Journal of Statistical Software

February 2010, Volume 33, Book Review 2. http://www.jstatsoft.org/

Reviewer: Joakim Ekström, Uppsala University

Computation of Multivariate Normal and t Probabilities

Alan Genz and Frank Bretz
Springer-Verlag, New York, 2009.
ISBN 978-3-642-01688-2. 126 pp. USD 59.95.
http://www.springer.com/978-3-642-01688-2

Computation of Multivariate Normal and t Probabilities is an introductory yet comprehensive book with a self-explanatory title. The book is small in size, a paperback of 126 pages, and is part of Springer's *Lecture Notes in Statistics* series (volume 195), with its familiar page layout, fonts and shades of orange on the cover.

The topic, being of course the numerical evaluation of the multivariate normal (MVN) integral, is both important and relevant. As readers of the Journal of Statistical Software likely are well aware, the normal distribution is the limit distribution of suitably normalized sums of random variables. The family of normally distributed random variables is, moreover, closed under addition, making the multivariate normal distribution ideally suited for linear modeling. Furthermore, this versatile distributional family has the remarkable property that all marginal and conditional distributions are also themselves normal.

Historically, the normal distribution has had a dominant position within the field of multivariate analysis. At least up until year 1913, according to Karl Pearson, the family of normal distributions was the only family of continuous multivariate probability distributions that had been effectively discussed, using Pearson's own wording. Furthermore, Pearson's influential mentor Francis Galton even had a philosophical argument why all variables in nature ought to be normally distributed, underscoring the distributional family's historically prominent position within multivariate analysis, and indeed empirical evidence-based sciences at large.

In despite of the wide use of the multivariate normal distribution for statistical modeling, the fact that the integral of its density function does not exist in closed form has historically posed a practical problem, as well as, it is probably safe to say, a source of annoyance for many statisticians. Nowadays, the inconvenience is all but mitigated by the availability of modern computers. In high dimensions, though, the numerical integration problem still poses a difficult task even for the fastest computers. An understanding of the numerical integration techniques and their potential pitfalls is, moreover, of value for statisticians in and of itself.

When starting to read the book, it quickly becomes clear that the book is well-structured, readable, and that the authors are knowledgeable; in fact 24 of the 270 bibliographical references are works by the authors. Following an interesting, although short, introduction the

authors discuss mathematical relations, identities and expressions that are useful for evaluating the MVN integral. While the text is rich in bibliographic references, the results are discussed briefly and none is explicitly shown. Relationships between the results presented are in general not discussed, other than possibly a brief mention such as, e.g., "the approach is based on Placket's identity". However, the purpose of the book is not to provide a thorough reference work on the subject but rather, in the words of the authors, to give the reader a glimpse into the subject. The matter is clearly structured, the material selection generous and well thought-out, and the bibliographical referencing first-rate.

The authors continue with a discussion on approximations of the integral, by means of inequalities, series expansions and reparametrizations. Thereafter, some numerical integration techniques are discussed. While the text is introductory, the technical nature of the subject, and the brief manner in which it is discussed, will likely give rise to difficulties in understanding for readers that do not have previous familiarity with, e.g., series expansion and numerical integration. Bits of pseudo-code and algorithms are provided, but the manner in which the algorithms are discussed is, again, brief.

The authors provide software implementations via their respective homepages, something which undoubtedly adds value to the book. The implementations in the languages R and MATLAB are presented, respectively, as subsections of the book. The functions are listed with descriptions and references to relevant sections of the book. Instructions and examples on how the functions are called are also provided. Furthermore, the software documentations are reproduced in appendices. However, the presentation of the software has an unpolished feel to it, and the integration with the other sections of the book could be better.

The strengths of the book are many. It is readable, its contents well-structured, the bibliography comprehensive, and the mathematical notation is clear and consistent throughout. The book will serve as a good introduction and bibliographical review, particularly for readers that have some previous familiarity with the subject. Though I hope that the authors, who without question are very knowledgeable, will publish a self-contained and thorough treatment of the subject, because such a work would serve as valuable and lasting reference-book on this important and historically prominent subject of statistics.

Reviewer:

Joakim Ekström
Uppsala University
Department of Statistics
Övre Slottsg 9
753 12 Uppsala, Sweden
E-mail: joakim.ekstrom@statistik.uu.se

Journal of Statistical Software http://www.jstatsoft.org/
published by the American Statistical Association http://www.amstat.org/
Volume 33, Book Review 2 *Published:* 2010-02-02
February 2010