

Rui Song

Associate Professor
Department of Statistics
Campus Box 8203
North Carolina State University
Raleigh, NC 27695-8203
Telephone: 919-515-1955
Fax: 919-515-1169
<http://rsong.wordpress.ncsu.edu>
E-Mail: rsong@ncsu.edu

Education

B.S. 2001 Peking University Probability and Statistics
Ph.D. 2006 University of Wisconsin-Madison Statistics

Positions and Employment

2006-2008 Postdoc, Department of Biostatistics, University of North Carolina at Chapel Hill
2008-2009 Postdoc, Department of Operation Research and Financial Engineering, Princeton University
2009-2012 Assistant Professor, Department of Statistics, Colorado State University
2012- Assistant/Associate Professor, Department of Statistics, North Carolina State University

Honors

ENAR Distinguished Student Paper Award, The International Biometric Society, 2006.
The Faculty Research and Professional Development Award, North Carolina State University, 2013-2014.

Research Interests

Machine Learning, Data Science, Precision Medicine, Financial Econometrics.

Research Support

NSF grant DMS-1555244, 2016-2021, \$400,000, "CAREER: Semiparametric and Machine Learning Approaches for Big Data Challenges in Precision Medicine," PI.
NCSU Research and Innovation Seed Funding Program, 2014-2015, \$15,000, "Machine Learning Methods for Annual Influenza Vaccine Update," co-PI (PI, Dr. Osman Ozaltin from ISE department).
NCI grant P01 CA142538, 2012-2017, "Statistical Methods for Cancer Clinical Trials," Co-I.
NSF grant DMS-1309465, 2013-2016, \$109,999, "High-dimensional Multi-stage Statistical Learning with Application to Dynamic Treatment Regimens," PI.
NCSU Faculty Research and Professional Development Award Grant, 2013-2014, \$4,000, "Statistical Models, Methodologies and Related Theory For Developing Dynamic Treatment Regimens," PI.
NSF grant DMS-1007698, 2010-2013, \$100,000, "Variable Selection Methods in High Dimensional Feature Space," PI.

Peer-Reviewed Publications

- Shi, C., Lu, W. and **Song, R.** (2018+) “Determining the Number of Latent Factors in Statistical Multi-Relational Learning”. Accepted at *Journal of Machine Learning Research*.
- Jiang, B., **Song, R.** , Zeng, D. and Li, J. (2018+) “Entropy Learning for Dynamic Treatment Regimes”. Accepted at *Statistica Sinica*. (under preparation as a discussion paper)
- Shi, C., **Song, R.** , Chen, K. and Li, R. “Linear hypothesis testing for high dimensional generalized linear models ” To appear in *the Annals of Statistics*.
- Liang, S., Lu, W. and **Song, R.** , (2018+) “Deep advantage learning for optimal dynamic treatment regime.” Accepted at *Statistics and its Related Fields*.
- Shi, C., Lu, W and **Song, R.** , (2018+) “On Testing Conditional Qualitative Treatment Effects ”. To appear at *the Annals of Statistics*.
- Zhu, W., Zeng, D. and **Song, R.** , (2018+) “Proper Inference for Value Function in High-Dimensional Q-Learning for Dynamic Treatment Regimes.” To appear at *Journal of American Statistical Association*.
- Shi, C., **Song, R.** , Lu, W., (2018+) “Discussion of ‘Optimal treatment allocations in space and time for on-line control of an emerging infectious disease’ ”. To appear at *Journal of the Royal Statistical Society, Series C*.
- Shi, C., **Song, R.** , Lu, W., and Fu, B., (2018+) “Maximin-Projection Learning for Optimal Treatment Decision with Heterogeneous Data.” To appear at *Journal of the Royal Statistical Society, Series B*.
- Liang, S., Lu, W., **Song, R.** . and Wang, L. (2018+). “Sparse concordance-assisted learning for optimal treatment decision.” To appear at *Journal of Machine Learning Research*.
- Luo, S., **Song, R.** , Styner, M., Gilmore, J. and Zhu, H., (2018) “FSEM: Functional Structural Equation Models for Twin Functional Data.” *Journal of American Statistical Association*, DOI: 10.1080/01621459.2017.1407773.
- Shi, C., Fan, A, **Song, R.** and Lu, W. (2017+) “High-dimensional A-learning for Dynamic Treatment Regimes.” To appear at *the Annals of Statistics*.
- Shi, C., **Song, R.** , Lu, W., (2017) “Discussion of ‘Random Projection Ensemble Classification’.” *Journal of the Royal Statistical Society, Series B*, 79, 959-1035.
- Kang, S., Lu, W. and **Song, R.** , (2017) “Subgroup Detection and Sample Size Calculation with Proportional Hazards Regression for Survival Data”, *Statistics in Medicine*. DOI: 10.1002/sim.7441
- Jiang, R., Lu, W., **Song, R.** , Hudgens, M.G. and Naprvavnik, S, (2017) “Doubly Robust Estimation of Optimal Treatment Regimes for Survival Data.” *the Annals of Applied Statistics*, 11, 1763–1786.
- Shi, C., Lu, W. and **Song, R.** , (2017) “A massive data framework for M-estimators with cubic-rate.” *Journal of American Statistical Association*, DOI: 10.1080/01621459.2017.1360779.
- Wang, L., Zhou, Y. **Song, R.** and Sherwood, (2017) “Quantile-Optimal Treatment Regimens.” *Jour-*

- nal of American Statistical Association*, DOI: 10.1080/01621459.2017.1330204.
- Lu, Z, **Song, R.** , Zeng, D and Zhang, J., (2017)“Principal Component Adjusted Screening for High-dimensional Data.” *Computational Statistics and Data Analysis*. 110:134-144.
- Song, R.** , Luo, S., Zeng, D., Zhang, H. H., Lu, W. and Li, Z, (2017)“Semiparametric single-index model for estimating optimal individualized treatment strategy.” *Electronic Journal of Statistics*, 11(1) 364-384.
- Fan, C., Lu, W., **Song, R.** and Zhou, Y, (2016) “Concordance-assisted learning for estimating optimal individualized treatment regimes.” *Journal of the Royal Statistical Society, Series B*: doi:10.1111/rssb.12216
- Bai, X, Tsiatis, Lu, W and **Song, R.** , (2016)“Optimal treatment regimes for survival endpoints using locally-efficient doubly-robust estimator from a classification perspective.” *Lifetime Data Analysis*. 1–20.
- Shi, C. **Song, R.** and Lu, W. (2016)“Robust Learning for Optimal Treatment Regimes with NP-Dimensionality.” *Electronic Journal of Statistics*. 10(2), 2894–2921.
- Jiang, R, Lu, W, **Song, R.** , and Davidian, M., (2016) “On estimation of optimal treatment regimes for maximizing t-year survival probability.” *Journal of the Royal Statistical Society, Series B*. DOI: 10.1111/rssb.12201
- Chen, J., Liu Y., Zeng, D., **Song, R.** , Zhao, Y. and Kosorok, M.R., (2016) “Comment on ‘Bayesian Nonparametric Estimation for Dynamic Treatment Regimes with Sequential Transition Times.’ ” To appear at *Journal of American Statistical Association*.
- Fan, A., **Song, R.** and Lu, W. , (2016) “Change-Plane Analysis for Subgroup Detection and Sample Size Calculation”. *Journal of American Statistical Association*. DOI:10.1080/01621459.2016.1166115
- Laber, EB, Zhao, Y., Regh, T., Davidian, M., Tsiatis, A., Stanford, J.B., Zeng, D., **Song, R.** and Kosorok MR., (2016) “Using pilot data to size a two-arm randomized trial to find a nearly optimal personalized treatment strategy.” *Statistics in Medicine*. DOI: 10.1002/sim.6783
- Song, R.** , Banerjee, M and Kosorok, MR, (2016) “Asymptotics for change-point models under varying degrees of mis-specification.” *the Annals of Statistics*. 44(1), 153-182
- Fan, A, Lu, W. **Song, R.** , (2016) “Sequential Advantage Selection for Optimal Treatment Regimens”. *the Annals of Applied Statistics*. 10(1), 32-53.
- Song, R.** , Zeng, D, Laber, E, Zhao, Y, Yuan M, and Kosorok, MR, (2015) “On Sparse Representation for Outcome-Weighted Learning.” *STAT* DOI: 10.1002/sta4.78.
- Bradic, J and **Song, R.** , (2015) “Gaussian Oracle Inequalities for Structured Selection in Non-Parametric Cox Model.” *Electronic Journal of Statistics*. 9(1) 492-534.
- Song, R.** , Wang, W, Zeng, D and Kosorok, MR, (2014) “Penalized Q-learning for Dynamic Treatment Regimes.” *Statistical Sinica*. 25(3):901-920.
- Zhao, Y, Zeng, D, Laber, E, **Song, R.** , Yuan M, and Kosorok, MR, (2014) “Doubly Robust

- Learning for Estimating Individualized Treatment with Censored Data.” *Biometrika*. doi: 10.1093/biomet/asu050
- Goldberg, Y, **Song, R.** , Zeng, D. and Kosorok, MR (2014), “Comment on ‘Dynamic treatment regimes: technical challenges and applications’” *Electronic Journal of Statistics*.
- Song, R.** , Lu, W, Ma, S and Jeng, J, (2014) “Censored Rank Independence Screening for High-dimensional Survival Data.” *Biometrika*. doi: 10.1093/biomet/asu047
- Song, R.** , Kosorok, M.R. and Fine, J.P. (2014) Comment on ”Multiscale change point inference” by Frick, Munk and Sieling. *Journal of the Royal Statistical Society, Series B*. 76(3), 564.
- Song, R.** , Yi, F and Zou, H, (2014) “On Varying-coefficient Independence Screening for High-dimensional data.” *Statistical Sinica*. 24(4), 1735-1752.
- Goldberg, Y, **Song, R.** and Kosorok, MR (2012), “Adaptive Q-learning,” *IMS Collections: From Probability to Statistics and Back: High-Dimensional Models and Processes 9*: 150-162.
- Song, R.** , Huang, J and Ma, S (2012), “Integrative Prescreening in Analysis of Multiple Cancer Genomic Studies,” *BMC Bioinformatics* **13:168**.
- Fan, J., Feng, Y. and **Song, R.** (2011), “Nonparametric Independence Screening in Sparse Ultra-High Dimensional Additive Models,” *Journal of American Statistical Association*, **106**, 544-557.
- Zhou, H., **Song, R.** and Qin, J. (2011), “Statistical inference for a two-stage outcome-dependent sampling design with a continuous outcome,” *Biometrics*, **67**, 194-202.
- Fan, J and **Song, R.** (2010) “Sure independence screening in generalized linear models with np-dimensionality,” *the Annals of Statistics*, **38(6)**, 3567-3604.
- Song, R.** and Cai, J. (2010). “Joint covariate-adjusted score test statistics for recurrent events and a terminal event,” In Special issue: Recurrent Events of *Lifetime Data Analysis*, **16(4)**, 491-508.
- Anand, I., Carson, P., Galle E., **Song, R.** , Boehmer, J., Ghali, K.J., Jaski, B., Lindenfeld, J., O’Connor, C., Steinberg, R.J., Leigh, J., Yong, P., Kosorok, M. R., Feldman, A.M., DeMets, D. and Bristow, M. R. (2009). “Cardiac Resynchronization Therapy Reduces the Risk of Hospitalizations in Patients With Advanced Heart Failure: Results From the Comparison of Medical Therapy, Pacing and Defibrillation in Heart Failure (COMPANION) Trial,” *Circulation*, **119**: 969-977.
- Song, R.** , Zhou, H. and Kosorok, M. R. (2009). “On Semiparametric efficient inference for Two-stage outcome-dependent-sampling with a continuous outcome,” *Biometrika*, **96**, 221-228.
- Song, R.** , Kosorok, M. R. and Fine, J. P. (2009). “On Asymptotically optimal tests under loss of identifiability in semiparametric models,” *Annals of Statistics*, **37**, 2409-2444.
- Song, R.** , Kosorok, M. R. and Cai, J. (2008). “Robust covariate-adjusted log-rank statistics and corresponding sample size formula for recurrent events data,” *Biometrics*, **64**, 741-750.
- Meunier, J., **Song, R.** , Scott, R. L., Doherty, E. K., David O. E., Andersen, E. and Bruggink, G. J. (2008). “Proximate cues for a short distance migratory species: A new application of survival analysis,” *Journal of Wildlife Management*, **72(2)**: 440-448.

Song, R. , Cook, T. D., Kosorok, M. R. (2007). “What we want Versus what we can get: A closer look at endpoints for cardiovascular studies,” *Journal of Biopharmaceutical Statistics*, **18(2)**, 370-381.

Kosorok, M. R. and **Song, R.** (2007). “Inference under right censoring for transformation models with a change-point based on a covariate threshold,” *Annals of Statistics*, **35**, 957-989.

Presentations and Posters

On Statistical Reinforcement Learning for Complex Data, August 2018, DiDi AI Labs, Beijing. **Invited**

Change-plane Analysis for Subgroup Detection and Sample Size Calculation, IMPACT Symposium, September 2017, Duke University, Durham NC; March 2018, Cary NC. **Invited**

Concordance Assisted Learning for Dynamic Treatment Regimes, JSM, Aug 2016; IMS-APRM, August 2016; Fudan University, Dec. 2017; Peking University, Beijing, May 2017; University of Illinois at Chicago, Apr. 2018. **Invited**

Maximin Dynamic Treatment Regimes for Heterogeneous Data, University of Southern California, Apr. 2016 University of South Carolina, Sep. 2017. **Invited**

On Estimation of Optimal Treatment Regimes For Maximizing t-Year Survival Probability, Joint Statistics Meeting, Aug. 2015, Seattle, WA. **Topic Contributed**

Semi-nonparametric methods for personalized medicine, Sep 2014, Department of Statistics, Virginia Tech; Dec 2014, Department of Biostatistics, Washington University at St. Louis; Dec 2014, Department of Statistics, Beijing Normal University; Dec 2014, IMA-HK-IAS Joint Program on Statistics and Computational Interface to Big Data; March 2014, Department of Statistics, University of South Carolina. **Invited**

Sure Independence Screening for Gaussian Graphical Models. ICSA meeting, June 2014, Portland, OR; The 3rd Institute of Mathematical Statistics Asia Pacific Rim Meeting, July 2014 Taiwan; Joint Statistics Meeting, Aug. 2014, Boston, MA, Department of Mathematics, Dec 2014, Washington University at St. Louis. **Invited**

Sequential Advantage Selection for dynamic treatment regimens. University of Pittsburgh, 2013. University of Georgia, 2014.

Variable screening methods for high-dimensional feature space. North Carolina State University, Ohio State University, 2012.

Penalized Q-learning for dynamic treatment regimes. Columbia University, North Carolina State University, University of North Carolina at Chapel Hill, Yale University, Impact Symposium, 2012.

Integrative Prescreening in Analysis of Multiple Cancer Genomic Studies. Joint Statistics Meeting, Aug. 2011, Miami Beach, FL. **Invited**

Penalized Q-learning for dynamic treatment regimes. Department of Statistics, University of Michigan. Feb. 2011.

- Penalized Q-learning for dynamic treatment regimes. Department of Statistics, the Pennsylvania State University. Sep. 2010.
- Nonparametric Independence Screening in Sparse Ultra-High Dimensional Additive Models. Joint Statistics Meeting, Aug. 2010, Vancouver, Canada. **Invited**
- Marginal Screening in Integrative Analysis of Multiple Heterogeneous Genomic Datasets. The First International Biostatistics Symposium, Jul., 2010, Beijing, China. **Invited**
- General inference under misspecification for a class of change-point models. Joint Statistics Meeting, Aug. 2009, Washington DC. **Invited**
- On Asymptotically optimal tests under loss of identifiability in semiparametric models. ICSA meeting, Jun., 2008, Piscataway, NJ. **Invited**
- Change-point Cox model with Current status data. University of Minnesota, Columbia University, Colorado State University, University of Michigan and Michigan State University, Jan.-Feb., 2008.
- Carson, J.L., Brighton, L. E., Jaspers, I., Hazucha, M., **Song, R.** , Zhou H. Persistent Upregulation of Nasal Epithelial Ciliary Beat Frequency Among Smokers and Individuals Exposed to Environmental Tobacco Smoke. (Poster presentation) Annual meeting of the American Lung Association & American Thoracic Society, May, 2008, Toronto, Canada.
- Anand, I.S., Carson, P. C., Galle, E., **Song, R.** , Boehmer, J., Ghali, J., Jaski, B., Lindenfeld, J., O'Connor, C., Steinberg, J. S., Leigh, J., Yong, P., Kosorok, M. R., Feldman, A. M., DeMets, and D. Bristow, M. R. Cardiac Resynchronization therapy reduces the risk of hospitalizations in patients with advanced heart failure – results from the companion trial. (Poster presentation) American Heart Association Scientific Sessions, Nov., 2007, Orlando, FL.
- Statistical inference for a two-stage outcome-dependent sampling design with a continuous outcome. ICSA meeting, Jun., 2007, Raleigh, NC. **Invited**
- Sample size calculations for recurrent events data. ENAR meeting, Mar., 2006, Tampa, FL. (**ENAR Distinguished Student Paper Award**)
- Inference for the proportional odds model with a change-point based on a covariate threshold. ENAR Meeting, Mar., 2005, Austin, TX.
- Chen, Y. F. , **Song, R.** , Cho, H. , Iversen, P. W. and Heuer, J. Tree-structured Survival Model on biomarker searching in rat 5/6 renal failure survival data. (Poster) May 2005, MBSW (Midwest Biopharmaceutical Statistics Workshop), MUNCIE, IN.
- Determining Prognostic Biomarkers from an In Vivo Study of Renal Failure. August 2005, *Eli Lilly & Company*, Indianapolis, IN.
- Recurrent events methods in the analysis of clinical trials for Cardiac-resynchronization therapy. Joint Statistics Meeting, August 2005, Minnesota, MN.
- Meunier, J., **Song, R.** , Scott, R. L., Doherty, E. K., David O. E., Andersen, E., Bruggink, G. J.

and Kosorok, M. R. Fall Migration Chronology of American Woodcock (*Scolopax minor*) in the Western Great Lakes Region. Sep., 2005, Wildlife Society Conference, Madison, WI.

Boehmer, J. P., Kosorok, M. R., **Song, R.**, Carson, R., Feldman, A., Saxon, L., DeMarco, T., Yong, P. Galle, E., Ecklund, E., and Bristow, M.. Cardiac Resynchronization Therapy with and without defibrillator: the effect on Cardiac Morbidity. Nov., 2005, Scientific Sessions of American Heart Association, Dallas, TX.

Courses Taught

At North Carolina State University:

ST 361 Introduction to Statistics for Engineers
ST 508 Statistics For the Behavioral Sciences II
ST 511 Introduction to Statistics for Biological Sciences
ST 745 Analysis of Survival Data
ST 790 Financial Statistics
ST 810 Topics in High-dimensional Statistical Inference

At Colorado State University:

ST 305 Sampling Techniques
ST 501 Statistical Science
ST 640 Design and Linear Modeling
ST 740 Introduction to Empirical Processes and Semiparametric Inference

Recent Graduate (Co-)Advising Experience

Runzhe Wan (since January 2018)
Hengrui Cai (since January 2018)
Chaowen Zheng (since January 2018)
Haoyu Chen (since September 2017)
Sheng Zhang (since September 2017)
Kevin Gunn (since January 2017)
Ye Liu, PhD (since January 2017)
Liangyu Zhu, PhD (since January 2017)
Chengchun Shi, PhD (since January 2015)
Shuhan Liang, PhD (co-advised with Dr. Wenbin Lu, graduated in May 2018, first job at Google)
Ailin Fan, PhD (co-advised with Dr. Wenbin Lu, graduated in May 2016, first job at Chase)
Shikai Luo, PhD (co-advised with Dr. Subhasis Ghoshal, graduated at May 2016, first job at Quantlab)
Runchao Jiang, PhD (co-advised with Dr. Wenbin Lu, graduated in May 2015, first job at Facebook)
Neal Jorgensen, MS (graduated in December 2010 at Colorado State University)

Professional Services

Associate Editor for *Biometrics* and *Journal of American Statistical Association* (book review section).

Referee for *Annals of Statistics*, *Applied Statistics*, *Biometrics*, *Biometrika*, *Lifetime Data Analysis* (5), *Electric Journal of Statistics*, *International Journal of Biostatistics*, *Journal of American Statistical Association*, *Journal of Biopharmaceutical Statistics*, *Journal of Multivariate Analysis* (3), *Metrika*, *Nonparametric Statistics*, *Scandinavia Journal of Statistics*, *Statistical sinica* (4), *Test*, *Journal of Royal Statistical Society*, volume B.

Treasurer for Nonparametric Statistics section of ASA for year 2011-2012 and Secretary for year 2012-2013.

Graybill conference 2011 organizing committee member, Department of Statistics, Colorado State University.

Committee member for 2010 ASA Nonparametric Statistics Section Student Paper Award.

Vilas award committee member, Graduate School, University of Wisconsin, Madison, 2005.