Lu Zhang

Science Center Room 702
One Oxford Street, Cambridge, MA 02138

lu_zhang@g.harvard.edu
(+1) 617-806-6131

EDUCATION

Harvard University, Cambridge, MA
Ph.D. in Statistics (expected in 2022)
Advisor: Lucas Janson   Aug. 2017 - present

University of Science and Technology of China (USTC), Hefei, Anhui
Bachelor of Science in Statistics
Thesis supervisor: Zemin Zheng   Aug. 2013 - Jun. 2017

RESEARCH INTERESTS

High dimensional inference, combinatorial inference, multiple testing, graphical models and reinforcement learning.
My research finds main applications in genetics and computational neuroscience.

HONORS & AWARDS

Arthur P. Dempster Award	Mar. 2021
QuantBio Student Fellowship, the NSF-Simons Center for Mathematical and Statistical Analysis of Biology at Harvard	Apr. 2019
Certificate of Distinction in Teaching Award, Harvard University	Apr. 2019
Guo Moruo Scholarship (summa cum laude), USTC	Apr. 2017
Gold Award of Outstanding Student Scholarship, USTC	Sept. 2015
National Scholarship, Chinese Ministry of Education	Sept. 2014

PUBLICATIONS & PREPRINTS

- Zhang, L. and Lu, J. (2021). StarTrek: Combinatorial Variable Selection with False Discovery Rate Control. *arXiv:2108.09904. In submission.*

- Zhang, L. and Janson, L. (2020). Floodgate: Inference for Model-Free Variable Importance. *arXiv:2007.01283. In submission.*

RECENT PROJECTS

A New Inferential Framework for Model-Free Targets
with Lucas Janson at Harvard   Sep. 2020 - present

- Working on a general inferential pipeline for a broad range of objects including conditional dependence, nonlinearity, sparsity, smoothness, privacy constraints and so on.

Model-X Approaches to Population Stratification in GWAS
with Lucas Janson at Harvard, Hilary Finucane, Alex Bloemendal and Nikolas Baya from the Broad Institute.   Mar. 2019 - present

- Developed a flexible and powerful model-X approach to select causal SNPs with error rate control guarantee, which addresses the population stratification problems in Genome-wide association study (GWAS) for admixed populations.

StarTrek: Combinatorial Variable Selection with False Discovery Rate Control
with Junwei Lu at Harvard   Feb. 2019 - present

- Established Cramér type Gaussian comparison bounds with two different norms and proved Cramér type deviation bounds for Gaussian multiplier bootstrap.

- Developed a computationally fast combinatorial variable selection procedure called StarTrek filter.

- Established asymptotic FDP/FDR control guarantee for the proposed method in high dimensional regimes for two specific examples: Gaussian graphical models and multitask regression.

- Conducted simulations to demonstrate the performance of our method in different types of networks and applied the StarTrek filter to the Genotype-Tissue Expression (GTEx) data to identify hub genes/central regulators on the gene expression network.

Floodgate: Inference for Model-Free Variable Importance
with Lucas Janson at Harvard   June. 2018 - Aug. 2020

- Proposed a model-X approach to provide asymptotic inference for the minimum mean squared error gap, an interpretable model-free measure of variable importance that is sensitive to arbitrary nonlinearities and interactions.
- Conducted simulations to demonstrate the properties of the method and applied it to the UK Biobank data to infer the strengths of dependence of platelet count on various groups of genetic mutations.

EXPERIENCE	**Summer Research Intern**	July. - Sept. 2017

Princeton University, Princeton, NJ
Advisor: Han Liu and Yang Ning.
Project: High Dimensional Inference for G-Block Latent Graphical Models.

TALKS
Floodgate: Inference for Model-Free Variable Importance, Bernoulli-IMS One World Symposium   Aug. 2020
Removing Population Structure from GWAS by Model-X Approaches, Quantitative Biology Initiative at Harvard.   Nov. 2019

CONSULTING
& TEACHING

Harvard Statistics Consulting Service	Spring & Fall 2019
STAT 111: Statistical Inference, Harvard	Spring 2019
STAT 211: Statistical Inference (graduate level), Harvard	Fall 2018 & 2019
Stochastic Processes, USTC	Fall 2016 & Spring 2017

LANGUAGES
& SKILLS

English, Mandarin.
R/Python/Mathematica/MATLAB/C/SAS/LaTeX/Bash.