Yating Zou

yating@live.unc.edu | (984)-291-3869 | Yating Zou
https://yatingz205.github.io/

Education

University of North Carolina - Chapel Hill, Gillings School of Public Health *Aug 2018 - Jun 2022*
Major: BS in Public Health (Biostatistics) & BS in Mathematics

Overall GPA: 3.97/4.00
Honors Carolina; Dean's List in all enrolled semesters
Graduated with Highest Distinction and Highest Honors in 1st Major

Research Interest

Precision Medicine | Semi- and Non-parametric Theory | Causal Inference | Statistical Machine Learning

Recent Coursework

Theory & **Application Focused**:

BIOS 760	- Advanced Probability and Statistical Inference I (on-going)	*Fall 2022*
BIOS 776	- Causal Inference in Biomedical Research	*Spring 2022*
STOR 515	- Dynamic Decision Analytics / Reinforcement Learning	*Spring 2022*
BIOS 740	- Precision Medicine (sit-in)	*Fall 2021*
STOR 612	- Foundations of Optimization	*Fall 2021*

Computation Focused:

BIOS 611	- Introduction to Data Science Project: Analysis of US Labor Market in the Pandemic Year	| R, Docker	*Fall 2021*
BIOS 669	- Working with Data in Public Health Research Project: Sentiment Analysis of Tweets at the onset of COVID-19	| SAS, Python	*Spring 2022*
COMP 562	- Introduction to Machine Learning	| Python	*Spring 2022*

Research & Academic Experience

Review of Causal Average Treatment Effect and Treatment Heterogeneity Estimation with Non-finite Action *Jan 2022 - Jun 2022*

(Thesis Advisor: Prof. Micheal Kosorok; Committee Members: Prof. Donglin Zeng, Dr. Jane Monaco)

Abstract: There is an increasing appeal in answering causal questions, and this interest has drawn perspectives from many related areas. However, estimation of the Average Treatment Effect (ATE) and the Conditional Average Treatment Effect (CATE) with continuous, or non-finite, action is still a relatively new setup with its specific challenges, including extra difficulty in the estimation of the Generalized Propensity Score, the diagnostics of covariate balancing, and the pathwise non-differentiability of the functional for ATE. This thesis, therefore, provides a review of recent advances on this topic. The main objectives are to (1) identify and coalesce recent related research, (2) provide a general introduction to methods such as machine learning, semi-parametric estimation, and optimization while discussing specific methods, and (3) outline the connections between existing methods of ATE and CATE estimation so that readers can better navigate themselves in this fast-expanding field.

University of North Carolina at Chapel Hill, Center for AIDS Research Chapel Hill, NC
Intern at Biostatistics Core (Supervisor: Dr. Jessica Keys, Prof. Bonnie Shook-Sa) *May 2022 - Present*

Research Project Involvement:

THINC (with Dr. Jessica Keys):

- Investigated the relationship between the level of HIV viral replication in the Cerebro-spinal Fluid (CSF) and neuropsychological outcomes for people on suppressive anti-retroviral therapy (ART).
- Cleaned and integrated biometric data from Yale, UCSF, and UNC; Measured correlation using Kendall's Tau, performed Kruskal–Wallis hypothesis test, and made figures for manuscript using R.
- Made a poster and presented the study at the 2022 Biostatistics Undergraduate Summer Internships (BUSI) Research Symposium.

R21 (with Prof. Bonnie Shook-Sa):

- Integrated and standardized household-level variables from questionnaires across 13 countries in sub-Saharan Africa.
- Built auto-regression model to identify women at highest risk for HIV infection.
- In the process of applying non-parametric estimator for Area under the Receiver Operating Characteristic (ROC) Curve (AUC) that accounts for complex sample survey design.

University of North Carolina at Chapel Hill, Department of Mathematics Chapel Hill, NC
Undergraduate Learning Assistant (Instructor: Prof. Jason Metcalfe) *Jan 2022 - May 2022*

Supportive Role:

- Assisted Professor Metcalfe in biweekly 1.5-hour student hours for two class sections for a total of 57 students on MATH 521: Advance Calculus I (or Real Analysis I), the first analysis class for undergraduate students.
- Followed up with homework, pre-class exercise, and lectures of the two sections.
- Answered questions from homework and exams on sets, convergence of sequence, series, and functions, continuity, and differentiability of one-variable functions.

UNC Lineberger Comprehensive Cancer Center, Dr. Yue Xiong's Lab Chapel Hill, NC
Statistical Assistant (Mentor: Dr. Yue Xiong) *Sep 2019 - Jul 2020*

Independent Work:

- Analyzed the distribution of mRNA fold change of 108 E3 Cullin Ring Ubiquitin Ligases (CRLs) across 33 cancer types using RNA-Seq data from the TCGA database using R.
- Conducted hypothesis tests to supplement biomedical research on the discovery of disease-linked CRL substrates and the assembly mechanisms of CRL regulators.
- Enabled interactive display of summarized statistics and plots using R to help identify marker genes for user-specified cancer types; Provided listings for the bio-pharmaceutical company Cullgen.

Supportive Role:

- Drew figures and plots using R for lab members to be used as part of their manuscripts.
- Assisted with the interpretation of statistical tests and statistical plots for lab members mostly with a biochemistry background.

Industry Experience

Fosun Pharmaceutical, International Investment Deputy Department Shanghai, China
Intern (Supervisor: Deborah Song) *Mar 2021 - May 2021*

Global M&A Project Involvement:

- Assisted in billion-dollar-scale global projects. Evaluated target companies using metrics such as the Discounted Cash Flow (DCF), Price-to-Earnings (P/E) Ratio, the Amortization (EV/EBITDA) multiple, and so on, based on data extracted from their financial statements.
- Analyzed the 5-year-trend of the above metrics of pharmaceuticals and biotech companies within the same sector as the target to support the bidding process.
- Kept track of query status from the virtual data room, a cloud for storing confidential files, during due diligence.

Independent Work:

- Created a comprehensive linked spreadsheet of 250+ major products of Fosun Pharma's subsidiaries to better align company needs with strategic planning of the International Investment Deputy Department.

Supportive Role:

- Organized and kept track of the internal project pipeline for the department.
- Extracted key information from incoming Confidential Information Memorandums (CIMs) from biotech companies and pharmaceuticals in Europe and North America.
- Followed-up with multiple project managers, collected related news and forum discussions for ongoing projects.

Meta Clinical Technology, Statistical Programming Department Chengdu, China
Intern (Supervisor: Dr. Bo Yan) *Aug 2020 - Feb 2021*

Clinical Trial Project Involvement:

- Joined 3 clinical trial studies as a Statistical Programmer and collaborated with members across sites.
- Created annotated Case Report Form (aCRF), offline listings, and SDTM and ADaM domains according to CDISC guidelines.

Independent Work:

- Developed a SAS macro to help determine the Best Overall Response (BOR) of solid tumors for any related clinical trial study based on the Response Evaluation Criteria in Solid Tumours 1.1 (RECIST1.1).
- The macro was tested on data from 3 existing clinical trials and has been a part of the company's shared macro repository since October 2020.

Leadership Role:

- Guided 7 new Statistical Programming interns who joined the company in November 2020.
- Provided updates to their lead managers on their training status, offered help on SAS programming techniques, answered questions regarding SOPs, specifications, the data flow across departments, and so on.

AsiaInfo Technologies Ltd Chengdu, China
Project Volunteer (Supervisor: Bo Gao) *Nov 2017 - Mar 2018*

Web Development Project Involvement:

- Connected SQL database to a web interface using Java and the Spring framework.
- Reported descriptive statistics of Electronic Health Records (EHR) and medical insurance data of 5 hospitals in Lianyungang, China.
- Detected abnormal prescription or insurance reimbursement activity through out-group identification.
- Participated in the UX design of the web interface.

Skills

Programming: Proficient at SAS, R, and Python; Familiar with SQL; Have experience with Java
Productivity: Proficient at LaTeX, Microsoft Office Suite including Word, Excel, and PowerPoint

Honors & Awards

2022 Salisbury Family Senior Thesis Research Award *2022*
Runner-up Prize of the 2019 UNC Carolina Challenge Makeathon *2019*

Professional Certification

GCP(Good Clinical Practice) training for Beijing Clinical Service Center (Certificate No. CSC202101-m02-561)
SAS Certified Specialist: Base Programming Using SAS 9.4

Volunteering Experience

Big Buddy for Incoming 2022 BSPH Students at UNC-CH *2022*
Girls Talk Math at UNC, Camp 2022 *2022*
Big Buddy for Incoming 2021 MS Students at UNC-CH *2021*
2019 Health Fair organized by Flourish, Student Health Action Coalition (SHAC) *2019*

Membership & Involvement

Kosorok Precision Health and Artificial Intelligence Research Lab (& the Empirical Process Theory Group)
UNC Center of AIDS Research, Biostatistics Core
American Statistical Association
Institute of Mathematical Statistics

References

Prof. Michael R. Kosorok
W.R. Kenan, Jr. Distinguished Professor, Department of Biostatistics
Professor, Department of Statistics and Operations Research
University of North Carolina at Chapel Hill
Email: kosorok@unc.edu

Prof. Michael G. Hudgens
Professor & Associate Chair, Department of Biostatistics
Director, Biostatistics Core of the UNC Center for AIDS Research (CFAR)
University of North Carolina at Chapel Hill
Email: hudgens@bios.unc.edu

Prof. Casey P. Rodriguez
Assistant Professor, Department of Mathematics
University of North Carolina at Chapel Hill
Email: crodrig@email.unc.edu