Chengchun Shi

CONTACT INFORMATION

Assistant Professor of Data Science
Department of Statistics
London School of Economics and Political Science
Columbia House
69 Aldwych, London

7555-508251
c.shi7@lse.ac.uk
callmespring.github.io

EDUCATION

North Carolina State University, Raleigh, NC, USA

Ph.D., Statistics (supervised by Dr. Wenbin Lu and Dr. Rui Song) — May 2019

Zhejiang University, Hangzhou, Zhejiang Province, China

B.S., Statistics — July 2014

POSITION & EMPLOYMENT

London School of Economics and Political Science, Department of Statistics, London, UK

Assistant professor of data science — Sep 2019 – present

AWARDS & HONORS

- Institute of Mathematical Statistics (IMS) Hannan Graduate Student Travel Award, 2018
- Institute of Mathematical Statistics (IMS) Travel Award, 2017
- Paige Plagge Award, NCSU 2017
- Mu Sigma Rho, National Statistics Honor Society, NCSU 2015–Present
- Outstanding Undergraduates of Zhejiang Province, 2014
- Meritorious Winner in 2013 Mathematical Contest in Modelling, 2013
- National scholarship, China, (2%) 2012

RESEARCH INTERESTS

- Statistical methods in reinforcement learning
- Statistical analysis of complex data

PUBLICATIONS & ACCEPTED MANUSCRIPTS

- **Shi, C.**, Wan, R., Song, R., Lu, W. and Leng, L. (2020). Does the Markov decision process fit the data: testing for the Markov property in sequential decision making, *ICML*, *accepted*.
- **Shi, C.**, Song, R., Lu, W. and Li. R. (2020). Statistical inference for high-dimensional models via recursive online-score estimation, *Journal of the American Statistical Association*, *accepted*.
- **Shi, C.**, Song, R. and Lu, W. (2020). Concordance and value information criteria for optimal treatment decision, *Annals of Statistics*, *accepted*.
- **Shi, C.**, Lu, W. and Song, R. (2020). A sparse random projection-based test for overall qualitative treatment effects, *Journal of the American Statistical Association*, *accepted*.
- **Shi, C.**, Song, R., Chen, Z. and Li, R. (2019). Linear hypothesis testing for high dimensional generalized linear models. *Annals of Statistics*, **46**, 2671-2703.
- **Shi, C.**, Lu, W., and Song, R. (2019). On testing conditional qualitative treatment effects. *Annals of Statistics*, **47**, 2348-2377.
- **Shi, C.**, Lu, W. and Song, R. (2019). Determining the number of latent factors in multi-relational learning, *Journal of Machine Learning Research*, **20**, 1-38.
- **Shi, C.**, Lu, W., and Song, R. (2018). A massive data framework for M-estimators with cubic-rate. *Journal of the American Statistical Association*, **113**, 1698-1709.
- **Shi, C.**, Song, R., Lu, W., and Fu, B. (2018). Maximin projection learning for optimal treatment decision with heterogeneous individualized treatment effects. *Journal of the Royal Statistical Society, Series B*, **80**, 681-702.
- **Shi, C.**, Fan, A., Song, R., and Lu, W. (2018). High-dimensional A-learning for optimal dynamic treatment regimes. *Annals of Statistics*, **46**, 925-957.

- **Shi, C.**, Song, R. and Lu, W. (2018). Discussion of "Optimal treatment allocations in space and time for on-line control of an emerging infectious disease". *Journal of the Royal Statistical Society, Series C*, **67**, 743-789.
- **Shi, C.**, Song, R. and Lu, W. (2017). Discussion of "Random projection ensemble classification". *Journal of the Royal Statistical Society, Series B*, **79**, 959-1035.
- **Shi, C.**, Song, R. and Lu, W. (2016). Robust learning for optimal treatment decision with NP-dimensionality. *Electronical Journal of Statistics*, **10**, 2894-2921.
- Zhang, P., Qiu, Z. and **Shi, C.** (2016). simplexreg: An R Package for Regression Analysis of Proportional Data Using Simplex Distribution. *Journal of Statistical Software*, **71**, 1-21.

Manuscripts in progress

- A supervised learning framework for batch reinforcement learning, to be submitted to *NIPS*.
- Spatiotemporal causal effects evaluation: A multi-agent reinforcement learning framework, to be submitted to *NIPS*.
- Double generative adversarial networks for conditional independence testing, to be submitted to *NIPS*.
- Are the causal discoveries valid? A meta algorithm for testing large directed acyclic graph.

Teaching

- Graduate Courses
  - ST445 Managing and Visualizing Data — 2019 MT

Student Advising

- Master Students
  - Georgia Stimpson — expected to graduate 2020
  - Wen Yun Ong, Wingchi Yip, Warunya Mahaisawariya — expected to graduate 2020

Grant

- Eden Catalyst Fund, LSE Eden Center for Education Enhancement — £2,071
  - Title: Module-Level Grade Inflation Analysis at LSE.

Department Services

- Graduate admissions committee, Dept. of Statistics, LSE — 2020
- Seminar Committee (co-chair), Dept. of Statistics, LSE — 2020 LT & ST
- PhD Search Committee, Dept. of Statistics, LSE — 2020

Professional Activities

- Review service: Annals of Statistics; Journal of the American Statistical Association; Journal of Royal Statistical Society, Series B; Biometrika; Journal of Machine Learning Research; Annals of Applied Statistics; Biometrics; Statistics Surveys; Statistics Methods in Medical Research; Biostatistics; Journal of Biopharmaceutical Statistics; POLS ONE.

Talks

- February, 2020, Stat seminar at Cambridge, Cambridge, UK, "On Statistical Learning for Individualized Decision Making with Complex Data".
- January, 2020, Stat seminar at Shanghai University of Finance and Economics, Shanghai, China, "On Testing Conditional/Overall Qualitative Treatment Effects".
- October, 2019, Stat seminar at NCSU, Raleigh, NC, USA, "On Testing Conditional/Overall Qualitative Treatment Effects".
- October, 2019, Stat seminar at York University, York, UK, "On Statistical Learning for Individualized Decision Making with Complex Data".
- Auguest, 2019, Stat seminar at Shanghai University of International Business and Economics, Shanghai, China, "On Statistical Learning for Individualized Decision Making with Complex

Data".

- August 2019, JSM Meeting (invited), Denver, CO, USA, "A Sparse Random Projection-based Test for Overall Qualitative Treatment Effects".
- June 2019, ICSA Meeting (invited), Raleigh, NC, USA, "A Sparse Random Projection-based Test for Overall Qualitative Treatment Effects".
- September 2018, Biostat seminar at NCSU, Raleigh, NC, USA, "On Statistical Learning for Individualized Treatment Regime".
- August 2017, JSM Meeting, Baltimore, MD, USA, "On Testing Conditional Qualitative Treatment Effects".
- August 2016, JSM Meeting, Chicago, IL, USA, "Minimax-Angle Learning for Optimal Treatment Decision with Heterogeneous Data".
- March 2016, ENAR Meeting, Austin, TX, USA, "High Dimensional A-learning for Optimal Treatment Regime".

SOFTWARE

PYTHON MODULES

- TestMDP: Testing for the Markov property in sequential decision making. Available on GitHub: https://github.com/RunzheStat/TestMDP

R PACKAGES

- arleGP: Approximated Restricted Likelihood Estimator for Gaussian Process
- JQL: Jump Q-Learning for Individualized Interval-Valued Dose Rule (version 3.6.9) Available on CRAN: https://cran.r-project.org/web/packages/JQL/index.html
- BayesSAE: Bayesian Analysis of Small Area Estimation (version 1.0-2). Available on CRAN: https://cran.r-project.org/web/packages/BayesSAE/index.html
- simplexreg: Regression Analysis of Proportional Data Using Simplex Distributions (version 1.3). Available on CRAN: https://cran.r-project.org/web/packages/simplexreg/index.html
- ITRSelect: Variable Selection for Optimal Individualized Dynamic Treatment Regime (version 1.0-1). Available on CRAN: https://cran.r-project.org/web/packages/ITRSelect/index.html
- ITRLearn: Statistical Learning for Individualized Treatment Regime (version 1.0). Available on CRAN: https://cran.r-project.org/web/packages/ITRLearn/index.html