

ZHEXUAN YANG

State College, PA, 16803

✉ zyang2023@yahoo.com  [linkedin.com/in/zxy028](https://www.linkedin.com/in/zxy028)  [zyang1919.github.io](https://github.com/zyang1919)

Education

Northern Illinois University

PhD in Mathematics Science (Concentration: Statistics)

Aug. 2019 – May 2025

DeKalb, Illinois

Anhui University of Science and Technology

Bachelor of Science in Applied Statistics

Sep. 2013 – Jun. 2017

Anhui, China

Experience

Assistant Teaching Professor

The Pennsylvania State University

July. 2025 – Present

State College, Pennsylvania

- Teaching course in statistics for undergraduate level student.
- Developed lecture material, homework, exams.
- Mentoring undergraduate students.

Biostatistics Intern

Jefferson Health

Jun. 2023 – Jul. 2023

Philadelphia, Pennsylvania

- Utilized Surveillance, Epidemiology, and End Results (SEER) Medicare database and performed survival analysis including Cox PH model with time-dependent covariates to assess whether patient metabolic co-morbidities (e.g., obesity and diabetes) and therapeutics (e.g., oral hypoglycemic agents and statins) have any significant associations with outcomes in patients with central nervous system tumors.
- Participated Sidney Kimmel Cancer Center's cancer trial protocol review session.

Teaching Assistant

Department of Statistics, Northern Illinois University

Aug. 2019 – Dec. 2024

DeKalb, Illinois

- Graded quizzes and homework for both graduate level and undergraduate level statistic courses.
- Ran recitation sessions to reinforce information taught in the classes and administered quizzes.
- Held office hours and tutoring sessions to assist students in need.

Consulting Assistant

Department of Statistics, Northern Illinois University

Aug. 2018 – May. 2024

DeKalb, Illinois

- Provided statistical consulting services to 20+ clients from both in-campus and off-campus students.
- Provided statistical consulting services to Chicago Pace Bus service.
- Provided statistical consulting services to Chicago Mercy Hospital.

Courses

Course Taught at the Pennsylvania State University:

- STAT 414: Introduction to Probability Theory (Fall 2025)
- STAT 416: Stochastic Modeling (Spring 2026)

Course Assisted at Northern Illinois University:

- STAT 100: Basic Statistics (Spring & Fall 2022, Spring & Fall 2023, Spring, Summer & Fall 2024)
- STAT 437: Categorical Data Analysis (Fall 2019)
- STAT 515: Computational Methods in Statistics (Fall 2021, Spring 2025)
- STAT 517: Statistical Learning (Fall 2019, Spring 2025)
- STAT 538: Time Series Analysis (Fall 2024)

Statistical Software

- **alpmixBayes**: Bayesian Estimation for Alpha-Mixture Survival Models. *The Comprehensive R Archive Network R Project.* 2025 - now
- **mmcmcBayes**: Multistage MCMC framework for differentially methylated regions detection. *The Comprehensive R Archive Network R Project.* 2025 - now

Publication

Article under review

- **Yang, Z.**, Ryu, D., Luan, F. (2026). mmcmcBayes: Multistage MCMC approach for Detecting the Differentially Methylated Regions Using R. (*Preprint*)
- **Yang, Z.** and Ryu, D. (2024). Analysis of Longitudinal Data with Outcome-Dependent Discrete Follow-up Process.
- **Yang, Z.**, Ryu, D. and Luan, F. (2024). Bayesian Multivariate Smoothing Spline for the Functional Clustering Analysis of Neural Activity.

Research

Outcome-Dependent Follow-up Study

- Proposed a follow-up process for discrete visit times with negative binomial regression under Bayesian framework of outcome dependent follow-up model.
- Applied the proposed the follow-up process model for the cardiotoxicity study and kidney function study.

DNA Methylation Analysis

- Constructed three-parameter skewed normal distribution and modified alpha-skew generalized normal distribution of modeling the DNA methylation level and proposed the Multistage MCMC within a comprehensive Bayesian framework to identify the differentially methylated genetic regions.
- Applied the proposed the model to 450K micro-array data set to identify the suspected gene symbol leading to lung cancer.

Multivariate Functional Clustering Analysis

- Constructed multivariate smoothing spline to model the neural activity in the brain of mice.
- Compared the multivariate smoothing spline with seemingly unrelated regression model.
- Constructed weighted distance matrix to perform the functional clustering analysis.

Bayesian Survival Analysis

- Constructed Bayesian α -mixture survival model.
- Developed `alpmixBayes` function for Bayesian alpha mixture survival models through CRAN

Conference

- “Analysis of longitudinal data with outcome-dependent discrete follow-up process” (poster). 32nd International Biometrics Conference (IBC), Atlanta, GA, December, 2024.
- “Bayesian Multivariate Smoothing Spline for the Functional Clustering Analysis of Neural Activity” (poster). ENAR 2025 Spring Meeting, New Orleans, Louisiana, March, 2025.

Mentorship

Schreyer Honors College

The Pennsylvania State University

- (Fall 2025) Ethan Sheaffer, STAT 414 Introduction to Probability Theory
- (Spring 2026) Daniel Brady, STAT 416 Stochastic Modeling
- (Spring 2026) Krishna Pagrut, STAT 416 Stochastic Modeling
- (Spring 2026) Andres Torres, STAT 416 Stochastic Modeling

Grant

- Travel Grant, Department of Statistics, Northern Illinois University, 2024
- Federal Grant, Higher Education Emergency Relief Fund, U.S. Department of Education, 2023

Service

- Graduate Student Council, Faculty Senate Meeting, Northern Illinois University.
- Non-tenure Promotion Committee, Committee Member, The Pennsylvania State University

Reward

- 2018-2019 Northern Illinois University, Department of Mathematical Sciences, *Certificate of Merit*.

Projects

Drug abuse study | *R*, *SAS*

- Performed three different methods (KM, Cox, AFT) on data of drug abuse study
- Utilized AIC values to determine the best survival model.

Ordinal logistic regression | *SAS*, *R*

- Performed ordinal logistic regression for both equal and unequal slope when the proportional odds assumption is not met.
- Interpreted the results based on the odds ratio.

Experiment design | *R*, *Python*

- Considered the error under various distributions when analyzing data from corn field experiments.
- Experiment included the specified fertilizer effect, column effect, and row effect.
- The normal and exponential distributions were considered for the error term.

Retrospective chart review analysis | *R*

- Using R to clean the data and perform the statistical analysis.
- Constructing contingency tables and performing binomial tests.

PCA and Sparse PCA | *R*

- Performed PCA and Sparse PCA to examine museum specimens of the pygmy marmoset to determine if there is one single species, or potentially two or more species.

COVID-19 Model Forecasting | *R*, *Python*

- Performed the PCA and negative binomial regression to do the model forecasting.
- Performed the time series analysis to forecast the daily cases.