

Mengxi Yi
Klaus Nordhausen *Editors*

Robust and Multivariate Statistical Methods

Festschrift in Honor of David E. Tyler

 Springer

Robust and Multivariate Statistical Methods



David E Tyler 1986 in Paris and 2015 in Myrtle Beach Published with the kind permission of © Coleen Tyler, 2022. All Rights Reserved.

Mengxi Yi • Klaus Nordhausen
Editors

Robust and Multivariate Statistical Methods

Festschrift in Honor of David E. Tyler

 Springer

Editors

Mengxi Yi
School of Statistics
Beijing Normal University
Beijing, China

Klaus Nordhausen
Department of Mathematics and Statistics
University of Jyväskylä
Jyväskylä, Finland

ISBN 978-3-031-22686-1 ISBN 978-3-031-22687-8 (eBook)
<https://doi.org/10.1007/978-3-031-22687-8>

Mathematics Subject Classification: 62Hxx, 62F35, 62E20, 62J07

© The Editor(s) (if applicable) and The Author(s), under exclusive licence to Springer Nature Switzerland AG 2023

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

Foreword

It is an honor and a pleasure to contribute this biographical note to the Festschrift dedicated to David (Dave) Tyler on the occasion of his pending retirement from Rutgers University.

I met Dave in the late 1980s through our mutual friend and colleague Javier Cabrera. Our frequent conversations regarding statistics and robustness led to the idea that workshops in robust statistics would be beneficial to the rapidly changing field of statistics. As a result, our first conference on *Robustness and Data Analysis* was held at Princeton University in 1994 with an outstanding list of invited speakers that included John Tukey and Frank Hampel. The success of this first conference encouraged us to continue with international meetings over more than a decade, and Dave's contributions were vital to our creation of the International Conference on Robust Statistics (ICORS) workshops. Today, with the effort and guidance of a younger generation of statisticians, these ICORS workshops continue to thrive.

But Dave's involvement in these conferences went well beyond the usual organizational stage. His deep understanding of statistical issues and his conviction that robust statistics is not merely a subfield of statistics, but rather a school of thought motivated by the realities of data analysis, provided a clear and firm foundation for these meetings. In every conference, Dave interacted extensively with the participants by exchanging ideas and engaging in discussions for the future direction of robust statistics as well as by stressing the importance to further integrate the concept of robustness into the everyday practice of data analysis.

Dave's insight and depth of understanding regarding statistical issues is attested by his many high-quality research publications. From his first paper, "Asymptotic theory of eigenvectors," published in the *Annals of Statistics* in 1981, to his latest research work on robust covariance matrices, his broad knowledge and independent approach are reflected in the long list of research papers published in the most prestigious statistical journals. Taken together, these papers illustrate the wide range of Dave's interests and his continuing influence in statistics, particularly in the areas of multivariate statistics and robustness. His international academic reputation in these areas is evidenced by his countless invitations to conferences, seminars, and special courses at major academic institutions worldwide. His distinguished

academic credentials led to appointments of associate editor at some of the most prestigious statistical journals, including the *Annals of Statistics*, the *Journal of the Royal Statistical Society B*, and the *Journal of Multivariate Analysis*.

Dave's supportive and collaborative attitude toward his colleagues and his deep statistical knowledge earned him the respect and appreciation of major international researchers in statistics, as is documented by his joint publications with distinguished scholars around the world. In addition, his generosity with his ideas is reflected in the numerous doctoral students that he inspired and supervised with their PhD dissertations. The contributors of this Festschrift, many of whom had the privilege to have worked with Dave and to have benefitted from his knowledge as well as his company, enthusiastically offered their research papers for this volume, and this constitutes a testimony of their appreciation and admiration for his stature as a scholar and for him as a person.

Dave's childhood and adolescence merit comment. From the third oldest child of an impoverished family with ten children to a PhD in statistics from Princeton University to a Distinguished Professorship at Rutgers University, Dave's life and career have been quite unusual. He is the only member of his family to have achieved an advanced university degree.

Born and raised in Pittsburgh, Pennsylvania, Dave spent his early school years in Catholic schools, where he developed an early interest in mathematics. When he was 11 years old, due to family circumstances, he and his siblings were sent to a Catholic orphanage, and he remained there for two years. While poverty was a defining state, his mother provided a stabilizing influence in the family. After graduating from an urban public high school, Dave was admitted to Indiana University of Pennsylvania, where he earned a BA in mathematics in 1972 and where later, in 2015, he was awarded a Distinguished Alumni Award for his career accomplishments. In 1972, he continued with graduate studies at the University of Massachusetts, Amherst, and earned a Master's degree from the Mathematics Department in 1974. That same year, at age 24, he married Coleen McCullough, an aspiring young artist of similar religious and social background. Dave then pursued a doctoral program at Princeton University, where in 1979 he was awarded a PhD in statistics. After Princeton, he served as Assistant Professor at the University of Florida (1978,1979) and Old Dominion University (1979–1983). Finally, in 1983 he joined the Statistics Department of Rutgers University, where he was named Distinguished Professor of Statistics in 2004.

Among Dave's striking personal characteristics are modesty, humanity, and total honesty. This was evident not only during my work with him, but also in the social setting, where I met Coleen, an accomplished artist with whom I had a lasting friendship, and his son, Ed. The interaction between our families made me appreciate Dave's human dimension in addition to his outstanding scholarship. Moreover, firmly bound to his modest origins but dedicated to the field of statistics, Dave complemented his colorful personality with numerous interests and activities such as swimming, basketball, chess, hiking, biking, and boating, among others.

On the occasion of his pending retirement from Rutgers University, the institution where he spent most of his career, I wish Dave many more productive years and I look forward to enjoying the pleasure of his professional and personal company for many more years to come.

Philadelphia, PA, USA
June 2022

Luisa Fernholz

Preface

We are honoured and delighted to edit this Festschrift dedicated to David (Dave) E. Tyler, Distinguished Professor of Statistics at Rutgers University. The idea for this Festschrift was born around the occasion of Dave's 70th birthday and coming retirement from Rutgers to celebrate his outstanding career with many significant contributions to the field of statistics, especially in the areas of multivariate and robust methods.

Dave has a remarkable research career, which he started in 1978, after obtaining his PhD from Princeton University, as an assistant professor at the University of Florida. Via the Old Dominion University, he came in 1983 to Rutgers University, where he currently is a distinguished professor in statistics. In 1994, Dave was elected as an IMS fellow for his distinctive contributions in statistics regarding his independent work on M-estimation of scatter. In particular, most of his work was supported by various grants from, e.g. National Science Foundation (NSF). Dave has a reliable intuition and ability to identify interesting and challenging research questions which are of general importance and relevance. Then he develops his ideas in an insightful as well as rigorous manner addressing all possible details. His attention to detail, while keeping an eye on the big picture and the relevant questions, has been passed on to early career stage researchers, with whom he collaborated and mentored. It is therefore also no surprise that all seven PhD students of Dave embarked on their careers in academia, most of whom are now associate and full professors at universities around the world.

Contributed by Dave's students, friends, coauthors and colleagues, this book includes 22 peer-reviewed papers. The topics of the contributions are mainly motivated by the research interests of Dave. Accordingly, the book consists of four parts. Part **I** begins with an analysis of Dave's publication and coauthor networks, followed by a review article on Dave's famous *Tyler's shape estimator*. Parts **II** and **III**, as the main body of this book, cover some recent advances in multivariate and robust methods. The final part, Part **IV**, includes some various other topics such as supervised learning and normal extremes.

Speaking of these cutting-edge articles, we would like to express our gratitude to the efforts and patience of all contributors in the publishing process, especially

because of the Covid-19 pandemic that disrupted most contributors' routine of work. Despite of those disruptions, upon joint work of authors and referees, we have reached a milestone with very interesting papers. We would like to thank therefore all contributors, who submitted their original and high-quality work to this Festschrift for Dave, and the referees, without whose generous help we would not have made it in time, given the tight schedule. We would like to thank also Veronika Rosteck and Daniel Ignatius from Springer who provided help and assistance whenever needed.

Finally, we want to salute Dave again for his intellectual contributions as well as his help as a mentor and as a friend. May Dave stay healthy and continue advancing the knowledge and boundaries of statistics!

Beijing, China
Jyväskylä, Finland
July 2022

Mengxi Yi
Klaus Nordhausen

Acknowledgements

The Editors would like to thank all the following referees for their excellent work:

Suchin Aeron	Andreas Alfons	Aurore Archimbaud
Andreas Artemiou	Marco Avella	James O. Berger
Germain van Bever	Ana Bianco	Eva Cantoni
Raymond Carroll	Raymond Chambers	Christophe Croux
Lutz Dümbgen	Kamila Facevicova	Hank Flury
Dominique Fourdeinier	Marc Hallin	Zifei Han
Jana Jureckova	Shogo Kato	John Kent
Arnab Kumar Laha	Gaorong Li	Chuan Liu
Markus Matilainen	Peter McCullagh	Jaakko Nevalainen
Hannu Oja	Esa Ollila	Davy Paindaveine
Una Radojicic	Peter Rousseeuw	Gerta Rucker
Marcelo Ruiz	Anne Ruiz-Gazen	Richard L. Smith
Sara Taskinen	Beavers Traymon	Daniel Vogel
Changxi Wang	Christian Weiss	Ami Wiesel
Ines Wilms	Victor Yohai	Teng Zhang
Ting Zhang	Guang Zhu	

Contents

Part I About David E. Tyler's Publications

An Analysis of David E. Tyler's Publication and Coauthor Network	3
Daniel Fischer, Klaus Nordhausen, and Mengxi Yi	
A Review of Tyler's Shape Matrix and Its Extensions	23
Sara Taskinen, Gabriel Frahm, Klaus Nordhausen, and Hannu Oja	

Part II Multivariate Theory and Methods

On the Asymptotic Behavior of the Leading Eigenvector of Tyler's Shape Estimator Under Weak Identifiability	45
Davy Paindaveine and Thomas Verdebout	
On Minimax Shrinkage Estimation with Variable Selection	65
Stavros Zinonos and William E. Strawderman	
On the Finite-Sample Performance of Measure-Transportation-Based Multivariate Rank Tests	87
Marc Hallin and Gilles Mordant	
Refining Invariant Coordinate Selection via Local Projection Pursuit	121
Lutz Dümbgen, Katrin Gysel, and Fabrice Perler	
Directional Distributions and the Half-Angle Principle	137
John T. Kent	

Part III Robust Theory and Methods

Power M-Estimators for Location and Scatter	157
Gabriel Frahm	
On Robust Estimators of a Sphericity Measure in High Dimension	179
Esa Ollila and Hyon-Jung Kim	

Detecting Outliers in Compositional Data Using Invariant Coordinate Selection	197
Anne Ruiz-Gazen, Christine Thomas-Agnan, Thibault Laurent, and Camille Mondon	
Robust Forecasting of Multiple Time Series with One-Sided Dynamic Principal Components	225
Daniel Peña and Víctor J. Yohai	
Robust and Sparse Estimation of Graphical Models Based on Multivariate Winsorization	249
Ginette Lafit, Javier Nogales, Marcelo Ruiz, and Ruben Zamar	
Robustly Fitting Gaussian Graphical Models—the R Package robFitConGraph	277
Daniel Vogel, Stuart J. Watt, and Anna Wiedemann	
Robust Estimation of General Linear Mixed Effects Models	297
Manuel Koller and Werner A. Stahel	
Asymptotic Behaviour of Penalized Robust Estimators in Logistic Regression When Dimension Increases	323
Ana M. Bianco, Graciela Boente, and Gonzalo Chebi	
Conditional Distribution-Based Downweighting for Robust Estimation of Logistic Regression Models	349
Weichang Yu and Howard D. Bondell	
Bias Calibration for Robust Estimation in Small Areas	365
Setareh Ranjbar, Elvezio Ronchetti, and Stefan Sperlich	
The Diverging Definition of Robustness in Statistics and Computer Vision	395
Peter Meer	
Part IV Other Methods	
Power Calculations and Critical Values for Two-Stage Nonparametric Testing Regimes	409
John Kolassa, Xinyan Chen, Yodit Seifu, and Dewei Zhong	
Data Nuggets in Supervised Learning	429
Kenneth Edward Cherasia, Javier Cabrera, Luisa T. Fernholz, and Robert Fernholz	
Improved Convergence Rates of Normal Extremes	451
Yijun Zhu and Han Xiao	
Local Spectral Analysis of Qualitative Sequences via Minimum Description Length	477
David S. Stoffer	

List of Contributors

Ana M. Bianco Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Buenos Aires, Argentina

Gracila Boente Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Buenos Aires, Argentina

Howard Bondell School of Mathematics and Statistics, University of Melbourne, Parkville, Australia

Javier Cabrera Department of Statistics, Rutgers University, Piscataway, NJ, USA

Gonzalo Chebi Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Buenos Aires, Argentina

Xinyan Chen Department of Statistics, Rutgers University, Piscataway, NJ, USA

Kenneth Edward Cherasia Department of Statistics, Rutgers University, Piscataway, NJ, USA

Lutz Dümbgen Institute of Mathematical Statistics and Actuarial Science, University of Bern, Bern, Switzerland

Luisa T. Fernholz Department of Statistics, Temple University, Philadelphia, PA, USA

Robert Fernholz Intech Corp., Princeton, NJ, USA

Daniel Fischer Applied Statistical Methods, Natural Resources Institute Finland (Luke), Jokioinen, Finland

Gabriel Frahm Department of Mathematics and Statistics, Helmut Schmidt University, Hamburg, Germany

Katrin Gysel SAKK Kompetenzzentrum, Bern, Switzerland

Marc Hallin ECARES and Département de Mathématique, Université libre de Bruxelles, Brussels, Belgium

John T. Kent Department of Statistics, University of Leeds, Leeds, UK

Hyon-Jung Kim Faculty of Information Technology and Communication Sciences, Tampere University, Tampere, Finland

John Kolassa Department of Statistics, Rutgers University, Piscataway, NJ, USA

Manuel Koller Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland
Seminar für Statistik, ETH Zürich, Zürich, Switzerland

Ginette Lafit Research Group of Quantitative Psychology and Individual Differences, KU Leuven, Leuven, Belgium

Thibault Laurent Toulouse School of Economics, French National Centre for Scientific Research, Toulouse, France

Peter Meer Department of Electrical Engineering, Rutgers University, Piscataway, NJ, USA

Camille Mondon Département de Mathématiques et Applications, Ecole Normale Supérieure, Paris, France

Gilles Mordant Institute for Mathematical Stochastics, Universität Göttingen, Göttingen, Germany

Javier Nogales Department of Statistics, Universidad Carlos III de Madrid, Getafe, Spain

Klaus Nordhausen Department of Mathematics and Statistics, University of Jyväskylä, Jyväskylä, Finland

Hannu Oja Department of Mathematics and Statistics, University of Turku, Turku, Finland

Esa Ollila School of Electrical Engineering, Aalto University, Espoo, Finland

Davy Paindaveine ECARES and Mathematics Department, Université libre de Bruxelles, Brussels, Belgium

Daniel Pena Department of Statistics, Universidad Carlos III de Madrid, Getafe, Spain

Fabrice Perler Bundesamt für Gesundheit BAG, Leibefeld, Switzerland

Setareh Ranjbar Faculty of Business and Economics, University of Lausanne, Lausanne, Switzerland

Elvezio Ronchetti Research Center for Statistics and Geneva School of Economics and Management, University of Geneva, Geneva, Switzerland

Marcelo Ruiz Department of Mathematics, Universidad Nacional de Río Cuarto, Río Cuarto, Argentina

Anne Ruiz-Gazen Toulouse School of Economics, University of Toulouse 1 Capitole, Toulouse, France

Yodit Seifu Bristol-Myers Squibb, Berkeley Heights, NJ, USA

Stefan Sperlich Geneva School of Economics and Management, University of Geneva, Geneva, Switzerland

Werner A. Stahel Seminar für Statistik, ETH Zürich, Zürich, Switzerland

David S. Stoffer Department of Statistics, University of Pittsburgh, Pittsburgh, PA, USA

William E. Strawderman Department of Statistics, Rutgers University, Piscataway, NJ, USA

Sara Taskinen Department of Mathematics and Statistics, University of Jyväskylä, Jyväskylä, Finland

Christine Thomas-Agnan Toulouse School of Economics, University of Toulouse 1 Capitole, Toulouse, France

Thomas Verdebout Mathematics Department, Université libre de Bruxelles, Brussels, Belgium

Daniel Vogel MEDICE Arzneimittel Pütter GmbH & Co. KG, Iserlohn, Germany
Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen, UK

Stuart J. Watt Mirador Analytics, Melrose, UK

Anna Wiedemann Department of Psychiatry, University of Cambridge, Cambridge, UK

Han Xiao Department of Statistics, Rutgers University, Piscataway, NJ, USA

Mengxi Yi School of Statistics, Beijing Normal University, Beijing, China

Victor J. Yohai Department of Mathematics and Instituto de Calculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina

Weichang Yu School of Mathematics and Statistics, University of Melbourne, Parkville, Australia

Ruben Zamar Department of Statistics, University of British Columbia, Vancouver, BC, Canada

Dewei Zhong Anheuser-Busch Companies, New York, NY, USA

Yijun Zhu Department of Statistics, Rutgers University, Piscataway, NJ, USA

Stavros Zinonos Cardiovascular Institute of New Jersey, RWJMS, New Brunswick, NJ, USA