliffe Infirmary, Oxford OX2 6HE) and Mr J G Griffith (Jesus College, University of Oxford) write: We note the use in a recent series of articles of the neologism 'algorithm'. The *Oxford English Dictionary* describes this as a 'recent pseudo-etymological perversion' in which 'algorism' is learnedly confused with 'ἄριθμός', 'number'. 'Algorism' is, of course, from the Arabic and without Greek derivation. 'Algorithm' could only be a fusion of 'ἄλγος' (pain) and 'ἄριθμός' (number) and would presumably denote 'measurement or counting of pain'.

(*British Medical Journal*, 23 June 1984)

There are many words in the English language that started life as a 'recent pseudo-etymological perversion'. The exact point at which any such word passes the boundary, and becomes correct usage, is difficult to determine, but we have no doubt at all that *algorithm* did so many years ago, and it would be inexcusable affectation to refuse to use it.

Can any word still be a neologism if it appeared in Dr Johnson's dictionary, admittedly not with the current meaning, but with both spellings allowed? He knew only one spelling for 'arithmetick', but Boswell's hope that 'the authority of the great Master of our language will stop that curtailing innovation, by which we see *critic*, *public*, &c., frequently written instead of *critick*, *publick*, &c.' remained unsatisfied. Those who think 'algorithm' an incorrect spelling must therefore regard 'arithmetic' as even more so.

Since Dr Johnson's day 'algorithm' has come to mean a recipe for computing operations. Even if it is this meaning that is objected to as neologistic, it goes back nearly half a century at least, as the Supplement to the *Oxford English Dictionary* shows by quoting a reference to 'Euclid's algorithm' from 1938.

Algorism has never taken on this new meaning but remains defined as 'The Arabic, or decimal, system of numeration'. Why should anyone be expected to take an almost unknown word with the wrong meaning, when a better-known word with the right meaning is available?

UNSELECTED ALGORITHMS

To those algorithm authors whose contributions have not been selected for the book we should like to give assurance that lack of space is the reason. No-one should assume that the absence of a given algorithm implies in any way that it lacks merit.

ACKNOWLEDGEMENTS

We thank not only all who have helped in the preparation of this book, but also all whose labours have contributed to the Algorithms Section of the journal. In particular John Nelder as founder and first editor, Howard Simpson as a subsequent editor, and Patrick Royston and Janet Webb who have now taken on the task, deserve gratitude. Maria de la Hunty, as the Society's Executive Editor (for all its journals) has always been a tower of strength. None of them, nor ourselves when in the editorial chair, could have done much without the selfless efforts of all the anonymous referees who so carefully examine and test all the contributions. We are deeply grateful to all those algorithm authors (and their employers in some instances) who have so willingly agreed to let us reproduce their work, and also to the Association for Computing Machinery for permission to use one algorithm from their *Communications*.

Finally, we thank our good friends the computers, but for whom our subject matter would not exist other than as a speculative dream of science fiction.

P. Griffiths

I. D. Hill

December 1984

