

# Yan Dora Zhang

## Contact Information

Department of Statistics and Actuarial Science  
Room 304, 3/F, Run Run Shaw Building  
The University of Hong Kong  
Pokfulam Road, Hong Kong  
Tel: (852) 3917-2473  
Email: [doraz@hku.hk](mailto:doraz@hku.hk)  
Website: <http://web.hku.hk/~doraz>

## Education

- |      |   |
|------|---|
| 2016 | <b>Ph.D. in Statistics</b> , North Carolina State University, <i>NC, USA</i><br>Thesis: <i>Bayesian Methods for High-dimensional Data</i><br>Advisor: <a href="#">Howard D. Bondell</a> |
| 2013 | <b>M.S. in Statistics</b> , North Carolina State University, <i>NC, USA</i>   |
| 2011 | <b>B.S. in Statistics</b> , Nankai University, <i>Tianjin, China</i>  |

## Research Interests

Statistical genetics, Bayesian methods, bioinformatics, biostatistics, Artificial Intelligence  
**Application areas:** cancer genetics, psychology disease, public health/biomedical research

## Professional Experience

- |           |  |
|-----------|--|
| 2019–     | <b>Assistant Professor</b><br>Department of Statistics & Actuarial Science, Faculty of Science<br>Centre for PanorOmic Sciences, LKS Faculty of Medicine<br>The University of Hong Kong, <i>Hong Kong SAR, China</i> |
| 2016–2019 | <b>Postdoctoral Fellow</b> (Mentor: <a href="#">Nilanjan Chatterjee</a> )<br>Department of Biostatistics, Bloomberg School of Public Health<br>Johns Hopkins University, <i>Baltimore, MD, USA</i>                   |

## 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

## Peer-Reviewed Publications

(\*First authors with equal contribution. †Senior authors with equal contribution.)

### First/Corresponding Author Paper

1. **Zhang YD**, Naughton BP, Bondell HD, Reich BJ. (2020). Bayesian regression using a prior on the model fit: The R2-D2 shrinkage prior. *Journal of the American Statistical Association*, (just-accepted), 1-37.
2. **Zhang YD**, Hurson AN, Zhang H, Choudhury PP, ..., Chatterjee N<sup>†</sup>, García-Closas M<sup>†</sup>. (2020). Assessment of polygenic architecture and risk prediction based on common variants across fourteen cancers. *Nature communications*, 11(1), 1-13.
3. **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.
4. **Zhang Y**, Bondell HD. (2018). Variable selection via penalized credible regions with Dirichlet–Laplace global-local shrinkage priors. *Bayesian Analysis*, 13(3), 823-844.

### Collaboration Paper

5. Meisner A, Kundu P, **Zhang YD**, Lan LV, Kim S, Ghandwani D, Choudhury PP, Berndt SI, Freedman ND, Garcia-Closas M, Chatterjee N. (2020). Combined Utility of 25 Disease and Risk Factor Polygenic Risk Scores for Stratifying Risk of All-Cause Mortality. *The American Journal of Human Genetics*, 107(3), 418-431.
6. Tubbs JD, **Zhang YD**, Sham PC. (2020). Intermediate confounding in trio relationships: The importance of complete data in effect size estimation. *Genetic Epidemiology*, 44(4), 395-399.
7. 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.
8. Pal Choudhury P\*, Wilcox AN\*, Brook MN, **Zhang Y**, Ahearn T, ..., Chatterjee N<sup>†</sup>,

García-Closas M<sup>†</sup>. (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.

## Teaching

### Lecturer

Fall 2020	STAT3604, Design and analysis of experiments, <i>HKU</i>
Spring 2020	STAT3621, Statistical data analysis, <i>HKU</i>

### Lab Instructor

2013–2014	Experimental Statistics for Biological Sciences II, <i>NCSU</i>
-----------	---

### 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, <i>NCSU</i>
-----------	---

## Presentations

### Invited Talk

Sep 2018	Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies, <i>School of Public Health, Peking University, Beijing, China</i>
Sep 2018	Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies, <i>SJTU-YALE Joint Center of Biostatistics and Data Science, Shanghai, China</i>
Sep 2018	Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies, <i>School of Statistics and Data Science, Nankai University, Tianjin, China</i>
Aug 2017	Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies, <i>JSM, Baltimore, MD, USA</i>
Dec 2015	Bayesian penalized credible region variable selection and global-local shrinkage priors, <i>The Institute of Statistics, Nankai University, Tianjin, China</i>

### Contributed Talk

Aug 2016	High-dimensional linear regression via the R2-D2 shrinkage priors, <i>JSM, Chicago, Illinois, USA</i>
Aug 2015	High-dimensional variable selection via penalized credible regions with global-local shrinkage priors, <i>JSM, Seattle, WA, USA</i>

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 implica-  
tions for the future, *ASHG, Orlando, FL, USA*

## Professional Service

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

2018–Current    **Journal Reviewer** *Nature Genetics, Nature Communications, Bayesian  
Analysis, International Journal of Biostatistics, Communications in Statis-  
tics – Theory and Methods, Human Heredity*

## 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
- [CancerEffectSize](#): Assessment of Polygenic Architecture and Risk Prediction based on Common Variants for Cancers
- [R2D2](#): Bayesian Regression Using a Prior on the Model Fit: The R2-D2 Shrinkage Prior