Yu Zhang

yxz141130@utdallas.edu
+86-18561876616

EDUCATION

University of Texas at Dallas **Dallas, USA**
PhD Candidate in Statistics Sep,2014-Feb,2021

- GPA:3.76
- Research Area in Statistical Genomics & Bioinformatics
- Graduate Tuition Scholarship and Teaching Assistantships

University of Minnesota, Twin Cities **Minneapolis, USA**
Bachelor of Science in Mathematics-Actuarial Specialization 2010 -2014

- GPA:3.71
- Dean's List Student
- Passed Two Actuarial Exams (SOA) Exam P & Exam FM

SKILL

Statistical and Machine Learning Skills: Generalized linear models, ANOVA, Meta-analysis, Time series analysis, PCA, LDA, KNN, SVM, Random forest, Clustering, Regularization, CNN, etc.
Programming Language: R, Python, C++, SAS, Shell
Operating System: UNIX / Linux, Windows, OS X

RESEARCH INTEREST

Statistical Genomics & Bioinformatics: Genome-wide association studies, network-based modeling, next generation sequencing data analysis, structural variations detection, bacterial genomics, statistical preprocessing of genomic data, cancer genomics.

RESEARCH EXPERIENCE

Research Project: An Expectation–Maximization Algorithm for Estimating Distributions of Deletion Mutations among Bacterial Populations **2016- Present**

- Detected structural genomic variations using Breseq, DELLY, PINDEL in a simulation dataset and a deep sequence dataset for a CRISPR array of Enterococcus faecalis.
- Built a model based on Expectation–Maximization algorithm in R and Python to estimate Distributions of Deletion Mutations among Bacterial Populations.
- Used simulation data to check preference of the EM Model and applied the EM Model with strain T11 CRISPR-Cas9 genomic dataset estimating population fraction of Deletion Mutations under antibiotic selection.
- Compared the proposed method to existed tools and extended the algorithm to more real dataset.

Research Project: Markov Random Field Modeling of Biological Pathways in Genome-Wide Association Studies **2019- Present**

- Propose a Markov Random Field (MRF) model in R to incorporate single biological pathway for association analysis. Proposed an iterated conditional modes algorithm as well as a decision theoretic approach for statistical inference of each gene's association with disease.

- Inferred gene status, like disease association status based on the marginal posterior probability obtained from Bayesian analysis.
- Integrated multiple biological pathways to enhance the power of detection and control the false positive rate of the model.
- Evaluated the proposed approaches with simulations studies, Crohn's disease and Lung cancer dataset.

RELEVANT COURSEWORK

Statistical Inference, Linear Statistical Model, Advanced Probability & Stochastic Process, Advanced Statistical Method, Machine Learning, Decision Theory and Bayesian Inference, Time Series Modeling & Filtering Algorithms: Design and Analysis, Introduction to Databases

PUBLICATIONS

- Yujing Cao[§], **Yu Zhang**[§], Xinlei Wang, Min Chen. (2021). Graphical Modeling of Multiple Biological Pathways in Genomic Studies. In: Zhao Y, Chen DG (eds) Modern Statistical Methods for Health Research. Springer International Publishing. (§ co-first authors)

- **Yu Zhang**, Michael Zhang, Kelli Palmer, Min Chen, et al. Estimating Bacterial Population Structure with Application to Study Antibiotic Resistance Gene Transfer through Evolving CRISPR-Cas System of *E. faecalis.* (under revision)

TEACHING / WORKING EXPERIENCE

The University of Texas at Dallas. **2016-2020**
Teaching Assistant
Courses: Integral Calculus, Differential Calculus, Statistical Methods for Data Science, Probability and Statistics in Computer Science and Software Engineering
- Effectively two teaching problem sections each semester
- Worked under the main class instructor's supervision to lead problem review sessions
- Graded assignments and exams

Hetu International Education LLC. **2014-2020**
Chief Academic Officer
Responsible for the overall planning, management, supervision and coordination of the program academic affairs, including
- Maintain effective communication and cooperation with the host universities.
- Supervise professor recruitment and the follow-up arrangement, college course development, intern training, etc.
- Coordinate and give guidance on various tasks and emergencies during the program implementation.