

YUSI FANG

914-314-9339 | yuf31@pitt.edu | 320 N Neville St. APT15-Pittsburgh-PA-15213 | [Personal Website](#)

EDUCATION

Graduate School in Public Health, University of Pittsburgh

Pittsburgh, PA

Ph.D Candidate in Biostatistics

Expected: Apr.2022

Cumulative GPA: 4.0/4.0

Research Interest: High-dimensional Statistics, Meta-analysis, Statistical Genomics and Genetics and Machine Learning Methods

Relevant coursework: Bayesian Data Science, Applied Mixed Model Analysis, SAS, High-Dimensional Statistics, Introduction to Genomics Analysis, High-Dimensional Data with Omics Application, Advanced R Computing, Nonparametric Theory, Asymptotic Methods, Survival analysis

School of Mathematical Sciences, Xiamen University

Xiamen, China

B.S. in Mathematics and Applied Mathematics

Aug. 2013 – May 2017

Major GPA: 3.94 / 4.00 (top 2/71 in the program)

EXPERIENCE

Research Assistant

Jul. 2017 – Present

Department of Biostatistics, School of Public health, University of Pittsburgh

Pittsburgh, PA

Advisors: George Tseng and Zhao Ren

- Developed Meta-analysis methods to address challenges arising from modern high-throughput data analysis
- Developed methodology for outcome-guided disease subtyping for high-dimensional omics data
- Collaborated with biologists for data pre-processing and analysis of clinical and genomics data
- Conduct daily maintenance and management of laboratory servers

Undergraduate Research Assistant

Jul. 2016 – May 2017

School of Mathematical Sciences, Xiamen University

Xiamen, China

Advisor: Wei Liang

- Developed algorithm for kernel-based semi-supervised Bayesian quantile regression with application to cell lineage data for the detection of abnormal asynchrony of division between sister cells

SELECTED RESEARCH PROJECTS

Robust testing for combining dependent p-values

Oct. 2019 – Sep. 2020

Department of Biostatistics, University of Pittsburgh

- Proposed a family of robust tests for combining dependent p-values via transformation of heavy-tailed distribution
- Theoretically proved the asymptotic robustness of type I error control in violation of independent assumption and the optimality for detecting sparse signals of our proposed family of tests
- Simulated extensively to investigate the finite sample performance of type I error control and power of our methods comparing to other existing methods
- Applied our methods to analyze a GWAS of neuroticism

Outcome-guided disease subtyping for high-dimensional omics data

Jun. 2019 – Jul. 2020

Department of Biostatistics, University of Pittsburgh

- Developed a unified latent generative model to perform outcome-guided disease subtyping for continuous clinical outcomes with feature selection for omics data
- Modified the model for survival outcomes by embedding the accelerated failure time model
- Ran simulations to evaluate the performance of our model for survival outcomes

Optimal detection of signals under multiple scenarios via combining p-values

Jun. 2020– Present

Department of Biostatistics, University of Pittsburgh

- Developed a novel adaptive Fisher's method for p-value combination
- Theoretically proved the asymptotic optimality for signal detection of our method under both the sparse and dense cases using concentration inequalities and results from empirical process.
- Simulated comprehensively for finite sample power comparison between our method and other methods in multiple cases
- Modified our method to achieve better finite sample performance and higher computational efficiency
- Developed algorithm based on importance sampling and cross-entropy method for efficient computation and real data application

Data analysis of women over 70 years old with clinically node negative breast cancer May. 2020– Sep. 2020

Magee-Womens Research Institute and Foundation, University of Pittsburgh Medical Center

- Implemented Cox-proportional hazards model for overall survival and disease free survival
- Conducted propensity score matching over selected baseline covariates and performed down-stream analysis

TEACHING EXPERIENCE

BIOST 2094 Advanced R Computing Teaching Fellow

Department of Biostatistics, University of Pittsburgh, 2020 Spring

PUBLICATIONS

- **Fang, Yusi** & Tseng, George & Chang, Chung. (2020). Robust Testing for Combining Dependent P-values Using Transformation of Heavy-tailed Distribution. *submitted to JASA*
- Liu, Peng & **Fang, Yusi** & Ren, Zhao & Tang, Lu & Tseng, George. (2020). Outcome-Guided Disease Subtyping for High-Dimensional Omics Data. *submitted to AOAS*
- Liu, Peng & Liu, Silvia & **Fang, Yusi** & Xue, Xiangning & Zou, Jian & Tseng, George & Konnikova, Liza. (2020). Recent Advances in Computer-Assisted Algorithms for Cell Subtype Identification of Cytometry Data. *Frontiers in Cell and Developmental Biology*. 8. 234. 10.3389/fcell.2020.00234.
- Lin, Chien-Wei & Chang, Lun-Ching & Ma, Tianzhou & Oh, Hyunjung & French, Beverly & Puralewski, Rachel & Mathews, Fasil & **Fang, Yusi** & Lewis, David & Kennedy, James & Mueller (Müller), Daniel J. & Marshe, Victoria & Jaffe, Andrew & Chen, Qiang & Ursini, Gianluca & Weinberger, Daniel & Newman, Anne & Lenze, Eric & Nikolova, Yuliya & Sibille, Etienne. (2020). Older molecular brain age in severe mental illness. *Molecular Psychiatry*. 1-11. 10.1038/s41380-020-0834-1.
- Taylor, Sarah & Wield, Alyssa & **Fang, Yusi** & Bhargava, Rohit & Lang, Susan & Tseng, George & Coffman, Lan & Oesterreich, Steffi. (2020). Endocrine biomarkers in low-grade serous ovarian cancers (LGSC) and serous ovarian tumors of low malignant potential (LMP).. *Journal of Clinical Oncology*. 38. e18045-e18045. 10.1200/JCO.2020.38.15_suppl.e18045.
- Grabosch, Shannon & Bulatovic, Mirna & Zeng, Feitianzhi & Ma, Tianzhou & Zhang, Lixin & Ross, Malcolm & Brozick, Joan & **Fang, Yusi** & Tseng, George & Kim, Eun & Gambotto, Andrea & Elishaev, Esther & Edwards, Robert & Vlad, Anda. (2019). Cisplatin-induced immune modulation in ovarian cancer mouse models with distinct inflammation profiles. *Oncogene*. 38. 10.1038/s41388-018-0581-9.
- Liang, Wei & Yuxiao Yang & **Yusi Fang** & Zhongying Zhao & Jie Hu. "Bayesian Detection of Abnormal Asynchrony of Division Between Sister Cells in Mutant *Caenorhabditis elegans* Embryos." *Journal of Computational Biology* 26, no. 5 (2019): 495-505.

TECHNICAL SKILLS

Languages: English and Chinese

Statistical Packages and Software: RShiny, Rcpp, STATA, SPSS and SAS

Programming Languages: R, Python, C, Linux bash shell script and LaTeX