

CURRICULUM VITAE

Yanxun Xu

Department of Applied Mathematics and Statistics
Johns Hopkins University
Baltimore, MD

Phone: 410-516-7341
Email: yanxun.xu@jhu.edu
Homepage: www.ams.jhu.edu/~yxu70

EDUCATION

Ph.D. in Statistics (Mentor: Prof. Yuan Ji)
Rice University, Houston, TX August 2010 - July 2013
The University of Texas MD Anderson Cancer Center (Joint program)

M.S. in Statistics, minor in Economics
Texas Tech University, Lubbock, TX August 2008 - August 2010

B.S. in Mathematics and Applied Mathematics
Beijing University of Aeronautics and Astronautics, Beijing, China September 2005 - July 2007

EMPLOYMENT

Assistant Professor
Department of Applied Mathematics and Statistics July 2015 - present
Johns Hopkins University, Baltimore, MD

Postdoctoral Fellow (Mentor: Prof. Peter Müller)
Department of Statistics and Data Sciences August 2013 - July 2015
The University of Texas at Austin, Austin, TX

Research Assistant
Department of Bioinformatics and Computational Biology June 2010 - June 2011
Department of Biostatistics July 2011 - July 2013
The University of Texas MD Anderson Cancer Center, Houston, TX

Visiting Student
Dr. Kevin White Lab, The Institute for Genomics and Systems Biology - IGSB January - February 2013
University of Chicago, Chicago, IL

Project Coordinator
Center for Clinical Research and Informatics June - August 2012
NorthShore University HealthSystem, Chicago, IL

RESEARCH INTERESTS

- **Bayesian models for big data analysis, high-throughput genomics and proteomics data**
 - Bayesian feature allocation models for tumor heterogeneity
 - Network analysis of gene regulation by integrating transcription factor binding, histone modifications, DNA methylation, copy number variation and miRNA targeting
 - Nonparametric Bayesian bi-clustering for ChIP-Seq count data
- **Bayesian adaptive designs for clinical trials**
 - Bayesian nonparametric estimation of causal inference for dynamic treatment regimes

- Subgroup-based adaptive designs for multi-arm biomarker trials
- Latent Gaussian process models with application to monitoring clinical trials

PUBLICATIONS

*: Co-first authors

1. **Xu Y**, Müller P, Wahed A and Thall P, Bayesian Nonparametric Estimation for Dynamic Treatment Regimes with Sequential Transition Times. **Journal of the American Statistical Association. Accepted (Winner of the 2015 David P. Byar Young Investigator Travel Award)**
2. Zhu Y*, **Xu Y***, Helseth D*, Gulukota K, Yang S, Pesce L, Mitra R, Müller P, Sengupta S, Guo W, Silverstein J, Foster I, Parsad N, White K and Ji Y, Zodiac: A Comprehensive Depiction of Genetic Interactions in Cancer by Integrating TCGA Data. **Journal of the National Cancer Institute**, 2015, 107(8): djv129.
3. **Xu Y** and Ji Y, A Latent Gaussian Process Model with Application to Monitoring Clinical Trials. **Biometrics. Under review**
4. **Xu Y**, Müller P, Yuan Y, Gulukota K and Ji Y, MAD Bayes for Tumor Heterogeneity – Feature Allocation with Exponential Family Sampling. **Journal of the American Statistical Association**, 2015, 110(510): 503-514.
5. **Xu Y**, Lee J, Yuan Y, Mitra R, Liang S, Müller P and Ji Y, Nonparametric Bayesian Bi-Clustering for ChIP-Seq Count Data. **Bayesian Analysis**, 2013, 8(2): 1-22.
6. Ji Y, **Xu Y**, Zhang Q, Tsui K-W, Yuan Y, Liang S and Liang H, BM-Map: Bayesian Mapping of Multireads for Next-Generation Sequencing Data. **Biometrics**, 2011 Dec; 67(4): 1215-24. (**Winner of the 2011 ENAR Distinguished Student Paper Award**)
7. **Xu Y**, Trippa L, Müller P and Ji Y, Subgroup-Based Adaptive (SUBA) Designs for Multi-Arm Biomarker Trial. **Statistics in Biosciences**, 2014, DOI: 10.1007/s12561-014-9117-1(**1st Place Winner of the 2014 JSM Biopharmaceutical Section Student Paper**)
8. Yuan Y, Van Allen EM, Omberg L, Wagle N, Sokolov A, **Xu Y** and *et al.*, Assessing the Clinical Utility of Cancer Genomic and Proteomic Data across Tumor Types. **Nature Biotechnology**. 2014, 32: 644-652
9. **Xu Y**, Zheng X, Yuan Y, Estecio M, Issa J-P, Ji Y and Liang S, BM-SNP: A Bayesian Model for SNP Calling using High Throughput Sequencing Data. **IEEE/ACM transactions on computational biology and bioinformatics**. In Press
10. Yuan Y, **Xu Y**, Xu J and Liang H, Predicting the Lethal Phenotype of the Knockout Mouse by Integrating Genomic Data. **Bioinformatics**, 2012 May 1; 28(9): 1246-52.
11. Han L, Yuan Y, Zheng S, Edgerton M, Yang Y, Diao L, Li J, **Xu Y**, Verhaak R, The Cancer Genome Atlas Research Network and Liang H, The Pan-Cancer Analysis of Pseudogene Expression Reveals Biologically and Clinically Relevant Tumor Subtypes. **Nature Communications**. 2014, 5: 3963
12. **Xu Y**, Zhang J, Yuan Y, Mitra R, Müller P and Ji Y, A Bayesian Graphical Model for Integrative Analysis of TCGA Data. **2012 IEEE Workshop on Genomic Signal Processing and Statistics**.
13. Costello JC, Heiser LM, Georgii E, Gnen M, Menden MP, Wang NJ, Bansal M, Ammad-ud-din M, Hintsanen P, Khan SA, Mpindi JP, NCI DREAM Community (including **Xu Y**), Kallioniemi O, Honkela A, Aittokallio T, Wennerberg K, Collins JJ, Gallahan CD, Singer D, Saez-Rodriguez J, Kaski S, Gray JW, Stolovitzky G, The Challenge of Drug Sensitivity Prediction: An Assessment of Methods Through Collaborative Competition. **Nature Biotechnology**. In Press
14. Mitra R, Müller P, Liang S, **Xu Y** and Ji Y, Towards the Discovery of the Histone Code - A Bayesian Graphical Model for Histone Modifications. **Circulation: Cardiovascular Genetics**, 2013; 6: 419-426.

15. **Xu Y**, Zheng X, Yuan Y, Estecio M, Issa J-P, Ji Y and Liang S, A Bayesian Model for SNP Discovery Based on Next-Generation Sequencing Data. **2012 IEEE Workshop on Genomic Signal Processing and Statistics**.
16. Yuan Y, Norris C, **Xu Y**, Tsui KW, Ji Y and Liang H, BM-Map: An Efficient Software Package for Accurately Allocating Multireads of RNA-Seq Data. **BMC Genomics**, 2012, 13(Suppl 8): S9.
17. Trindade A.A. and **Xu Y**, Quantile Versions of Holt-Winters Forecasting Algorithms. **Journal of Statistics: Advances in Theory and Applications**, 2011, 5(1): 15-35.

BOOK CHAPTERS

1. **Xu Y**, Zhu Y and Ji Y, Graphical Models for Integrating Omics Data. In: Integrating Omics Data: Statistical and Computational Methods. Editors: George C. Tseng, Xianghong Jasmine Zhou and Debashis Ghosh. This book is under copyediting and production in Cambridge University Press; expected in early 2015.
2. **Xu Y**, Ji Y and Müller P, Biomarker-Driven Adaptive Design. In: Nonparametric Bayesian Methods in Biostatistics and Bioinformatics. Editors: Mitra R and Müller P. Springer-Verlag, 2015.

PAPERS IN PREPARATION

1. **Xu Y**, Müller P and Telesca D, Determinantal Point Process (DPP) prior for Diversity in Bayesian Latent Variable Models.
2. **Xu Y**, Müller P, Berry D. and et. al., A Population-Finding Design with Covariate-Dependent Random Partitions.
3. **Xu Y**, Müller P, and Thall P, Bayesian Subgroup-Based Adaptive Designs in Sequential, Multiple Assignment, Randomized Biomarker Trials (SMART).

HONORS AND AWARDS

- 2015 David P. Byar Young Investigator Travel Award January 2015
- 1st Place Winner in the ASA Biopharmaceutical Section Student Paper (2014 JSM) August 2014
- ENAR Student Distinguished Paper Award March 2011
- John & Eileen Tietze Fellowship, Rice University August 2010 - July 2011
- Southern Regional Council on Statistics 2014 Travel Award June 2014
- Women in Machine Learning 2013 Travel Award December 2013
- GENSIPS'12 Travel Award December 2012
- TTU Math Graduate Scholarship for Academic Excellence August 2008 - August 2010
- Outstanding Student Scholarship October 2006

TEACHING EXPERIENCE

Courses Taught (as Sole Instructor)

- STAT 632 Bayesian Statistics 2015 Fall
- MATH 1330 Introductory Mathematical Analysis I 2009 Fall
- MATH 2300 Introductory Statistics 2010 Spring

Teaching Assistant

- MATH 1330 Introductory Mathematical Analysis I 2009 Spring
- STAT 532 Mathematical Statistics I 2010 Fall
- STAT 581 Mathematical Probability I 2011 Spring
- STAT 553 Biostatistics 2011 Fall

Guest Lecturer

- M 362M Introduction to Stochastic Processes 2013 Fall
- SSC 386D Monte Carlo Methods in Statistics 2014 Spring
- SSC 325H Honors Statistics 2014 Spring

INVITED TALKS

- MAD Bayes for Modeling Tumor Heterogeneity Using Haplotype Variants
Johns Hopkins University, Baltimore, MD Jan 2015
- Subgroup-Based Adaptive (SUBA) Designs with Covariate-Dependent Random Partitions
Yale University, New Haven, CT Dec 2014
- MAD Bayes for Modeling Tumor Heterogeneity Using Haplotype Variants
University of Minnesota, Minneapolis, MN Nov 2014
- Bayesian Nonparametric Estimation for Dynamic Treatment Regimes with Sequential Transition Times
Rice University, Houston, TX July 2014
- Subgroup-Based Adaptive (SUBA) Designs for Multi-Arm Biomarker Trials
The University of Texas Medical School, Houston, TX July 2014
2014 Joint Applied Statistics Symposium of ICSA and KISS, Portland, OR June 2014
2014 Joint Statistical Meetings (JSM), Boston, MA August 2014
- Nonparametric Bayesian Bi-Clustering for ChIP-Seq Count Data
Johns Hopkins University, Baltimore, MD March 2013
- A Bayesian Graphical Model for Integrative Analysis of TCGA Data.
Northshore University HealthSystem, Evanston, IL June 2012
University of Chicago, Chicago, IL July 2012

CONFERENCE PRESENTATIONS

- 2014 Southern Regional Council on Statistics, Galveston, TX* June 2014
Poster Session: Bayesian Nonparametric Estimation for Dynamic Treatment Regimes with Sequential Transition Times.
- 2013 Women in Machine Learning Workshop, Lake Tahoe, NV* Dec 2013
Poster Session: MAD Bayes for Tumor Heterogeneity – Feature Allocation with Exponential Family Sampling.
- 2013 Eastern North American Region Meetings (ENAR), Orlando, FL* March 2013
Session Speaker: A Bayesian Graphical Model for Integrative Analysis of TCGA Data.
- 2012 IEEE International Workshop on Genomic Signal Processing and Statistics, DC* December 2012
Session Speaker: A Bayesian Model for SNP Discovery Based on Next-Generation Sequencing Data.
Session Speaker: A Bayesian Graphical Model for Integrative Analysis of TCGA Data.
- 2012 Joint Statistical Meetings (JSM), San Diego, CA* July 2012
Session Speaker: Hierarchical Gaussian Process Latent Variable Model for Clinical Decisions.
- 2011 Eastern North American Region Meetings (ENAR), Miami, FL* March 2011
Session Speaker: BM-Map: Bayesian Mapping of Multireads for Next-Generation Sequencing Data.

Fourth Annual Bayesian Biostatistics Conference, Houston, TX

January 2011

Poster Session: An Efficient Bayesian Algorithm for Accurately Allocating Multireads of RNA-Seq Data.

PEER REVIEWER

- Journal of the American Statistical Association
- Biometrical Journal
- Biostatistics
- BMC Bioinformatics
- Statistical Applications in Genetics and Molecular Biology
- IEEE Signal Processing Letters
- Current Bioinformatics
- Statistics in Medicine
- Journal of Machine Learning Research
- Journal of Biopharmaceutical Statistics