

SELECTED STATISTICAL PAPERS OF SIR DAVID COX.

D.J. Hand and A.M. Herzberg (Eds.).

Volume I: Design of Investigations, Statistical Methods and Applications; Volume II: Foundations of Statistical Inference, Theoretical Statistics, Time Series and Stochastic Processes.

Cambridge University Press, 2005, pp. xi + 591; pp. xi + 590, £180.00/US\$320.00 Set.

Contents:

Volume 1:

1. Design of investigations (10 papers)
2. Statistical methods (25 papers)
3. Applications (6 papers)

Volume 2:

4. Foundations of statistical inference (8 papers)
5. Theoretical statistics (22 papers)
6. Time series (4 papers)
7. Stochastic processes (11 papers)

Readership: Statisticians

Professors Hand and Herzberg should be thanked for the efforts which have culminated with the publication of these two volumes. The work presents eighty-six of David Cox's papers, selected from those published up to the end of 1993. Also included is a complete list of David Cox's publications through 2004.

There can be little doubt that the papers of David Cox provide a rich vein of statistical insight and their availability, in a very convenient form, will be very much appreciated by current and future statistical researchers. And to those who remember the papers, and perhaps their own, now rather dog-eared, reprints, publication in the original format seems particularly appropriate.

The selection process, a joint endeavor of David Cox and the editors, must have been difficult. The choice of 86 papers out of 236 reflects a reasonable set of priorities but will, for many readers and perhaps inevitably, leave a personal favourite or two 'out in the cold'. A two page paper from *Technometrics* in 1965 was one I would have been glad to see included.

What makes this work of special value is David Cox's willingness to contribute commentaries on the papers. With characteristic modesty he describes his famous paper on *Regression models and life tables* as "cited a fairly large number of times although no doubt read rather less often". However the task of providing the commentaries has been taken seriously and provides valuable additional insight and perspective on the papers, as well as some personal recollections. Discussion of the genesis of the paper on *An analysis of transformations*, joint with George Box, provides an interesting case of the latter. However, I must resist the temptation to say more about the commentaries. Do read and enjoy them for yourself. But, don't miss the remark about the paperless office in the Preface!

I think many will be pleased to have these two volumes on their bookshelves.

MRC Biostatistics Unit
Cambridge, U.K.

V.T. Farewell

DATA MINING ET STATISTIQUE DÉCISIONNELLE:

L'INTELLIGENCE DANS LES BASES DU DONNÉE.

S. Tufféry.

Paris: Editions Technip, 2005, pp. xviii + 379.

Table des matières:

1. Panorama du data mining
2. Le déroulement d'une étude de data mining
3. L'exploration et la préparation des données
4. L'utilisation des données commerciales
5. Aperçu sur les techniques de data mining
6. L'analyse factorielle
7. Les réseaux de neurones

8. Les techniques de classification automatique
9. La recherche d'associations
10. Les techniques de classement et de prédiction
11. Une application du data mining: Le scoring
12. Les facteurs de succès d'un projet de data mining
13. Les logiciels de statistique et data mining
14. Le text mining
15. Le web mining

ANNEXE A: Rappels de Statistique

ANNEXE B: Data Mining, Informatique et Libertés

Lecture: Gestionnaires de bases de données

Ce livre de data mining et statistique décisionnelle est un ouvrage de référence pour les gestionnaires de bases de données dans tous les secteurs d'activité. Le 'forage de données' est une discipline moderne qui veut extraire des informations utiles d'une grande base de données en utilisant la statistique et l'informatique. Le livre est surtout basé sur les méthodes classiques mais il y a aussi des méthodes plus modernes comme arbres de décision, réseaux de neurones, support vector machines, algorithmes génétiques, bagging et boosting, etc. Le livre contient beaucoup d'exemples et applications avec les logiciels SAS, SPSS et SPAD. Il y a aussi des annexes intéressantes sur les bases des statistiques et sur les aspects juridiques de l'informatique. Ce livre est utile pour le grand public des enseignants et praticiens dans ce domaine.

Universiteit Hasselt
Diepenbeek, Belgium

N.D.C. Veraverbeke

THE ANALYSIS OF MEANS. A Graphical Method for Comparing Means, Rates, and Proportions.

P.R. Nelson, P.S. Wludyka and K.A.F. Copeland.

Philadelphia: Society for Industrial and Applied Mathematics/Alexandria, Virginia: American Statistical Association, 2005, pp. xii + 247, US\$85.00.

Contents:

1. Introduction
2. One-factor balanced studies
3. One-factor unbalanced studies
4. Testing for equal variances
5. Complete multifactor studies
6. Incomplete multifactor studies
7. Axial mixture designs
8. Heteroscedastic data
9. Distribution-free techniques

APPENDIX A: Figures

APPENDIX B: Tables

APPENDIX C: SAS Examples

Readership: Practitioners who want to compare means arising from a variety of contexts

1. "The analysis of means (ANOM) is a graphical procedure used to quantify differences among treatment groups in a variety of ... situations."

2. "The book is a very practical comprehensive guide to the ANOM procedure. It has widespread application and is written at a level that can be comprehended by those who do not have a background in statistics. The ANOM procedures presented in the book provide the reader a means of communicating data to management, industry regulators, and peers in simple, graphical terms. The procedures combine the power of statistics with the simplicity of basic plotting techniques." (Sheri L. Meredith, Senior Quality Engineer, Vistakon, Division of Johnson & Johnson Vision Care, Inc., Jacksonville, Florida.)

Paragraphs 1 and 2 above are taken from the back of the book. I agree with everything Ms. Meredith says, and it seems pointless for me to rephrase it. I add