

Frontiers in Probability and the Statistical Sciences

György Terdik

# Multivariate Statistical Methods

Going Beyond the Linear

 Springer

# Frontiers in Probability and the Statistical Sciences

## **Editor-in-Chief**

Somnath Datta, Sch of Public Health & Info Sci, University of Louisville, Louisville, KY, USA

## **Series Editors**

Frederi G. Viens, 1399 Math, Science Building, Purdue University, WEST LAFAYETTE, IN, USA

Dimitris N. Politis, Dept Math, APM 5701, University of California, San Diego, La Jolla, CA, USA

Konstantinos Fokianos, Mathematics & Statistics, University of Cyprus Mathematics & Statistics, Nikosia, Cyprus

Michael Daniels, J.T. Patterson Labs Bldg PAT 141-MC, University of Texas, Austin, USA

The “Frontiers” is a new series of books (edited volumes and monographs) in probability and statistics designed to capture exciting trends in current research as they develop. Some emphasis will be given on novel methodologies that may have interdisciplinary applications in scientific fields such as biology, ecology, economics, environmental sciences, finance, genetics, material sciences, medicine, omics studies and public health.

More information about this series at <http://www.springer.com/series/11957>

György Terdik

# Multivariate Statistical Methods

Going Beyond the Linear

- Vector-Moments and Vector-Cumulants
- Nonlinear Statistics of Normal Multivariates
- Testing Skewness and Kurtosis

György Terdik  
Faculty of Informatics, Dept. of IT  
University of Debrecen  
Debrecen, Hungary

ISSN 2624-9987                      ISSN 2624-9995 (electronic)  
Frontiers in Probability and the Statistical Sciences  
ISBN 978-3-030-81391-8              ISBN 978-3-030-81392-5 (eBook)  
<https://doi.org/10.1007/978-3-030-81392-5>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2021

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG.  
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*To  
Judit,  
and  
Bori and Bálint and Ábel*

# Foreword

As the author rightly points out, much of what we call linear modeling and inference in multivariate statistics is closely tied to the multivariate Gaussian distribution and allied topics. Indeed, the classic text by Professor C.R. Rao is aptly titled *Linear Statistical Inference* (1973, Wiley) as is one of the books recently co-authored by me which is titled *Linear Models and Regression with R* (2020, World Scientific Press). Non-Gaussian multivariate distributions are characterized by nonlinearity, where the cumulants play a prominent role. This monograph is a very deep and substantial piece of work analyzing the cumulants and related statistical measures like the skewness and kurtosis for the most commonly discussed non-Gaussian multivariate distributions. This is also a great place to learn about and polish mathematical prerequisites like multilinear algebra and tensor products for higher-order partial derivatives of vector-valued functions. A large number of exercises at the end of each chapter is an added bonus for those who want to use this as a text for an advanced course. The extensive set of nearly two hundred references, which range from a manuscript that goes back to 1931 to the most recent papers appearing in 2020 and 2021, demonstrates the breadth and depth of coverage.

This monograph will be a standard reference for those who are interested in multivariate distribution theory that goes beyond the multidimensional normal, and will remain so for many years to come!

Santa Barbara, CA, USA  
February 25, 2021

S. Rao Jammalamadaka  
Distinguished Professor,  
Dept. of Statistics and Applied Probability  
University of California, Santa Barbara

# Preface

Linear theory in statistics is closely connected to the normal distribution. One of the main reasons for this is that the best predictor is linear when the random variables are jointly normal, and an important characterization of the normal distribution is that all cumulants of order higher than 2 are zero. The linear theory is pretty much well defined, unlike the nonlinear theory, which can proceed in many different ways. A careful study of the cumulants is a necessary and typical part of nonlinear statistics. Such a study of cumulants for multivariate distributions is made complicated by the index notations. One solution to this problem is the usage of tensor analysis. In this book we offer an alternate method, which we believe is simpler to follow. The higher-order cumulants with the same degree for a multivariate vector can be collected together and kept as a vector. To be able to do so, we introduce a particular differential operator on a multivariate function, called the  $T$ -derivative, and use it to obtain cumulants and provide results which are somewhat analogous to well-known results in the univariate case. We demonstrate this technique through the characterization of several multivariate distributions via their cumulants and by extending the discussion to statistical inference for multivariate skewness and kurtosis.

The book is organized as follows:

Chapter 1 introduces some basic notions and methods which are used in permutations, matrix theory, multilinear algebra, and set partitions.

Chapter 2 deals with the method of tensor products for the higher-order partial derivatives of vector-valued functions. Faà di Bruno's formula is also discussed here.

Chapter 3 concerns the basic theory of  $T$ -moments and  $T$ -cumulants. Besides connections between cumulants and moments, there are results which relate the cumulants of products to products of cumulants, conditional cumulants, etc.

Chapter 4 covers the elementary theory of the nonlinear Hilbert space of Gaussian variates. Multivariate vector-valued Hermite polynomials are introduced with their basic properties and their moments and cumulants are derived.

Chapters 5 and 6 deal with various applications of the material of the previous chapters. In particular, Chap. 5 deals with the cumulants of multivariate skew-distributions, including skew-normal, skew-spherical, skew-t, scale mixtures of



skew normal, skew-normal-Cauchy, and Laplace distributions. The final Chap. 6 is devoted to statistical inference for multivariate skewness and kurtosis. We study the estimators of higher-order cumulant vectors and their asymptotic normality. Explicit formulae for the asymptotic covariances of estimated skewness and kurtosis are given.

The results presented in the book can be followed rather easily because specific references to the formulae that have been used previously are given step by step. Nevertheless, a close reading of the core Chaps. 3 and 4 is required for a proper understanding.

The exercises are of service to practising the methods; some of them are alluded to as simple facts in the text but left unproved. Selected solutions, some useful formulae, as well as basic notations are listed in the Appendix of the book under the titles Solutions, Formulae, and Notations, respectively.

The book is designed both as text for advanced graduate-level courses in multivariate statistical analysis and as a reference book for research workers interested in this area. The R package MultiStatM is available for applications of the theory considered in this book.

During the summer of 2001, when Professor Tata Subba Rao visited the University of Debrecen, we discussed the idea and agreed that statisticians need to pay more attention to cumulants in general. We proposed to write a book on the elementary treatment of cumulants, but unfortunately that work was never continued. In the present book, we use some of the ideas and even some material on cumulants for scalar-valued variates that we discussed at that time.

It is a great pleasure to thank Professor Sreenivasa Rao Jammalamadaka at the University of California, Santa Barbara, and Professor Emanuele Taufer at the University of Trento for numerous discussions and joint work concerning multivariate skewness and kurtosis.

Special thanks go to Professor Sreenivasa Rao Jammalamadaka for a careful reading of the manuscript and useful suggestions which improved the presentation of the book.

Debrecen, Hungary  
January 2021

György Terdik

# Contents

<b>1</b>	<b>Some Introductory Algebra</b> .....	1
1.1	Permutations .....	1
1.2	Tensor Product, <i>vec</i> Operator, and Commutation .....	5
1.2.1	Tensor Product .....	5
1.2.2	The <i>vec</i> Operator .....	6
1.2.3	Commutation Matrices .....	7
1.2.4	Commuting T-Products of Vectors .....	10
1.3	Symmetrization and Multilinear Algebra .....	13
1.3.1	Symmetrization .....	13
1.3.2	Multi-Indexing, Elimination, and Duplication .....	16
1.4	Partitions and Diagrams .....	26
1.4.1	Generating all Partitions .....	29
1.4.2	The Number of All Partitions .....	30
1.4.3	Canonical Partitions .....	34
1.4.4	Partitions and Permutations .....	35
1.4.5	Partitions with Lattice Structure .....	37
1.4.6	Indecomposable Partitions .....	38
1.4.7	Alternative Ways of Checking Indecomposability .....	40
1.4.8	Diagrams .....	43
1.5	Appendix .....	49
1.5.1	Proof of Lemma 1.1 .....	49
1.5.2	Proof of Lemma 1.3 .....	50
1.5.3	Proof of Lemma 1.5 .....	50
1.5.4	Star Product .....	51
1.6	Exercises .....	52
1.7	Bibliographic Notes .....	59
<b>2</b>	<b>The Tensor Derivative of Vector Functions</b> .....	61
2.1	Derivatives of Composite Functions .....	61
2.1.1	Faà di Bruno's Formula .....	63
2.1.2	Mixed Higher-Order Derivatives .....	67

2.2	T-derivative.....	70
2.2.1	Differentials and Derivatives.....	70
2.2.2	The Operator of T-derivative.....	72
2.2.3	Basic Rules.....	74
2.2.4	T-derivative of T-products.....	77
2.2.5	Taylor Series Expansion.....	88
2.3	Multi-Variable Faà di Bruno's Formula.....	90
2.4	Appendix.....	98
2.4.1	Proof of Faà di Bruno's Lemma.....	98
2.4.2	Proof of Faà di Bruno's T-formula.....	99
2.4.3	Moment Commutators.....	100
2.5	Exercises.....	101
2.6	Bibliographic Notes.....	106
<b>3</b>	<b>T-Moments and T-Cumulants.....</b>	<b>107</b>
3.1	Multiple Moments.....	107
3.2	Tensor Moments.....	110
3.3	Cumulants for Multiple Variables.....	118
3.3.1	Definition of Cumulants.....	118
3.3.2	Definition of T-cumulants.....	120
3.3.3	Basic Properties.....	124
3.4	Expressions between Moments and Cumulants.....	126
3.4.1	Expression for Cumulants via Moments.....	126
3.4.2	Expressions for Moments via Cumulants.....	138
3.4.3	Expression of the Cumulant of Products via Products of Cumulants.....	143
3.5	Additional Matters.....	149
3.5.1	Expressions of Moments and Cumulants via Preceding Moments and Cumulants.....	149
3.5.2	Cumulants and Fourier Transform.....	151
3.5.3	Conditional Cumulants.....	154
3.5.4	Cumulants of the Log-likelihood Function.....	161
3.6	Appendix.....	167
3.6.1	Proof of Lemma 3.6 and Theorem 3.7.....	167
3.6.2	A Hint for Proof of Lemma 3.8.....	170
3.6.3	Proof of Lemma 3.2.....	172
3.6.4	Proof of Lemma 3.5.....	173
3.7	Exercises.....	174
3.8	Bibliographic Notes.....	180
<b>4</b>	<b>Gaussian Systems, T-Hermite Polynomials, Moments, and Cumulants.....</b>	<b>183</b>
4.1	Hermite Polynomials in One Variable.....	183
4.2	Hermite Polynomials of Several Variables.....	185
4.3	Moments and Cumulants for Gaussian Systems.....	190

4.3.1	Moments of Gaussian Systems and Hermite Polynomials .....	190
4.3.2	Cumulants for Product of Gaussian Variates and Hermite Polynomials .....	193
4.4	Products of Hermite Polynomials, Linearization .....	197
4.5	T-Hermite Polynomials .....	202
4.6	Moments, Cumulants, and Linearization .....	216
4.6.1	Cumulants for T-Hermite Polynomials .....	220
4.6.2	Products for T-Hermite Polynomials .....	223
4.7	Gram–Charlier Expansion .....	226
4.8	Appendix .....	232
4.8.1	Proof of Theorem 4.2 .....	232
4.8.2	Proof of (4.79) .....	233
4.9	Exercises .....	234
4.10	Bibliographic Notes .....	238
<b>5</b>	<b>Multivariate Skew Distributions</b> .....	<b>241</b>
5.1	The Multivariate Skew-Normal Distribution .....	241
5.1.1	The Inverse Mill’s Ratio and the Central Folded Normal Distribution .....	242
5.1.2	Skew-Normal Random Variates .....	244
5.1.3	Canonical Fundamental Skew-Normal (CFUSN) Distribution .....	247
5.2	Elliptically Symmetric and Skew-Spherical Distributions .....	252
5.2.1	Elliptically Contoured Distributions .....	253
5.2.2	Multivariate Moments and Cumulants .....	261
5.2.3	Canonical Fundamental Skew-Spherical Distribution .....	266
5.3	Multivariate Skew-t Distribution .....	275
5.3.1	Multivariate t-Distribution .....	275
5.3.2	Skew-t Distribution .....	277
5.3.3	Higher-Order Cumulants of Skew-t Distributions .....	278
5.4	Scale Mixtures of Skew-Normal Distribution .....	285
5.5	Multivariate Skew-Normal-Cauchy Distribution .....	287
5.5.1	Moments of $h( Z )$ .....	292
5.6	Multivariate Laplace .....	294
5.7	Appendix .....	297
5.7.1	Spherically Symmetric Distribution .....	297
5.7.2	T-Derivative of an Inner Product .....	303
5.7.3	Proof of (5.44) .....	304
5.7.4	Proof of Lemma 5.6 .....	306
5.8	Exercises .....	307
5.9	Bibliographic Notes .....	310
<b>6</b>	<b>Multivariate Skewness and Kurtosis</b> .....	<b>313</b>
6.1	Multivariate Skewness of Random Vectors .....	313
6.2	Multivariate Kurtosis of Random Vectors .....	321

6.3	Indices Based on Distinct Elements of Cumulant Vectors .....	327
6.4	Testing Multivariate Skewness .....	328
6.4.1	Estimation of Skewness .....	329
6.4.2	Testing Zero Skewness .....	333
6.5	Testing Multivariate Kurtosis .....	337
6.5.1	Estimation of Kurtosis .....	338
6.5.2	Testing Zero Kurtosis .....	341
6.6	A Simulation Study .....	343
6.7	Appendix .....	346
6.7.1	Estimated Hermite Polynomials .....	346
6.8	Exercises .....	347
6.9	Bibliographic Notes .....	348
<b>A</b>	<b>Formulae</b> .....	<b>351</b>
A.1	Bell Polynomials .....	351
A.1.1	Incomplete (Partial) Bell Polynomials .....	351
A.1.2	Bell Polynomials .....	352
A.2	Commutators .....	353
A.2.1	Moment Commutators .....	353
A.2.2	Commutators Connected to T-Hermite Polynomials .....	356
A.3	Derivatives of Composite Functions .....	359
A.4	Moments, Cumulants .....	361
A.4.1	T-Moments, T-Cumulants .....	361
A.5	Hermite Polynomials .....	363
A.5.1	Product of Hermite Polynomials .....	363
A.5.2	T-Hermite Polynomials .....	365
A.6	Function G .....	368
A.6.1	Moments, Cumulants for Skew-t Generator R .....	370
A.6.2	Moments of Beta Powers .....	374
A.7	Complementary Error Function .....	376
A.8	Derivatives of i-Mill's Ratio .....	378
	<b>Notations</b> .....	<b>381</b>
	<b>Solutions</b> .....	<b>385</b>
	<b>References</b> .....	<b>409</b>
	<b>Index</b> .....	<b>417</b>