数学与系统科学研究院

计算数学所学术报告

报告人： **Prof. Ming-Hui Chen**

（ *Department of Statistics, University of Connecticut* ）

报告题目：

Online Updating of Statistical Inference in the Big Data Setting

邀请人： 胡星标 研究员

报告时间：**2015** 年 **7** 月 **1** 日（周三）

上午 **10:00~11:00**

报告地点： 数学院南楼二层

202 会议室

Abstract:

We present statistical methods for big data arising from online analytical processing, where large amounts of data arrive in streams and require fast analysis without storage/access to the historical data. In particular, we develop iterative estimating algorithms and statistical inferences for linear models and estimating equations that update as new data arrive. These algorithms are computationally efficient, minimally storage-intensive, and allow for possible rank deficiencies in the subset design matrices due to rare-event covariates. Within the linear model setting, the proposed online-updating framework leads to predictive residuals tests that can be used to assess the goodness-of-fit of the hypothesized model. We also propose a new online-updating estimator under the estimating equation setting. Theoretical properties of the goodness-of-fit tests and proposed estimators are examined in detail. In simulation studies and real data applications, our estimator compares favorably with competing approaches under the estimating equation setting.

This is the joint work with Elizabeth D. Schifano, Jing Wu, Chun Wang, and Jun Yan of University of Connecticut.

Biosketch

Ming-Hui Chen, Ph.D. is Professor and the Director of the Statistical Consulting Services (a statistical consulting center) in the Department of Statistics, University of Connecticut. He was elected to Fellow of the Institute of Mathematical Statistics in 2007, Fellow of American Statistical Association in 2005, an elected ordinary member of the International Statistical Institute (ISI) in 1999. He has been awarded several US NIH and NSF grants since 1994. He received the AAUP (American Association of the University Professors) Research Excellence Awardand the CLAS (College of Liberal Arts and Sciences) Excellence in ResearchAward in the Physical Sciences Division,University of Connecticut, 2013 and the UCONN Alumni Association's University Award for Faculty Excellence in Research and Creativity (Sciences), University of Connecticut, 2014.

He co-authored three books: *Bayesian Survival Analysis* (with J. Ibrahim, D. Sinha), Springer, 2001*; Monte Carlo Methods in Bayesian Computation* (with Q. Shao, J. Ibrahim), Springer, 2000; and *Applied Statistics for Engineers* (with J. Petruccelli, B. Nandram), Prentice-Hall, 1999. He also co-edited two books*: Frontiers of Statistical Decision Making and Bayesian Analysis --- In Honor of James O. Berger* (with Dey, D.K., Müller, P., Sun, D., and Ye, K.), Springer, 2010 and *Bayesian Phylogenetics: Methods, Algorithms, and Applications* (with Kuo, L. and Lewis, P.O.), Chapman and Hall/CRC, 2014.He has published over 325 articles in mainstream statistics, biostatistics, and medical journals, including *Annals of Statistics, Journal of the American Statistical Association, Biometrika, Journal of the Royal Statistical Society, Series B, C, and D, Biometrics, New England Journal of Medicine, The Lancet, The Journal of the American Medical Association, etc.*

Currently, he serves as an Editor of*Bayesian Analysis* and*Statistics andIts Interface*as well asAssociate Editors for *Journal of the American Statistical Association, Lifetime Data Analysis, and Journal of Computational and Graphical Statistics.* He was the President (2013) of *the International Chinese Statistical Association (ICSA,* 泛华统计协会) and served on the board of directors of*the International Society for Bayesian Analysis (ISBA)* for 2011-2013.

Dr. Chen has special interest in the areas of Bayesian Statistical Methodology, Categorical Data Analysis, Design of Bayesian clinical trials, Bayesian DNA Microarray Data Analysis, Bayesian Phylogenetics, Meta analysis, Missing Data Analysis (EM, MCEM, and Bayesian), Monte Carlo Methodology, Prior Elicitation, Statistical Analysis and Methodology for Prostate Cancer Data, Statistical Modeling and computing, Survival data analysis, and Variable Selection.

欢迎大家参加！