

SPRINGER BRIEFS IN STATISTICS
JSS RESEARCH SERIES IN STATISTICS

Yasunori Fujikoshi
Vladimir V. Ulyanov

Non-Asymptotic Analysis of Approximations for Multivariate Statistics



 Springer

SpringerBriefs in Statistics

JSS Research Series in Statistics

Editors-in-Chief

Naoto Kunitomo, Economics, Meiji University, Chiyoda-ku, Tokyo, Tokyo, Japan

Akimichi Takemura, The Center for Data Science Education and Research, Shiga University, Bunkyo-ku, Tokyo, Japan

Series Editors

Genshiro Kitagawa, Meiji Institute for Advanced Study of Mathematical Sciences, Nakano-ku, Tokyo, Japan

Shigeyuki Matsui, Graduate School of Medicine, Nagoya University, Nagoya, Aichi, Japan

Manabu Iwasaki, School of Data Science, Yokohama City University, Yokohama, Tokyo, Japan

Yasuhiro Omori, Graduate School of Economics, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

Masafumi Akahira, Institute of Mathematics, University of Tsukuba, Tsukuba, Ibaraki, Japan

Masanobu Taniguchi, Department of Mathematical Sciences/School, Waseda University/Science & Engineering, Shinjuku-ku, Japan

Hiroe Tsubaki, The Institute of Statistical Mathematics, Tachikawa, Tokyo, Japan

Satoshi Hattori, Faculty of Medicine, Osaka University, Suita, Osaka, Japan

Kosuke Oya, School of Economics, Osaka University, Toyonaka, Osaka, Japan

Yasunori Fujikoshi · Vladimir V. Ulyanov

Non-Asymptotic Analysis of Approximations for Multivariate Statistics

 Springer

Preface

This book provides readers with recent non-asymptotic results for approximations in multivariate statistical analysis. There are many traditional multivariate methods based on large-sample approximations. Furthermore, in recent years more high-dimensional multivariate methods have been proposed and utilized for cases where the dimension p of observations is comparable with the sample size n or even exceeds it. Related to this, there are also many approximations under high-dimensional frameworks when $p/n \rightarrow c \in (0, 1)$ or $(0, \infty)$.

An important problem related to multivariate approximations concerns their errors. Most results contain only so-called order estimates. However, such error estimates do not provide information on actual errors for given values of n , p , and other parameters. Ideally, we need non-asymptotic or computable error bounds that relate to these actual errors, in addition to order estimates. In non-asymptotic bounds, the pair (n, p) , as well as other problem parameters, are viewed as fixed, and statistical statements such as tail or concentration probabilities of test statistics and estimators are constructed as a function of them. In other words, these results are applied for actual values of (n, p) . In general, non-asymptotic error bounds involve an absolute constant. If the absolute constant is known, then such an error bound is called the computable error bound.

Our book focuses on non-asymptotic bounds for high-dimensional and large-sample approximations. A brief explanation of non-asymptotic bounds is given in Chap. 1. Some commonly used notations are also explained in Chap. 1. Chapters 2–6 deal with computable error bounds. In Chap. 2, the authors consider computable error bounds on scale-mixed variables. The results can be applied to asymptotic approximations of t - and F -distributions, and to various estimators. In Chap. 3, error bounds for MANOVA tests are given based on large-sample results for multivariate scale mixtures. High-dimensional results are also given. In Chap. 4, the focus is on linear and quadratic discriminant contexts, with error bounds for location and scale mixture variables. In Chaps. 5 and 6, computable error bounds for Cornish–Fisher expansions and \mathcal{A} -statistics are considered, respectively.

Next, in Chaps. 7–11, new directions of research on non-asymptotic bounds are discussed. In Chap. 7, the focus is on high-dimensional approximations for bootstrap procedures in principal component analysis. Then, in Chap. 8 we consider the Kolmogorov distance between the probabilities of two Gaussian elements to hit a ball in Hilbert space. In Chap. 9, the focus is on approximations of statistics based on observations with random sample sizes. In Chap. 10, the topic is large-sample approximations of power-divergence statistics including the Pearson chi-squared statistic, the Freeman–Tukey statistics, and the log-likelihood ratio statistic. Finally, Chap. 11 proposes a general approach for constructing non-asymptotic estimates and provides relevant examples for several complex statistics.

This book is intended to be used as a reference for researchers interested in asymptotic approximations in multivariate statistical analysis contexts. It will also be useful for instructors and students of graduate-level courses as it covers important foundations and methods of multivariate analysis.

For many approximations, detailed derivations would require a lot of space. For the sake of brevity and presentation, we therefore mainly give their outline. We believe and hope that the book will be useful for stimulating future developments in non-asymptotic analysis of multivariate approximations.

We are very grateful to our colleagues R. Shimizu, F. Götze, G. Christoph, V. Spokoiny, H. Wakaki and A. Naumov to be our co-authors for years. Our joint works are widely used in the book. We express our sincere gratitude to Prof. Naoto Kunitomo, Meiji University, Tokyo, and Dr. Tetsuro Sakurai, Suwa University of Science, Nagano, for their valuable comments on various aspects of the content of this book. We are also grateful to Mr. Y. Hirachi for his assistance in the preparation of this book.

The research was partially supported by the Ministry of Education, Science, Sports, and Culture through a Grant-in-Aid for Scientific Research (C), 16K00047, 2016–2018, and results of Chapters 8 and 9 have been obtained under support of the RSF Grant No. 18-11-00132. The study was done within the framework of the Moscow Center for Fundamental and Applied Mathematics, Moscow State University and HSE University Basic Research Programs.

Hiroshima, Japan
Moscow, Russia
April 2020

Yasunori Fujikoshi
Vladimir V. Ulyanov

Contents

1	Non-Asymptotic Bounds	1
1.1	Errors and Non-Asymptotic Bounds	1
	References	4
2	Scale-Mixed Distributions	5
2.1	Introduction	5
2.2	Error Bounds in Sup-Norm	6
2.3	Special Cases	10
2.3.1	Scale-Mixed Normal	10
2.3.2	Scale-Mixed Gamma	13
2.3.3	Scale-Mixed F	16
2.4	Error Bounds Evaluated in L_1 -Norm	17
	References	21
3	MANOVA Test Statistics	23
3.1	Introduction	23
3.2	Multivariate Scale Mixtures for T_{LH} and T_{LR}	25
3.3	Error Bounds for Approximations of T_{LH} and T_{LR}	28
3.4	Error Bound for T_{BNP}	29
3.5	Error Bounds for High-Dimensional Approximations	31
	References	32
4	Linear and Quadratic Discriminant Functions	35
4.1	Introduction	35
4.2	Location and Scale Mixture Expression for EPMC	36
4.3	General Approximation and Error Bounds	38
4.4	Error Bounds for EPMC	43
4.5	Some Related Topics	45
	References	46

5	Cornish–Fisher Expansions	49
5.1	Introduction	49
5.2	Cornish–Fisher Expansion	50
5.3	Error Bounds for Cornish–Fisher Expansion	52
5.4	Proofs for Error Bounds	54
5.5	Transformations for Improved Approximations	55
5.6	Examples	57
	References	58
6	Likelihood Ratio Tests with Box-Type Moments	61
6.1	Introduction	61
6.2	Large-Sample Asymptotic Expansions	62
6.3	High-Dimensional Asymptotic Expansions	63
6.4	Error Bound	66
	References	70
7	Bootstrap Confidence Sets	73
7.1	Introduction	73
7.2	Bootstrap Procedure	76
7.3	Confidence Sets for Spectral Projectors: Bootstrap Validity	77
	References	80
8	Gaussian Comparison and Anti-concentration	81
8.1	Introduction	81
8.2	Motivation: Prior Impact in Linear Gaussian Modeling	82
8.3	Gaussian Comparison	84
8.4	Anti-concentration Inequality	86
8.5	Proofs	87
	8.5.1 Proof of Theorem 8.1	88
	8.5.2 Proof of Lemma 8.1	89
	References	90
9	Approximations for Statistics Based on Random Sample Sizes	93
9.1	Introduction	93
9.2	Notation and Examples	94
9.3	Two Transfer Propositions	97
9.4	Edgeworth and Cornish–Fisher Expansions with Student’s Limit Distribution	98
9.5	Edgeworth and Cornish–Fisher Expansions with Laplace Limit Distribution	102
	References	106
10	Power-Divergence Statistics	109
10.1	Introduction	109
10.2	Rates of Convergence	110

- 10.3 Refinements of Convergence Rates 112
- 10.4 Approximating the Number of Integer Points
in Convex Sets 114
- References 115
- 11 General Approach to Constructing Non-Asymptotic Bounds 117**
- 11.1 Introduction 117
 - 11.1.1 Notation 119
- 11.2 Results for Symmetric Functions 119
 - 11.2.1 Limit Theorem with Bound for Remainder Term 120
 - 11.2.2 Ideas of the Proof of Theorem 11.1 120
 - 11.2.3 Asymptotic Expansions 121
- 11.3 Applications in Probability and Statistics 123
 - 11.3.1 Expansion in the Central Limit Theorem
for Weighted Sums 123
 - 11.3.2 Expansion in the Free Central Limit Theorem 126
 - 11.3.3 Expansion of Quadratic von Mises Statistics 126
 - 11.3.4 Expansions for Weighted One-Sided
Kolmogorov–Smirnov Statistics 129
- References 130