The International Workshop on Precision Medicine

精准医疗国际研讨会

December 11, 2022 · Shanghai

Committees

General Chairs

Jun Shao	The University of Wisconsin-Madison, USA
Yong Zhou	East China Normal University, China
Donglin Zeng	The University of North Carolina at Chapel Hill, USA
Wenbin Lu	NC State University, USA
Yincai Tang	East China Normal University, China

Program Chairs

Zhou Yu	East China Normal University, China
Yingchun Zhou	East China Normal University, China
Yukun Liu	East China Normal University, China
Fang Fang	East China Normal University, China
Yanlin Tang	East China Normal University, China

Organization Committee

Wei Zhao	East China Normal University, China
Shanping Wang	East China Normal University, China
Shuya Cai	East China Normal University, China
Luyao Wang	East China Normal University, China
Ying Zhang	East China Normal University, China

Program Committee

(in alphabetical order)

Tianxi Cai	Harvard University, USA
Yufeng Liu	University of North Carolina at Chapel Hill, USA
Shujie Ma	University of California at Riverside, USA
Lu Tian	Stanford University, USA
Menggang Yu	The University of Wisconsin-Madison, USA
Jian Huang	The Hong Kong Polytechnic University, China
Xiaohua Zhou	Peking University, China
Yingchun Zhou	East China Normal University, China
Jin Xu	East China Normal University, China
Huijuan Ma	East China Normal University, China

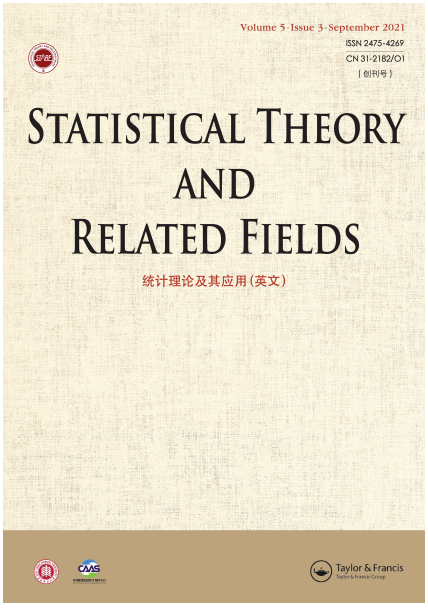
General Information

Precision medicine, also known as "personalized medicine" is an innovative approach to tailoring disease prevention and treatment that takes into account differences in people's genes, environments, and lifestyles. The goal of this workshop is to introduce current researches and developments in precision medicine with focus on both the theory and applications of methodologies in statistics and deep learning, and promote international communication between academic circles and industry. The 2022 International Workshop on Precision Medicine will be hosted by East China Normal University (ECNU) on December 11, 2022. The conference will be held virtually through Zoom platform.

Sponsors

- School of Statistics, East China Normal University
- The Chinese Association for Applied Statistics (CAAS)
- Editorial Office of Statistical Theory and Related Fields
- Key Laboratory of Advanced Theory and Application in Statistics and Data Science - MOE

About *Statistical Theory and Related Fields*



Statistical Theory and Related Fields（《统计理论及其应用(英文)》）由华东师范大学主办并由华东师范大学与英国 Taylor & Francis 合作出版，是国内首家统计学领域英文学术期刊, 于 2017 年 8 月正式创刊，2019 年 4 月被世界最大引文与评价数据库 Scopus 收录，2022 年 4 月被中国科学引文数据库(CSCD)收录, 2022 年 9 月被世界开放存取期刊数据库(DOAJ)收录，2022 年 11 月被 Web of Science 的 ESCI 收录。2019 年 11 月入选《中国科技期刊卓越行动计划》"高起点新刊"，2021 年 2 月得到国内出版许可刊号(CN31-2182/O1), 2021 年 9 月中国大陆创刊，从 2022 年第一期刊开始采用免费开放获取(OA)方式。

欢迎大家踊跃参会并投稿！

网站：https://www.tandfonline.com/toc/tstf20/current
https://xblk.ecnu.edu.cn/starf/

编辑部：上海市中山北路 3663 号
华东师范大学电化教学楼 513 室

Program

Zoom conference ID: 81204911599, Password: 137698

Events	Time	Speaker/Title	Chair
Opening Ceremony	08:10-08:20	**Yong Zhou,** ECNU, Welcome speech **Wenbin Lu**, North Carolina State University, Welcome speech	**Yincai Tang**, ECNU
Morning Session	08:20-09:00	**Donglin Zeng**, The University of North Carolina at Chapel Hill, USA **Title**: Learning Optimal Group-Structured Individualized Treatment Rules	**Yingchun Zhou**, ECNU
	09:00-09:40	**Tian Lu**, Stanford University, USA **Title**: Estimation and Validation of a Class of Conditional Average Treatment Effects Based on Observational Data	
	09:40-10:00	**Tea Break**	
	10:00-10:40	**Peter Song**, University of Michigan, USA **Title**: Synergistic Self-Learning Approach to Establishing Individualized Treatment Rules from Multiple Benefit Outcomes	**Yukun Liu**, ECNU
	10:40-11:20	**Jian Huang**, The Hong Kong Polytechnic University, China **Title**: Beyond the Cox Model: A Generative Learning Approach to Nonparametric Analysis of Right Censored Data	
Afternoon Session	13:30-14:00	**Tao Wang**, Shanghai Jiao Tong University, China **Title**: Analysis of compositions of microbiomes: Testing hypotheses regarding unobservable absolute abundances	**Guanghui Wang**, ECNU
	14:00-14:30	**Wensheng Zhu**, Northeast Normal University, China **Title**: Robust Covariate-Balancing Method in Learning Optimal Individualized Treatment Regimes	
	14:30-15:00	**Jin Xu**, East China Normal University, China **Title**: Model-free Screening for Variables with Treatment Interaction	
	15:00-15:20	**Tea Break**	
	15:20-15:50	**Xingjie Shi**, East China Normal University, China **Title**: Dimension reduction, clustering and annotation for spatial transcriptomics data	**Yuqi Qiu**, ECNU
	15:50-16:20	**Jingsi Ming**, East China Normal University, China **Title**: Flexible Integration of single-cell RNA-sequencing data for large-scale Multi-tissue cell atlas datasets	
	16:20-16:50	**Shuyi Zhang**, East China Normal University, China **Title**: Distributed Algorithms for U-statistics-based Empirical Risk Minimization	**Yan Zhong**, ECNU
	16:50-17:20	**Mengjiao Peng**, East China Normal University, China **Title**: Estimating optimal treatment regimes in semi-supervised framework	

Abstracts of Invited Talks

Speaker	Abstract
 **Donglin Zeng**, The University of North Carolina at Chapel Hill, USA	**Title**: Learning Optimal Group-Structured Individualized Treatment Rules **Abstract**: One essential problem in precision medicine is to determine optimal Individualized Treatment Rules (ITRs) that tailor treatment decisions to patient-specific characteristics so as to maximize rewarding outcome. In practice, many treatment options are usually available, and this poses a significant challenge for learning reliable treatment rules given limited data. In this work, we propose GRoup Outcome Weighted Learning (GROWL) to simultaneously learn both group structure among the treatments and the resulting optimal group-based ITRs. Our approach combines treatment clustering and value optimization through one single algorithm. We provide theoretical guarantee for GROWL by establishing the results for Fisher consistency, excess risk bound, and non-asymptotic convergence rate. Extensive simulation studies and real data analysis are performed to demonstrate the superior performance of this method.
	Bio: Donglin Zeng is a professor in Department of Biostatistics, University of North Carolina at Chapel Hill. He received his Ph.D. degree in Statistics from University of Michigan in 2001. His research interests include method development on machine learning, personalized medicine, semiparametric inference and high dimensional inference, with particular applications to EHR data, clinical trials, survival data and genetics in biomedical studies. He has published nearly 200 articles in various statistical journals.
 **Tian Lu**, Stanford University, USA	**Title**: Estimation and Validation of a Class of Conditional Average Treatment Effects Based on Observational Data **Abstract**: While sample sizes in randomized clinical trials are large enough to estimate the average treatment effect, they are oftentimes insufficient for estimating the treatment-covariate interactions critical to the development of precision medicine. Observational data from real world practice may play an important role in alleviating this problem. One common approach in trials is to predict the outcome of interest with separate regression models in each treatment arm, and recommend interventions based on the contrast of the predicted outcomes. Unfortunately, in contrast to clinical trial setting, this simple approach may induce spurious treatment-covariate interaction in observational studies when the regression model is mis-specified. Motivated by the need of modeling the number of relapses in multiple sclerosis patients, where the ratio of relapse rates is a natural choice of the treatment effect, we propose to estimate the conditional average treatment effect in this setting, and derive a doubly robust estimator of this conditional average treatment effect via a semiparametric regression model containing a treatment-covariate interactions. We also provide a validation procedure to check the quality of the estimator on an independent sample. We conduct simulations to demonstrate the finite sample performance of the proposed methods, and illustrate the advantage of this approach on real data examining the treatment effect of dimethyl fumarate compared to teriflunomide in multiple sclerosis patients.
	Bio: Tian Lu received his Ph.D. in Biostatistics from Harvard University in 2002. Before joining the Stanford community in 2007, he was an assistant professor in the Department of Preventive Medicine at Northwestern University. He is now a professor in the Department of Biomedical Data Science (DBDS) and, by courtesy, in the Department of Statistics. He is the elected fellow of American Statistical Association. His methodological research interests include survival analysis, resampling method, meta analysis, high dimensional data analysis, precision medicine for disease diagnosis etc.

Peter Song, Michigan University, USA

Title: Synergistic Self-Learning Approach to Establishing Individualized Treatment Rules from Multiple Benefit Outcomes

Abstract: Precision health is an emerging research field in healthcare. This talk presents a novel method, termed synergistic self-learning (SS-learning), to address two major challenges in the derivation of individualized treatment rule (ITR) in the presence of multiple clinical outcomes, including heterogeneous multiview outcomes and complex missing data patterns. Our proposed SS-learning can effectively synergize heterogeneous features of multiple training data sources in the derivation of ITR. Applying the proposed SS-learning method to a nutritional supplement trial, we identified several important biomarkers that can be used to form an ITR that would lead to a higher health benefit should it be implemented in the whole study population. We also examined the sensitivity and stability of SS-learning through comprehensive simulation studies.

Bio: Peter Song is a Professor of Biostatistics in the Department of Biostatistics, School of Public Health at the University of Michigan. He received his Ph.D. in Statistics from the University of British Columbia in 1996. He has published over 190 peer-reviewed papers. Pro. Song's current research interests include data integration, distributed inference, longitudinal data analysis, high-dimensional statistics, statistical learning, statistical methods for bioinformatics and genetics, and spatiotemporal modeling. He is interested in methodological developments related to modelling, statistical inference and applications in biomedical sciences. Pro. Song now serves as Associate Editor of the Journal of American Statistical Association, the Annals of Applied Statistics, and the Journal of Multivariate Analysis.

Jian Huang, The Hong Kong Polytechnic University, China

Title: Beyond the Cox Model: A Generative Learning Approach to Nonparametric Analysis of Right Censored Data

Abstract: We propose a deep generative learning approach to nonparametric estimation of conditional survival and hazard functions with right-censored data. The key idea of the proposed method is to first learn a conditional generator for the joint conditional distribution of the observed time and censoring indicator given the covariates, and then construct the Kaplan-Meier and Nelson-Aalen estimators based on this conditional generator for the conditional hazard and survival functions. Our method combines ideas from the recently developed deep generative learning approach and classical nonparametric estimation in survival analysis. We analyze the convergence properties of the proposed method and establish the consistency of the generative nonparametric estimators of the conditional survival and hazard functions. Our numerical experiments validate the proposed method and demonstrate its superior performance in a range of simulated models. We also illustrate the applications of the proposed method in constructing prediction intervals for survival times with two data examples.

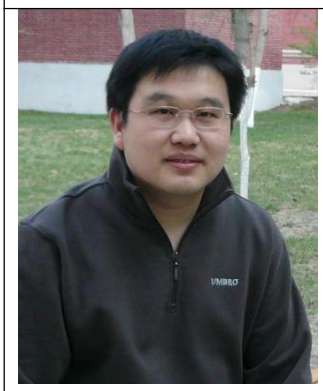
Bio: Jian Huang is a chair professor of Data Science and Analytics in the Department of Applied Mathematics at The Hong Kong Polytechnic University. He obtained his Ph.D. degree in Statistics from the University of Washington in Seattle. His research interests include machine learning, high-dimensional statistics, computational statistics, biostatistics, and bioinformatics. He has published more than 100 papers widely in the fields of Statistics, Biostatistics, Machine Learning, Bioinformatics and Econometrics. He was designated a highly cited researcher in the field of Mathematics from 2015 to 2019 by the Web of Science group at Clarivate and included in the list of top 2% of the world's most cited scientists by Elsevier BV and Stanford University (2021). Professor Huang is a fellow of the American Statistical Association and a fellow of the Institute of Mathematical Statistics.

Tao Wang, Shanghai Jiao Tong University, China

Title: Analysis of compositions of microbiomes: Testing hypotheses regarding unobservable absolute abundances

Abstract: Detection of differentially abundant microbes between two or more environments, known as differential abundance analysis, is at the core of statistical analysis of microbiome data. A major hurdle in performing differential abundance testing is that the observed abundances are compositional representing only relative information. ANCOM is the most popular method for drawing inferences regarding absolute abundances at the ecosystem level using the specimen level relative abundances. Despite its impressive performance, there are two drawbacks to ANCOM. First, with K microbes it requires fitting K(K-1)/2 models for log-ratios of counts, and so can be computationally intensive. Second, it does not output P-values for microbes detected as differentially abundant. We propose a fast implementation of ANCOM, fastANCOM, that fits only K models for log-transformed counts. fastANCOM provides P-values to declare statistical significance and outputs log fold changes of abundance between groups. fastANCOM compares favorably with existing differential abundance testing methods.

Bio: Tao Wang is a tenured associate professor in Department of Bioinformatics and Biostatistics and Department of Statistics, Shanghai Jiao Tong University. He achieved his Ph.D. degree in statistics at Hong Kong Baptist University, and was a Postdoctoral Associate in the Department of Biostatistics at Yale School of Public Health. He is a reviewer for Mathematical Reviews, an ISI elected member, and an Assistant Professor Adjunct of Biostatistics at Yale School of Public Health.



Wensheng Zhu, Northeast Normal University, China

Title: Robust Covariate-Balancing Method in Learning Optimal Individualized Treatment Regimes

Abstract: One of the most important problems concerned in precision medicine is to find the optimal individualized treatment rule, which is designed to recommend treatment decisions and maximize the overall clinical benefit to the patients based on their individual characteristics. Typically, the expected clinical outcome (i.e. the value function) is required to be estimated first, in which an outcome regression model or a propensity score model usually need to be assumed for the most of the existing statistical methods. However, if any of the above assumptions are invalid, the estimated treatment regime is not reliable. In this article, we first define a contrast value function, which is the basis of the study for individualized treatment regime (ITR). Then we construct a general framework of a hybrid estimator to estimate the contrast value function by combining two types of estimation methods. We further propose a robust covariate-balancing (RCB) estimator of the contrast value function by combining the inverse probability weighted (IPW) method and matching method, which is based on the Covariate Balancing Propensity Score (CBPS). The theoretical results show that the RCB estimator is doubly robust, that is, it is consistent if either the propensity score model or the matching is correct. Based on a large number of simulation studies, we demonstrate that the RCB estimator outperforms existing methods. Lastly, the proposed method is illustrated through analysis of AIDS clinical trial data.

Bio: Wensheng Zhu received his Ph.D. degree from Northeast Normal University in December 2006 and has been a professor at the School of Mathematics and Statistics of Northeast Normal University since December 2013. He was a postdoctoral fellow at Yale University from 2008 to 2010 and visited the University of North Carolina at Chapel Hill from 2015 to 2017. His research interests are biostatistics and bioinformatics. He has published many academic papers in the international top journals of statistics, such as Journal of the American Statistical Association (JASA) and NeuroImage. He has chaired and completed a number of National Natural Science Foundation

	projects.
 **Jin Xu**, East China Normal University, China	**Title**: Model-free Screening for Variables with Treatment Interaction **Abstract**: Precision medicine is a medical paradigm that focuses on making effective treatment decision based on individual patient characteristics. When there are a large amount of patient information, such as patient's genetic information, medical records and clinical measurements, available, it is of interest to select the covariates which have interactions with the treatment, for example, in determining the individualized treatment regime where only a subset of covariates with treatment interactions involves in decision making. We propose a marginal feature ranking and screening procedure for measuring interactions between the treatment and covariates. The method does not require imposing a specific model structure on the regression model and is applicable in a high dimensional setting. Theoretical properties in terms of consistency in ranking and selection are established. We demonstrate the finite sample performance of the proposed method by simulation and illustrate the applications with two real data examples from clinical trials.
	Bio: Jin Xu is a professor at School of Statistics, East China Normal University. He received his Ph.D. at Bowling Green State University and was a postdoctoral fellow at University of California, Riverside. He was a senior biostatistician at Ambion Biotechnology Company in the United States. His research interests include personalized medicine, clinical trials, oncology clinical trial design, survival analysis, multivariate analysis, etc. He has published dozens of papers in biomedical statistical journals, including JASA, JRSSC, Statistical Methods in Medical Research, Statistics in Biopharmaceutical Research, Bioinformatics etc.
 **Xingjie Shi**, East China Normal University, China	**Title**: Dimension reduction, clustering and annotation for spatial transcriptomics data **Abstract**: With rapid advancement in spatially resolved transcriptomics (SRT) technologies, it becomes feasible to comprehensively characterize the gene expression profiles of tissues while retaining their information on physical locations. These SRT technologies provide new opportunities to study how spatial organizations in tissues relate to tissue functions. As the multiplexing and resolution of the experimental technologies continue to improve, there is an emerging need for the development of analytical approaches that take into account the specific biological questions at hand as well as the distinct features and limitations of different measurement methods. In this talk, I will introduce some recent developed statistical methods for dimension reduction, clustering and annotation for spatial transcriptomics data.
	Bio: Xingjie Shi received his Ph.D. from Shanghai University of Finance and Economics in 2014. He was co-trained by the Department of Biostatistics at Yale University from 2012 to 2014, and joined the Academy of Statistics and Interdisciplinary Sciences at East China Normal University in June 2021. He is mainly engaged in the research of omics genetics data fusion, high-dimensional big data statistical calculation, survival analysis and other aspects. More than 20 papers have been published in important journals at home and abroad, such as Nucleic Acids Research, Bioinformatics, Briefing in Bioinformatics, etc. Now, he has been elected as a member of the International Statistical Association.

Jingsi Ming, East China Normal University, China

Title: Flexible Integration of single-cell RNA-sequencing data for large-scale Multi-tissue cell atlas datasets

Abstract: Single-cell RNA-sequencing (scRNA-seq) is being used extensively to measure the mRNA expression of individual cells from deconstructed tissues, organs, and even entire organisms to generate cell atlas references, leading to discoveries of novel cell types and deeper insight into biological trajectories. These massive datasets are usually collected from many samples using different scRNA-seq technology platforms, including the popular SMART-Seq2 (SS2) and 10X platforms. Inherent heterogeneities between platforms, tissues, and other batch effects makes scRNA-seq data difficult to compare and integrate, especially in large-scale cell atlas efforts; yet, accurate integration is essential for gaining deeper insights into cell biology. Through comprehensive data exploration, we found that accurate integration is often hampered by differences in cell-type compositions. Herein we describe FIRM, an algorithm that addresses this problem and achieves efficient and accurate integration of heterogeneous scRNA-seq datasets across multiple tissue types, platforms, and experimental batches. We applied FIRM to numerous large-scale scRNA-seq datasets from mouse, mouse lemur, and human, comparing its performance in dataset integration with other state-of-the-art methods. FIRM-integrated datasets show accurate mixing of shared cell type identities and superior preservation of original structure without overcorrection, generating robust integrated datasets for downstream exploration and analysis. It is also a facile way to transfer cell type labels and annotations from one dataset to another, making it a reliable and versatile tool for scRNA-seq analysis, especially for cell atlas data integration.

Bio: Jingsi Ming is an assistant professor at the Academy of Statistics and Interdisciplinary Sciences, Faculty of Economics and Management, East China Normal University. She received his Ph.D. from Hong Kong Baptist University in 2018 and worked as a postdoctoral researcher at Hong Kong University of Science and Technology from 2018 to 2020. She joined East China Normal University in September 2020. Her research interests include statistical genetics, bioinformatics, statistical machine learning, etc. Her research results are published in Briefings in Bioinformatics, Bioinformatics, Nature Computational Science, Journal of Computational and Graphical Statistics and other journals. And she is selected for Shanghai Sailing Program.

Shuyi Zhang, East China Normal University, China

Title: Distributed Algorithms for U-statistics-based Empirical Risk Minimization

Abstract: Empirical risk minimization, where the underlying loss function depends on a pair of data points, covers a wide range of application areas in statistics including pairwise ranking and survival analysis. The common empirical risk estimator obtained by averaging values of a loss function over all possible pairs of observations is essentially a U-statistic. One well-known problem with minimizing U-statistic type empirical risks, is that the computational complexity of U-statistics increases quadratically with the sample size. When faced with big data, this poses computational challenges as the colossal number of observation pairs virtually prohibits centralized computing to be performed on a single machine. This paper addresses this problem by developing two computationally and statistically efficient methods based on the divide-and-conquer strategy on a decentralized computing system, whereby the data are distributed among machines to perform the tasks. One of these methods is based on a surrogate of the empirical risk, while the other method extends the one-step updating scheme in classical M-estimation to the case of pairwise loss. We show that the proposed estimators are as asymptotically efficient as the benchmark global U-estimator obtained under centralized computing. As well, we introduce two distributed iterative algorithms to facilitate the implementation of the proposed methods, and

	conduct extensive numerical experiments to demonstrate their merit.
	Bio: Shuyi Zhang is an assistant professor at the Academy of Statistics and Interdisciplinary Sciences, East China Normal University. She was a postdoctoral fellow at the Department of Statistics of Harvard University. And she was a joint doctoral student at Guanghua School of Management of Peking University and the Department of Statistics of Iowa State University. Her research interests include big data statistical analysis, semi-supervised learning, high-dimensional statistics, spatio-temporal statistics, and environmental risk measurement. She has published 5 academic papers in top international journals and owned 1 invention patent.
 **Mengjiao Peng**, East China Normal University, China	**Title**: Estimating optimal treatment regimes in semi-supervised framework **Abstract**: Finding the optimal individualized treatment rule mapping from the individual characteristics or contextual information to the treatment assignment has been studied intensively in the literature, with important applications in practice. We consider the problem of estimating the optimal treatment regime in a semi-supervised learning setting, where a very small proportion of the entire set of observations are labeled with the true outcome but features predictive of the outcome are available among all observations. We propose a model-free robust inference approach for optimal treatment regime by the aid of the unlabeled data with only covariate information to improve estimation efficiency. The proposed estimation of OPT primarily involves a flexible nonparametric imputation by single index kernel smoothing which works well even for high-dimensional covariates; and a follow-up estimation for optimal treatment regime based on concordance-assisted learning, including optimization of the estimated concordance function up to a threshold and finding the optimal threshold to maximize the inverse propensity score weighted (IPSW) estimator of the value function. Moreover, when the propensity score function is unknown, a doubly robust estimation method is developed under a class of monotonic index models. Our estimators are shown to be consistent and asymptotically normal. Simulations exhibit the efficiency and robustness of the proposed method compared to existing approaches in finite samples.
	Bio: Mengjiao Peng received her Ph.D. in statistics from Nanyang Technological University in 2019, and then worked as a lecturer at Purdue University in the United States for one year. In September 2020, she joined the Academy of Statistics and Interdisciplinary Sciences, East China Normal University. Her research interests include survival analysis, high-dimensional statistics, statistical machine learning and causal inference. She has published five papers in some international journals, including Statistical Methods in Medical Research, Computational Statistics Data Analysis, Expert Systems with Applications and others.