Yan Dora Zhang

CONTACT INFORMATION

Department of Statistics and Actuarial Science
RR304, 3/F, Run Run Shaw Building
The University of Hong Kong
Pokfulam Road, Hong Kong
Email: doraz@hku.hk
Website: http://web.hku.hk/~doraz

EDUCATION

2016	**Ph.D. in Statistics**, North Carolina State University, *NC, USA* Thesis: *Bayesian Methods for High-dimensional Data* Advisor: *Howard D. Bondell*
2013	**M.S. in Statistics**, North Carolina State University, *NC, USA*
2011	**B.S. in Statistics**, Nankai University, *Tianjin, China*

RESEARCH INTERESTS

Statistical genetics, bioinformatics, biostatistics, Bayesian methods, risk prediction
Application areas: cancer genetics, psychology disease, public health/biomedical research

PROFESSIONAL EXPERIENCE

2019–	**Assistant Professor** Department of Statistics & Actuarial Science, Faculty of Science Centre for PanorOmic Sciences, LKS Faculty of Medicine The University of Hong Kong, *Hong Kong SAR, China*
2016–2019	**Postdoctoral Fellow** (Mentor: *Nilanjan Chatterjee*) Department of Biostatistics, Bloomberg School of Public Health Johns Hopkins University, *Baltimore, MD, USA*
2014–2015	**Research Assistant** Target Sciences Statistics, GlaxoSmithKline, *Durham, NC, USA*
2014–2014	**Research Scientist Intern** Avaya Labs Research, Avaya, *Basking Ridge, NJ, USA*
2013–2013	**Research Assistant** Department of Civil Engineering North Carolina State University, *Raleigh, NC, USA*

SELECTED HONORS & AWARDS

2017   Honorable Mention for poster Estimating effect-size distribution from summary-level statistics for large genome-wide association studies" , Genetics Research Day, The Maryland Genetics, Epidemiology, and Medicine Training Program (MD-GEM)

2013   Nomination for the University Outstanding Teaching Assistant Award, NCSU

2013   Mu Sigma Rho Statistics Society Honor, NCSU

2011   Outstanding Graduates Honor, Nankai University, China

2010   Excellent Student Leader Honor, Nankai University, China

2009   National Scholarship, China

PUBLICATIONS

Peer-Reviewed

1. Tubbs JD, **Zhang YD**, Sham PC. (2020). Intermediate confounding in trio relationships: The importance of complete data in effect size estimation. *Genetic Epidemiology.*

2. Ji Y, Azuine RE, **Zhang Y**, Hou W, Hong X, Wang G, Riley A, Pearson C, Zuckerman B, Wang X. (2020). Association of cord plasma biomarkers of in utero acetaminophen exposure with risk of attention-deficit/hyperactivity disorder and autism spectrum disorder in childhood. *JAMA psychiatry*, 77(2), 180-189.

3. Pal Choudhury P, Wilcox AN, Brook MN, **Zhang Y**, Ahearn T, ..., Chatterjee N, García-Closas M. (2020). Comparative validation of breast cancer risk prediction models and projections for future risk stratification. *JNCI: Journal of the National Cancer Institute*, 112(3), 278-285.

4. **Zhang Y**, Qi G, Park JH, Chatterjee N. (2018). Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. *Nature genetics*, 50(9), 1318-1326.

5. **Zhang Y**, Bondell HD. (2018). Variable selection via penalized credible regions with Dirichlet–Laplace global-local shrinkage priors. *Bayesian Analysis*, 13(3), 823-844.

TEACHING

Lecturer

Spring 2020   STAT3621, Statistical data analysis, *HKU*

Lab Instructor

2013–2014   Experimental Statistics for Biological Sciences II, *NCSU*

Teaching Assistant

2011–2016   Applied Bayesian Statistics, Applied Multivariate Statistical Analysis, Experimental Statistics for Biological Sciences I, Introduction to Mathematical Statistics I, Experimental Statistics for Engineers II, Introduction to Statistics, *NCSU*

PRESENTATIONS

Invited Talk

Sep 2018   Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies,*School of Public Health, Peking University, Beijing, China*

Sep 2018   Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies, *SJTU-YALE Joint Center of Biostatistics and Data Science, Shanghai, China*

Sep 2018   Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies,*School of Statistics and Data Science, Nankai University, Tianjin, China*

Aug 2017   Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies, *JSM, Baltimore, MD, USA*

Dec 2015   Bayesian penalized credible region variable selection and global-local shrinkage priors, *The Institute of Statistics, Nankai University, Tianjin, China*

Contributed Talk

Aug 2016   High-dimensional linear regression via the R2-D2 shrinkage priors, *JSM, Chicago, Illinois, USA*

Aug 2015   High-dimensional variable selection via penalized credible regions with global-local shrinkage priors, *JSM, Seattle, WA, USA*

Aug 2014   Adaptive tensor regression with applications in neuroimaging data analysis, *JSM, Boston, MA, USA*

Poster Presentation

Oct 2018   Estimation of the polygenetic architecture of ten cancers and its implications for future discoveries, *ASHG, San Diego, CA, USA*

Oct 2017   Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits and implications for the future, *ASHG, Orlando, FL, USA*

PROFESSIONAL SERVICE

Editorial Activities

2020–2020   **Guest Editor**   *PLOS Computational Biology*

Journal Reviewer

Bayesian Analysis, Nature Communications

Other Professional Service

2017–2018   Committee Representative of Department of Biostatistics, School of Public Health, Johns Hopkins Postdoctoral Association

2013–2014   Student Seminar Representative, Department of Statistics, NCSU

2013–2014   Executive of STATCOM - NCSU Chapter, a student-run volunteer organization supported by ASA

MEMBERSHIP

2017–   American Society of Human Genetics (ASHG)

2016–2017   International Chinese Statistical Association (ICSA) Community

2012–2016   American Statistical Association (ASA) Community

SOFTWARE

- GENESIS: GENetic Effect-Size distribution Inference from Summary-level data