YUSI FANG

914-314-9339 | yuf31@pitt.edu | 320 N Neville St. APT15-Pittsburgh-PA-15213

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

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, Introduction to Bioinformatics Programming in Python, 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 methodology for combining p-values for dependent tests with heavy-tailed distributions
- Developed methodology for optimal detection of weak and sparse signals via adaptive Fisher's method
- Developed methodology for outcome-guided disease subtyping for high-dimensional omics data
- Developing methodology for uniformly optimal detection of signals under multiple scenarios via p-values combination
- Collaborated with biologists for data pre-processing and analysis of clinical and genetics data
- Laboratory server daily maintenance and management

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 and applied it to cell lineage data for detection of abnormal asynchrony of division between sister cells

Selected Reseach Projects

Combining p-values for dependent tests with heavy-tailed distributions Oct. 2019 – Sep. 2020
Department of Biostatistics, University of Pittsburgh

- Developed a family of tests for combining p-values for dependent tests with heavy-tailed distributions
- Theoretically proved the proposed family of tests enjoys asymptotic robustness under arbitrary dependency structure and optimality for detection of sparse signals
- Extensive simulations to demonstrate type-I error control and power of our methods comparing to other existing methods
- Application to a neuroticism GWAS application

Outcome-guided disease subtyping for high-dimensional omics data Jun. 2019 – Jul. 2020
Department of Biostatistics, University of Pittsburgh

- Developed graphical model for outcome-guided disease subtyping for continuous clinical outcome implemented by EM algorithm, with simultaneously feature selection for omics data and latent subtype characterization constructed from omics data
- Modified the model for survival outcome by embedding the accelerated failure time model into our model
- Running simulations to compare the performance of our model modified for survival outcome with other methods

Optimal detection of weak and sparse signals via adaptive Fisher's method Jan. 2020– Present
Department of Biostatistics, University of Pittsburgh

- Developed a novel adaptive Fisher's method for the detection of weak and sparse signals with improvement of computational efficiency
- Theoretically proved the proposed tests enjoys asymptotic optimality for detection of weak and sparse signals
- Implemented comprehensive simulations for finite sample power comparison between our proposed method and other methods
- Application to genome-wide association studies datasets

Optimal detection of signals under multiple scenarios via p-values combination Jun. 2020– Present
Department of Biostatistics, University of Pittsburgh

- Proposed a novel p-values combination method via modification on Fisher's method
- Theoretically proved the proposed method enjoys asymptotic optimality for multiple scenarios (scenarios of weak and sparse signals and scenarios of dense signals)
- Comprehensive simulations to investigate finite sample performance of our proposed method comparing to other methods in both cases with weak and sparse or dense signals
- Developing algorithm based on importance sampling and cross-entropy method for efficient computation of our proposed method

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, disease free survival
- Implemented mediation analysis
- Implemented propensity score matching over selected baseline covariates and rerun the survival analysis on the propensity score matched cohort

Teaching experience

Department of Biostatistics, University of Pittsburgh

- 2020 Spring, BIOST 2094 Advanced R Computing Teaching Fellow.

Publication

- **Fang, Yusi** & Tseng, George & Chang, Chung. (2020). Combining p-values for dependent tests with heavy-tailed distributions. *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 JASA*
- 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.
- 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: R, STATA, SPSS and SAS
Programming Languages: Python, C, shell scripting(Linux) and LATEX