

Chengchun Shi

Assistant Professor of Data Science
Department of Statistics
London School of Economics and Political Science
☎ +44 7555 508251
✉ c.shi7@lse.ac.uk

All knowledge is, in the final analysis, history
All sciences are, in the abstract, mathematics
All methods of acquiring knowledge are, essentially, through statistics

Education

- Aug. 2014 **B.S. in Statistics**, Zhejiang University, Hangzhou, China.
- May 2019 **Ph.D in Statistics**, North-Carolina State University, Raleigh, USA.
 - Thesis: On Statistical Learning for Individualized Decision Making with Complex Data
 - Advisor: Dr. Wenbin Lu & Dr. Rui Song

Position & Employment

- Sep. 2019 **Assistant Professor of Data Science**, Department of Statistics, London School of Economics and Political Science, London, UK.

Honors

- Mar. 2021 **Royal Statistical Society Research Price.**
- Aug. 2020 **Finalist in Reinforcement Learning Competition Track in 2020 KDD Cup**, Ranked 7 out of more than 1000 teams.
- Mar. 2018 **Institute of Mathematical Statistics Hannan Student Travel Award.**
- April. 2017 **Institute of Mathematical Statistics Travel Award.**
- April. 2017 **Paige Plagge Citizenship Award**, NCSU, awarded to a graduate student with an outstanding academic record, who in the judgment of the committee has especially enhanced the life of fellow students with encouragement, generosity and/or humor.
- Feb. 2014 **Outstanding Undergraduates of Zhejiang Province.**
- Feb. 2013 **Meritorious Winner in 2013 Mathematical Contest in Modelling.**
- Oct. 2012 **National Scholarship**, 2%.

Research & Interests

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

Publications & Accepted Papers

- May. 2021 Wan, R., Zhang, S., **Shi, C.**, Luo, S. and Song, R. Pattern Transfer Learning for Reinforcement Learning in Order Dispatching, *IJCAI RL4ITS Workshop (spotlight)*.
- May. 2021 **Shi, C.**, Wan, R., Chernozhukov, V. and Song, R. Deeply-debiased off-policy interval estimation, *ICML (long talk, top 3% of submissions)*.
- Feb. 2021 **Shi, C.** and Li, L. Testing mediation effects using logic of boolean matrices, *Journal of the American Statistical Association*, **accepted**.
- Feb. 2021 **Shi, C.**, Song, R. and Lu, W. Concordance and value information criteria for optimal treatment decision, *Annals of Statistics*, **49**, 49-75.

- Nov. 2020 **Shi, C.**, Wan, R., Song, R., Lu, W. and Leng, L. Does the Markov decision process fit the data: testing for the Markov property in sequential decision making, *ICML*, 8807-8817.
- Jul. 2020 **Shi, C.**, Lu, W. and Song, R. Breaking the curse of nonregularity with subagging — inference of the mean outcome under optimal treatment regimes, *Journal of Machine Learning Research*, **21**, 1-67.
- Jul. 2020 **Shi, C.**, Lu, W. and Song, R. A sparse random projection-based test for overall qualitative treatment effects, *Journal of the American Statistical Association*, **115**, 1201-1213.
- Dec. 2019 **Shi, C.**, Song, R., Lu, W. and Li, R. Statistical inference for high-dimensional models via recursive online-score estimation, *Journal of the American Statistical Association*, **accepted**.
- Oct. 2019 **Shi, C.**, Song, R., Chen, Z. and Li, R. Linear hypothesis testing for high dimensional generalized linear models, *Annals of Statistics*, **46**, 2671-2703.
- Aug. 2019 **Shi, C.**, Lu, W., and Song, R. On testing conditional qualitative treatment effects, *Annals of Statistics*, **47**, 2348-2377.
- Feb. 2019 **Shi, C.**, Lu, W. and Song, R. Determining the number of latent factors in multi-relational learning, *Journal of Machine Learning Research*, **20**, 1-38.
- Oct. 2018 **Shi, C.**, Lu, W., and Song, R. A massive data framework for M-estimators with cubic-rate, *Journal of the American Statistical Association*, **113**, 1698-1709.
- Jul. 2018 **Shi, C.**, Song, R. and Lu, W. 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.
- Jun. 2018 **Shi, C.**, Fan, A., Song, R., and Lu, W. High-dimensional A-learning for optimal dynamic treatment regimes, *Annals of Statistics*, **46**, 925-957.
- May 2018 **Shi, C.**, Song, R., Lu, W., and Fu, B. Maximin projection learning for optimal treatment decision with heterogeneous individualized treatment effects, *Journal of the Royal Statistical Society, Series B*, **80**, 681-702.
- Jun. 2017 **Shi, C.**, Song, R. and Lu, W. Discussion of “Random projection ensemble classification”, *Journal of the Royal Statistical Society, Series B*, **79**, 959-1035.
- Oct. 2016 **Shi, C.**, Song, R. and Lu, W. Robust learning for optimal treatment decision with NP-dimensionality, *Electronical Journal of Statistics*, **10**, 2894-2921.
- Aug. 2016 Zhang, P., Qiu, Z. and **Shi, C.** simplexreg: An R package for regression analysis of proportional data using simplex distribution, *Journal of Statistical Software*, **71**, 1-21.

Software & Python Modules & R Packages

- Jun. 2021 **SAVE**, Sequential Value Evaluation, available on *GitHub*.
- May 2021 **D2OPE**, Deeply-Debiased Off-Policy Interval Estimation, available on *GitHub*.
- Feb. 2021 **LOGAN**, Testing Mediation Effects Using the Logic of Boolean Matrices, available on *GitHub*.
- Jun. 2020 **TestMDP**, Testing for the Markov Property in Sequential Decision Making, available on *GitHub*.
- Nov. 2018 **ITRLearn**, Statistical Learning for Individualized Treatment Regime (version 1.0), available on *CRAN*.
- Sep., 2018 **ITRSelect**, Variable Selection for Optimal Individualized Dynamic Treatment Regime (version 1.0-1), available on *CRAN*.

- Feb., 2018

BayesSAE, *Bayesian Analysis of Small Area Estimation* (version 1.0-2), available on CRAN.
- Aug. 2016

simplexreg, *Regression Analysis of Proportional Data Using Simplex Distribution*, available on CRAN.
- Apr. 2016

arleGP, *Approximated Restricted Likelihood Estimator for Gaussian Process*.

Talks

- Jun. 2021

Stat seminar at HKU (invited), HongKong, China , *Statistical Inference in Reinforcement Learning*.
- Apr. 2021

Seminar at Purdue Reinforcement Learning Group (invited), West Lafayette, USA, *Statistical Inference in Reinforcement Learning*.
- Dec. 2020

CMStatistics (invited), Online, *Does the Markov Decision Process Fit the Data: Testing for the Markov Property in Sequential Decision Making*.
- Nov. 2020

Applied Reinforcement Learning (ARL) Seminar (invited), Online, *Discussion on ‘Statistically Efficient Offline Reinforcement Learning*.
- Aug. 2020

Stat seminar at Fudan University (invited), Shanghai, China, *On Statistical Learning for Individualized Decision Making with Complex Data*.
- Aug. 2020

JSM Meeting, Online, *Does the Markov Decision Process Fit the Data: Testing for the Markov Property in Sequential Decision Making*.
- Feb. 2020

Stat seminar at Cambridge (invited), Cambridge, UK, *On Statistical Learning for Individualized Decision Making with Complex Data*.
- Jan. 2020

Stat seminar at Shanghai University of Finance and Economics (invited), Shanghai, China, *On Testing Conditional/Overall Qualitative Treatment Effects*.
- Oct. 2019

Biostat seminar at NCSU (invited), Raleigh, NC, USA, *On Testing Conditional/Overall Qualitative Treatment Effects*.
- Oct. 2019

Stat seminar at York University (invited), York, UK, *On Statistical Learning for Individualized Decision Making with Complex Data*.
- Aug. 2019

Stat seminar at Shanghai University of International Business and Economics (invited), Shanghai, China, *On Statistical Learning for Individualized Decision Making with Complex Data*.
- Aug. 2019

JSM Meeting (invited), Denver, CO, USA, *A Sparse Random Projection-based Test for Overall Qualitative Treatment Effects*.
- Jun. 2019

ICSA Meeting (invited), Raleigh, NC, USA, *A Sparse Random Projection-based Test for Overall Qualitative Treatment Effects*.
- Sep. 2018

Biostat seminar at NCSU (invited), Raleigh, NC, USA, *On Statistical Learning for Individualized Treatment Regime*.
- Aug. 2017

JSM Meeting, Baltimore, MD, USA, *On Testing Conditional Qualitative Treatment Effects*.
- Aug. 2016

JSM Meeting, Chicago, IL, USA, *Minimax-Angle Learning for Optimal Treatment Decision with Heterogeneous Data*.
- Mar. 2016

ENAR Meeting, Austin, TX, USA, *High Dimensional A-learning for Optimal Treatment Regime*.

Grants

- Apr. 2021

LSE New Research Support Fund (PI), *Sparse Reinforcement Learning for Optimal Decision Making In Mobile Health*, £19,553.

Mar. 2020

Eden Catalyst Fund, LSE Eden Center for Education Enhancement (PI),
Module-Level Grade Inflation Analysis at LSE, £2,071.

Teaching

- Graduate Courses
 - ST445 Managing and Visualizing Data 2019,2020 MT
 - ST510 Foundations of Machine Learning 2021 LT
- Undergraduate Courses
 - ST202/206 Probability, Distribution Theory and Inference 2020 – 2021
- Undergraduate Workshops
 - Ou(R) Study Groups 2020 – 2021

Student Advising

- Master Students
 - Kevin Li expect to graduate in 2021
 - Thesis Title: Reinforcement Learning for Variable Selection
 - Valentin Nachev expect to graduate in 2021
 - Thesis Title: Reinforcement Learning in Mobile Health
 - Henry Tse, Priyanshi Gupta, Anuj Srivastava expect to graduate in 2021
 - Thesis Title: Federated Learning / Reinforcement Learning in Cancer Data
 - Georgia Stimpson graduated in 2020
 - Thesis Title: Double Machine Learning for Optimal Treatment Decision Rules
 - Wen Yun Ong, Wingchi Yip, Warunya Mahaisawariya graduated in 2020
 - Thesis Title: Explore Recommender Systems and Their Embeddings

Department Services

- Graduate Admissions Committee, Dept. of Statistics, LSE 2020 – now
- Statistics Seminar Committee (co-chair), Dept. of Statistics, LSE 2020 LT & ST
- Data Science Seminar Committee, Dept. of Statistics, LSE 2021 LT & ST
- PhD Search Committee, Dept. of Statistics, LSE 2020

Professional Activities

- Review service: NeurIPS; ICML; ICLR; 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; Science China Mathematics; Australian & New Zealand Journal of Statistics; The Canadian Journal of Statistics; Science China Mathematics; Communications in Statistics - Simulation and Computation.