

734-968-8570
Ann Arbor
pycai@umich.edu

Peiyao Cai

Personal Website
GitHub: peiyaocai

EDUCATION

Master of Mathematics
University of Michigan, Ann Arbor

Jan 2022 - June 2023 (expected)

GPA: 4.0/4.0
Main courses: Real Analysis(A), Algebraic Combinatorics(A+).

Master of Statistics
University of Michigan, Ann Arbor

Sep 2021 - June 2023 (expected)

GPA: 3.98/4.0
Main courses: Probability Distribution Theory(A), Statistical Learning: Regression(A), Computational Methods & Tools in Statistics(A), Statistical Inference(A).

Bachelor of Finance
Peking University

Sep 2017 - June 2021

Core GPA: 3.6/4.0
Main courses: Macroeconomics(A), Data Analysis and Statistical Software(A), Mathematical Methods in Finance(A+), Analysis of Financial Time Series(A).

Exchange Student
Northeastern University

Sep 2019 - Dec 2019

Main courses: International Finance, Investment Theories.

PUBLICATIONS

Cai, P., Lu, J., Tan, K.M., 2022. Inferring Hub Nodes on Differential Score Matching Graphical Models.
(Under review in *Journal of Machine Learning Research*.)

RESEARCH EXPERIENCE

Inferring Hub Nodes on Differential Score Matching Graphical Models **Apr 2022 - Present**
Supervisors: Kean Ming Tan (University of Michigan), Junwei Lu (Harvard University) *Ann Arbor, MI*

- Proposed a novel estimator for differential edge parameters between two independent probabilistic graphical models based on score matching method, which applies to general exponential graphical models.
- Designed a joint ADMM algorithm to solve the estimation problem with extra l_1 penalties.
- Further developed a general framework for estimation, getting de-biased estimator and making inferences for differential graph.
- Proved theorems, designed numerical simulations, data applications, and wrote the paper draft independently.

High-Frequency Price Jumps and News Impacts
Supervisor: Chenxu Li (Guanghua School of Management, Peking University)

Apr 2021 - Sep 2022

- Cleaned the original high-frequency trading data sets for 44 Dow Jones Index companies with the sample size reaching million level, reorganized the data structure and constructed firm-idiosyncratic features. Constructed the well organized panel data for further analysis.
- Combined the high-frequency stock return time series and company-related news from Thomson Reuters News Database together to detect abnormal stock return residuals and their characteristics for each company.

- Applied machine learning algorithms including probit-lasso regression and random forest to study the micro-structure of high-frequency stock market.

The Tournament Promotion Model in China's Prefecture-Level Cities

Mar 2021 - Jul 2021

Supervisor: Li-An Zhou (Guanghua School of Management, Peking University)

Beijing

- Conducted a DID-based method to evaluate the Tournament Promotion Model in China's prefecture-level cities.
- Established database of government annual report at the level of different prefecture-level city throughout the whole country.

US Mutual Fund Seasonalities

Jun 2021 - Sep 2021

Supervisors: Yingguang Zhang (Peking University), Jiacui Li (University of Utah)

Beijing

- Used Python to pre-process database containing million of observations, including labeling each observation with right signs, selecting useful variables and constructing important statistical variables.
- Used R to run various tests and regressions with multiple fixed effects to detect and verify seasonalities in US stock return.

WORK EXPERIENCE

Equity Capital Market Intern

Jun 2020 - Sep 2020

Northeast Security

- Responsible for updating weekly capital market database and writing weekly reports. Designed a Python Crawler program to help collect important data on website automatically.
- Participated in a directional private placement project of a domestic investment bank with a volume of 6 billion RMB, being responsible for writing and checking important compliance documents.

Data Analytics Intern

Jan 2020 - Mar 2020

PwC China

- Extracted financial data of Chinese enterprises from the data of the past 12 years with the volume of millions. Constructed the corresponding factors according to the Beneish model.
- Applied models including logistics regression, random forest and xgb regressor to quantify the probabilities of degrading of Chinese bonds.
- Constructed 27 different factors that could influence the probability of bond degrading. Constructed ROC curve to detect the goodness of fit with AUC of 0.92.

SKILLS

Programming

Python, R, Git, L^AT_EX, Matlab, Markdown

Communication

Chinese (native), English (business)

Other

Github, Microsoft Office