

Curriculum Vitae of Longhai Li

30 July 2017

Biographic Information

Current Position and Contact Information

Associate Professor
Department of Mathematics and Statistics
University of Saskatchewan
106 Wiggins Road, Saskatoon SK, S7N 5E6 CANADA
Phone: 306-966-6095 (direct)
Fax: 306-966-6086 (shared)
Email: longhai@math.usask.ca
Homepage: <http://math.usask.ca/~longhai>

Birth and Citizenship

- I was born in **November 1979** in Fujian China.
- I have been a Canadian Citizen since 2012.

Academic Credentials

- Ph.D., University of Toronto, 2007, Department of Statistics, Statistics
Supervisor: [Radford Neal](#)
Thesis title: *Bayesian Classification and Regression with High-Dimensional Features*
- M.Sc., University of Toronto, 2003, Department of Statistics, Statistics
- B.Sc., University of Science and Technology of China, 2002, Department of Statistics, Statistics

Employment History

Time frame	Job Title	Organization
07/2012 – present	Associate Professor	Department of Mathematics and Statistics University of Saskatchewan, Saskatoon, CANADA
07/2007 – 07/2012	Assistant Professor	Department of Mathematics and Statistics University of Saskatchewan, Saskatoon, CANADA
11/2013 – 01/2014	Visiting Associate Professor	Department of Biostatistics, MD Anderson Cancer Center, University of Texas, Houston, USA.
11/2005 – 01/2006	Research Intern	Machine Learning and Applied Statistics Group, Microsoft Research, Redmond, USA.
07/2005 & 07/ 2006	Sessional lecturer	Department of Statistics, University of Toronto, Toronto, CANADA

Teaching Experience

Courses Taught

Undergraduate service courses

- STAT 244.3 Elementary Statistical Concepts (taught in: 2009)
- STAT 245.3 Introduction to Statistical Methods (taught in: 2011, 2012, 2014, 2015)

Undergraduate major courses

- STA 247HF Probability with Computer Applications (taught in: 2005, 2006 at the Uni. of Toronto)
- STAT 241.3 Probability Theory (taught in: 2009, 2011, 2012)
- STAT 242.3 Statistical Theory and Methodology (taught in: 2010, 2015)
- STAT 342.3 Mathematical Statistics (taught in: 2007, 2008, 2011, 2014, 2015)
- STAT 348.3 Sampling Techniques (taught in: 2012, 2014, 2016)
- STAT 442.3 Statistical Inference (taught in: 2014, 2016)

Graduate courses

- STAT 812.3 Computational Statistics (taught in: 2007, 2010, 2014, 2016). We used number STAT 846 before I formally proposed this course in 2012.
- STAT 841.3 Probability Theory (measure-theoretic probability, taught in: 2009, 2011)
- STAT 846.3 Statistical Inference (taught in: 2014, 2016)
- STAT 848.3 Multivariate Data Analysis (taught in: 2008, 2009)

Research and Scholarly Work and Activities

Research Interests

I am an applied statistician with interest in analyzing big and complex-structured data using machine learning approach. My application areas include bioinformatics, public health, and finance. The goal of my research is to solve questions with practical implication by innovative statistical and computational methods. I have research experience on these topics:

- Predictive Analysis and Feature Selection in High-dimensional Data
- Predictive Model Assessment and Comparison
- Applications of Machine Learning Techniques
- Analysis of Temporal and Spatial Data
- Mixture Modeling and Clustering

External Research Grants and Contracts

- Sub-project entitled “Genotype & Environment to Phenotype” in a Canada First Research Excellence Fund (CFREF) Project “Designing Crops for Global Food Security”, \$ 756,918 for the sub-project, 2016-2019, Investigator in a team with Prof. Kusalik as PI.
- MITACS Accelerate internship supporting research on "Applications of Neural Network Curve Fitting Methods for Least-squares Monte Carlo Simulations in Financial Risk Management", \$15,000, 2016. PI. Partner for intern: PathWise™ Solutions Group, Aon Securities, Toronto, ON.
- NSERC Individual Discovery Grant supporting research on "Bayesian Methods for High-Dimensional and Correlated Data", \$70,000, 2014 – 2019 (\$14,000/year). PI.
- NSERC ECR Supplemental grant supporting research on "Efficient Bayesian Analysis for Complex Models", \$15,000 for 2011 – 2014 (\$5000/year). PI.
- NSERC Individual Discovery Grant supporting research on "Efficient Bayesian Analysis for Complex Models", \$80,000, 2009 – 2014 (\$16,000/year). PI.
- CFI Leaders Opportunity Funding supporting "A Computer Cluster for Research on Efficient Bayesian Statistical Methods", \$160,000, 2009. PI.
- MITACS Accelerate Internship supporting research on "Clustering Analysis for Detecting the Types of Vehicles", \$15,000, 2008. Co-PI with Bill Laverty. Partner for intern: International Road Dynamics, Saskatoon, SK.

Students Supervised

Graduate students

- Xiaoying Wang, M.Sc., expected in 2018.
Thesis project: Bayesian analysis of microbiom data.
- Wei Bai, M.Sc., expected in 2018. Co-supervised with Prof. Cindy Feng.
Thesis project: Predictive p-values for longitudinal data
- Arash Shamloo, M.Sc., expected in 2017.
Thesis project: Randomized Quantile Residuals for Diagnosis of Survival Models.
- Yunyang Wang, M.Sc., Statistics, finished in November 2016.
Thesis title: *Comparison of Stochastic Volatility Models Using Integrated Information Criteria*
Employments: Statistician at Montreal office of Evidera (a PPD company), Montreal, QC, from March 2017; Intern at PathWise™ Solutions Group, Aon Securities, Toronto, ON, July – November, 2016
- Naorin Islam, M.Sc., Statistics, finished in November 2016. Co-

supervised with Prof. Shahedul Khan.

Thesis title: *Substance Abuse and Health: A Structural Equation Modeling Approach to Assess Latent Health Effects*

- Alireza Sadeghpour, M.Sc., Statistics, finished in September 2016. Co-supervised with Prof. Cindy Feng.

Thesis title: *Empirical Investigation of Randomized Quantile Residuals for Diagnosis of Non-Normal Regression Models.*

Employment: Research assistant, Department of Pharmacy and Nutrition, School of Public Health, May 2017 - present.

- Lai Jiang, Ph.D., Statistics, finished in September 2015.

Thesis title: *Fully Bayesian T-probit Regression with Heavy-tailed Priors for Selection in High-Dimensional Features with Grouping Structure.*

Employment (lastly known): Postdoctoral fellow in a research team led by Dr. Celia Greenwood at Lady Davis Institute, Jewish General Hospital, McGill University, Montreal, QB, Canada.

- Shi Qiu, M.Sc., Statistics, finished in March 2015. Co-supervised with Prof. Cindy X. Feng.

Thesis title: *Cross-validatory Model Comparison and Divergent Regions Detection using lIS and $iWAIC$ for Disease Mapping.* Employment (lastly known): Data Service Specialist at International Road Dynamics Inc., Saskatoon, SK, Canada.

- Masud Rana, M.Sc., Statistics, finished in August 2012. Co-supervised with Prof. Shahedul Khan.

Thesis title: *Spatial-Longitudinal Bent-Cable Model with an Application to Atmospheric CFC Data.*

Employment (lastly known): Ph.D. Student at the University of Victoria, BC, Canada.

- Zhengrong Li, M.Sc., Statistics, finished in June 2012.

Thesis title: *A Non-MCMC Procedure for Fitting Dirichlet Process Mixture Models.*

Employment (lastly known): Statistical Analyst at Alberta Health Services in Calgary, AB, Canada.

Undergraduate research assistants

- Jiaqi Xiao, B.Sc., Economics. Undergraduate research assistant (May – August, 2015).
- Zhouji Zheng, B.Sc., Economics. Undergraduate research assistant (June – October, 2014, July – August, 2015).
- Bei Zhang, B.Sc, Economics and Statistics. Undergraduate research assistant (May – August, 2013).

Book and Thesis

- Soltanifar, M., Li, L., and Rosenthal, J., 2010. *A Collection of Exercises in Advanced Probability Theory*, World Scientific Publishing, Singapore. The book is a solutions manual for the book *A First Look at Rigorous Probability Theory* by Rosenthal, J. (Second Edition, 2006)
- Li, L., 2007. *Bayesian Classification and Regression with High Dimensional Features*. (Ph.D. Thesis), Toronto: University of Toronto.

Research Articles

(Trainees are marked with * and embolden)

Selected Working Articles

- Li, L. and **Wang, Y. *** (2017+), Predictive Information Criteria for Comparing Bayesian Stochastic Volatility Models. A revised version of Wang's thesis.

Articles Under Review

- Feng, C. X., Li, L., **Sadeghpour, A.*** (2017+), Randomized Quantile Residuals: an Omnibus Model Diagnostic Tool with Unified Reference Distribution. Under review by *Statistical Methods for Medical Research*.
- **Jiang, L.***, Li, L., and Yao, W. (2017+). Fully Bayesian Classification with Heavy-tailed Priors for Selection in High-Dimensional Features with Grouping Structure. Under review by *Canadian Journal of Statistics*.
- Li, L. and Yao, W. (2017+), Fully Bayesian Logistic Regression with Hyper-Lasso Priors for High-dimensional Feature Selection. Under review by *Journal of Statistical Computation and Simulation*.
- **Jin, L.***, McQuillan, I., and Li, L. (2017+) Computational Identification of Harmful Mutation Regions to the Activity of Transposable Elements. Under review by *BMC Genomics*

Refereed Journal Articles, and Conference Papers

- Li, L., Feng, C.X., and, **Qiu, S.*** (2017). Estimating Cross-validators Predictive P-values with Integrated Importance Sampling for Disease Mapping Models. *Statistics in Medicine*, Volume 36, Issue 14, 2220-2236.
- Feng, C. X., **Rostami, M.***, and Li, L. (2017), Impact of Misspecified Residual Correlation Structure on the Parameter Estimates in a Shared Spatial Frailty Model. *Journal of Statistical Computation and Simulation*, 87(12), 2384-2410.
- **Jin, L.***, McQuillan, I., and Li, L. (2016), Computational Identification of Regions that Influence Activity of Transposable Elements in the Human Genome. *2016 IEEE International Conference on Bioinformatics and Biomedicine* (acceptance rate: 0.19), October 2016.
- Feng, C. X. and Li, L. (2016). Modeling Zero Inflation and Overdispersion

in the Length of Hospital Stay for Patients with Ischaemic Heart Disease, in the book *Advanced Statistical Methods in Big-Data Sciences*, Springer.

- Li, L., **Qiu, S.***, **Zhang, B.***, and Feng, C.X., 2016. Approximating Cross-validatory Predictive Evaluation in Bayesian Latent Variables Models with Integrated IS and WAIC. *Statistics and Computing*, Volume 26, Issue 4, pp 881-897.
- Yao, W. and Li, L., 2014. A New Regression Model: Modal Linear Regression. *Scandinavian Journal of Statistics*, 41(3), 656-671.
- Yao, W. and Li, L., 2014. An Online Bayesian Mixture Labeling Method by Minimizing Deviance of Classification Probabilities to Reference Labels. *Journal of Statistical Computation and Simulation*, 84 (2), 310–323.
- Li, L., 2012. Bias-corrected Hierarchical Bayesian Classification with a Selected Subset of High-dimensional Features. *Journal of American Statistical Association*, 107 (497), 120–134.
- **Sajobi, T.T.***, Lix, L. M., Dansu, B. M., Lavery, W., Li, L., 2012. Robust Descriptive Discriminant Analysis for Repeated Measures Data. *Computational Statistics & Data Analysis*, 56(9), 2782–2794.
- Khan, S. A., **Rana, M.***, Li, L., Dubin, J. A., 2012. A Comparative Case Study to Monitor and Understand Atmospheric CFC Decline with the Spatial-Longitudinal Bent-Cable Model, *International Journal of Statistics and Probability*, 1(2), 56–68.
- **Sajobi, T. T.***, Lix, L. M., Li, L., and Lavery, W., 2011. Discriminant Analysis for Repeated Measures Data: Effects of Mean and Covariance Misspecification on Bias and Error in Discriminant Function Coefficients. *Journal of Modern Applied Statistical Methods*, 10 (2), 571–582.
- Li, L., 2010. Are Bayesian Inferences Weak for Wasserman's Example? *Communications in Statistics – Simulation and Computation*, 39 (4): 655–667.
- Li, L., and Neal, R.M., 2008. Compressing Parameters in Bayesian High-order Models with Application to Logistic Sequence Models. *Bayesian Analysis*, 3(4): 793–822.
- Li, L., Zhang, J., and Neal, R.M., 2008. A method for avoiding bias from features selection with application to naive Bayes classification models. *Bayesian Analysis*, 3(1): 171–196.

Invited Talks

- Randomized Quantile Residuals: an Omnibus Model Diagnostic Tool with Unified Reference Distribution, June 2017, Seminar talk, School of Mathematical Sciences, Xiamen University, China.
- Fully Bayesian Classification with Heavy-tailed Priors for Selection in High-Dimensional Features with Grouping Structure, June 2017, Seminar

- talk, School of Mathematical Sciences, Xiamen University, China.
- Randomized Quantile Residuals: an Omnibus Model Diagnostic Tool with Unified Reference Distribution, June 2017, Seminar talk, Department of Biostatistics, Southern Medical University, Guangzhou, China.
 - Estimating Cross-validators Predictive P-values with Integrated Importance Sampling for Disease Mapping Models, June 2017, Annual Meeting of Statistical Society of Canada, University of Manitoba, Canada.
 - Fully Bayesian Classification with Heavy-tailed Priors for Selection in High-Dimensional Features with Grouping Structure, December, 2016, Seminar talk, Department of Mathematics, Wuhan University, China.
 - Cross-validators Model Comparison and Divergent Regions Detection using iIS for Disease Mapping, April 2016, Seminar of Department of Mathematics and Statistics, University of Calgary, Canada
 - Cross-validators Model Comparison and Divergent Regions Detection using iIS for Disease Mapping, April 2016, Seminar of Department of Mathematics and Statistics, University of Alberta, Canada
 - Cross-validators Model Comparison and Divergent Regions Detection using iIS for Disease Mapping, January 2016, Seminar of Department of Statistics, University of Manitoba, Canada
 - Bias-corrected Hierarchical Bayesian Classification with a Selected Subset of High-dimensional Features, August 2015, ICSA Canada Chapter Annual Meeting, University of Calgary, AB, Canada.
 - Approximating Cross-validators Predictive Evaluation in Bayesian Latent Variables Models with Integrated IS and WAIC, December 2014, Colloquia, Department of Mathematics, Tongji University, Shanghai, China.
 - An Introduction to Microarray Data. Workshop on “Statistical Issues in Biomarker and Drug Co-development”, November 2014, Fields Institute, Toronto, ON, Canada.
 - Approximating Cross-validators Predictive Evaluation in Bayesian Latent Variables Models with Integrated IS and WAIC. Statistics Seminar, April 2014, Kansas State University, Manhattan, Kansas, USA.
 - High-dimensional Feature Selection using Hierarchical Bayesian Logistic Regression with Heavy-tailed Priors. CRM-ISM-GERAD Colloque de Statistique, April 2012, McGill University, Montreal, Quebec, Canada.
 - High-dimensional Classification using Hierarchical Bayesian Polychotomous Logistic Regression Models. Colloquia talk, January, 2011, Department of Statistical and Actuarial Sciences, The University of Western Ontario, London, ON, Canada.
 - High-dimensional Classification using Hierarchical Bayesian

Polychotomous Logistic Regression Models. Colloquia talk, September 2010, Department of Statistics, Penn State University, University Park, PA, USA.

- Avoiding Bias from Feature Selection. CRISM workshop on “Bayesian Analysis of High-dimensional Data”, April, 2008, University of Warwick, Coventry, UK.
- Calculating the Confidence Intervals Using Bootstrap. Internal project meeting of a research group studying safety of nuclear power, November 2004, Ontario Power Generation, Toronto, Canada.
- Approximating Cross-validators Predictive Evaluation in Bayesian Latent Variables Models with Integrated IS and WAIC. Annual Meeting of Statistical Society of Canada, May 27, 2014, Toronto, ON, Canada.
- High-dimensional Classification using Hierarchical Bayesian Polychotomous Logistic Regression Models. The 8th ICSA International Conference, Dec. 20, 2010, Guangzhou, China.
- Discriminant Analysis for Repeated Measures Data: Effects of Covariance Structure on Bias and Error in Discriminant Function Coefficients. Annual Meeting of Statistical Society of Canada, May 24, 2010, Quebec City, QC, Canada.
- Are Bayesian Inferences Weak for Wasserman’s Example? Annual Meeting of Statistical Society of Canada, May 25, 2010, Quebec City, QC, Canada.
- Calibrating Predictions Based on a Selected Subset of Features from Bayesian Gaussian Classification Models. Annual meeting of Statistical Society of Canada, January, 2009, Vancouver, BC, Canada.
- Calibrating Predictions Based on a Selected Subset of Features from Bayesian Gaussian Classification Models. Bayesian Biostatistics Conference, January, 2009, Houston, TX, USA.
- Compressing Parameters in Bayesian High-order Models. Annual Meeting of Statistical Society of Canada, May, 2008, Ottawa, ON, Canada.
- Compressing Parameters in Bayesian Models with High-order Interactions. The 3rd Monte Carlo Workshop, Harvard University, May, 2007, Cambridge, MA, USA.
- Avoiding Bias from Feature Selection in Regression and Classification Models. Joint Statistical Meeting, August, 2006, Seattle, WA, USA.
- Analysis of Obstructive Sleep Apnea Data with Bayesian Neural Network. Annual Meeting of Statistical Society of Canada, June, 2006, London, ON, Canada.

**Contributed
Conference
Presentation**

Software Packages

- **Li, L.** (2016), R code for computing predictive p-values in disease mapping models. Publicly released R code for demonstrating the method introduced in the paper Li, Feng, and Qiu (2016+) resubmitted to *Statistics in Medicine*.
- **Li, L.** (2011), Bias-corrected Bayesian Classification with Selected Features. (994 source lines). An R add-on package. First version released on 27 July 2011. This package is for the method described in Li (2012, JASA), collected in task view "Bayesian".
- **Li, L.** (2010), C Function for Adaptive Rejection Sampling (ARS). Released through my website.
- **Li, L.** (2008), Bayesian Prediction with High-order Interactions. (2872 source lines). An R add-on package. First version released on 21 February 2008. This package is for the method described in Li and Neal (2008), collected in CRAN task views "Bayesian" and "Machine Learning".
- **Li, L.** (2007), Classification Rule Based on Bayesian Naive Bayes Models with Features Selection Bias Corrected. (653 source lines). An R add-on package. First version released on 8 November 2007. This package is for the method described in Li, Zhang and Neal (2008), collected in CRAN task views "Bayesian", "Machine Learning" and "Multivariate".
- **Li, L.** (2007), Classification Rule based on Bayesian Mixture Models with Feature Selection Bias Corrected. (756 source lines). An R add-on package. First version released on 11 November 2007. This package is for a method described in my thesis (Li, 2007), collected in CRAN task view "Bayesian".

Journal Refereeing (incomplete list)

Biometrika / BMC Bioinformatics / Computational Statistics / Computational Statistics and Data Analysis / Electronic Journal of Statistics / IEEE Transactions on Neural Networks / Statistica Sinica / Statistics in Medicine / Statistics and Computing / PLOS ONE

Grant Refereeing

- MITACS Accelerate Grant (2017)
- MITACS Accelerate Grant (2017)
- NSERC Individual Discovery Grant (2016).
- NSERC Individual Discovery Grant (2011).

Conference Organization

- Co-chair, 2017 ICSA Canada Chapter Symposium in Vancouver
- Case study judge for Statistical Society of Canada, June 2017

- Organizer, Statistics and Probability Alumni Networking Day, Nov. 2016
- Organizer, Invited session entitled “*Recent Advances in Statistical Inference Methods in Regression Models for Complex and Big Data*” for China Statistics Conference, June 2016, Qingdao, China.
- Organizer, Statistics Students Seminar Day, Department of Mathematics and Statistics, University of Saskatchewan, 2015 – 2016.
- Organizer, Statistics Seminar Series, Department of Mathematics and Statistics, University of Saskatchewan, 2014 - 2015.
- Organizer, Statistics Seminar Series, Department of Mathematics and Statistics, University of Saskatchewan, 2011 - 2012.
- Organizer, Statistics Seminar Series, Department of Mathematics and Statistics, University of Saskatchewan, 2010 - 2011.
- Organizer, Statistics Seminar Series, Department of Mathematics and Statistics, University of Saskatchewan, 2009 – 2010.

Consulting Services

- Jan. 2013 – Feb. 2013, I provided reviewing and revision services for an SHRF Establishment Grant application: “Supporting and Engaging Professionals in Palliative Care: Designing a Professional Quality of Life Intervention for Nurses in Rural and Urban Practice Settings”, which was led by Prof. Kelly Penz at the College of Nursing of the University of Saskatchewan. The application was funded by SHRF.
- Sept. 2004 – Nov. 2004, I provided data analysis and methodology advise for a research group of Ontario Power Generation (OPG) at Toronto, who studied safety of nuclear power. I conducted data analysis for estimating the probabilities that nuclear waste containers are defective and their expected lifetimes, with confidence intervals found by bootstrap methods. I was invited to give a 30-min talk to introduce the methods and analysis results at their project meeting held at Toronto in November 2004.

Graduate Committee Memberships

(excluding my own students)

Name	Degree	Program	Research Area	Grad. Year
Sajobi, T.	Ph.D	Biostatistics	Discriminant Analysis	2012
Rostami, M.	M.Sc	Biostatistics	Spatial Survival Data Analysis	2016
Fan, W.	M.Sc	Biomed. Eng.	Bioinformatics	2014
Janzen, M.R.	Ph.D	Computer Sci.	Virtual Cinematography	2012
Chisti, A.	M.Sc	Computer Sci.	Medical Image Processing	2013
Maseri, M.	Ph.D	Computer Sci.	Provenance Systems	2014
Jin, L.	Ph.D	Computer Sci.	Bioinformatics	in progress
Maleki, F.	Ph.D.	Computer Sci.	Bioinformatics	in progress

Li, Z.	Ph.D	Geography	Remote Sensing Data Analysis	2017
Yu, X.	Ph.D	Geography	Remote Sensing Data Analysis	in progress
Schmirler, M.	M.Sc	Statistics	Stochastic Simulation for DNA Unknotting	2012
Obeidat, M.	Ph.D	Statistics	Bayesian Modeling of Temporal Counts	2014
Kendall, C.	M.Sc	Statistics	Factor Analysis	2014
Dong, Y.	M.Sc	Statistics	Bayesian LASSO for ZIP Models	2016
Schmirler, M.	Ph.D	Statistics	Stochastic Simulation for DNA Unknotting	in progress
Khosa, S. K.	Ph.D.	Statistics	Survival Analysis	in progress
Achath, S.	M.Sc.	Statistics	Imprecise Probability	2017
Rijal, S.	M.Sc.	Statistics	Complex Sampling Survey	in progress
Kar, S.	M.Sc.	Statistics	Unknown	in progress

University Administration Services

University Level

- Member, Academic Programming Committee, University of Saskatchewan, 2016-2017
- Team Leader, application for accreditation of courses offered by the University of Saskatchewan that can be used toward A. STAT designation by the Statistical Society of Canada (SSC), 2015-2016.
- Member, Advisory Committee of Bioinformatics Program, University of Saskatchewan, 2015 – 2016

College Level

- Member, Search Subcommittee for a joint position in “data science/big data”, College of Arts and Science, University of Saskatchewan, 2016-2017.
- Member, Academic Programming Committee, College of Arts and Science, University of Saskatchewan, 2014 – 2015.

Department Level

- Member, Search Sub committee for APA and lecturers position, Department of Mathematics and Statistics, University of Saskatchewan, 2016-2017.
- Member, Search Committee, Department of Mathematics and Statistics, University of Saskatchewan, 2015-2016.
- Member, Graduate committee, Department of Mathematics and

Statistics, University of Saskatchewan, 2014 - 2016.

- Member, Curriculum mapping committee, Department of Mathematics and Statistics, University of Saskatchewan, 2012 - 2013.
- Member, Salary Review Committee, Department of Mathematics and Statistics, University of Saskatchewan, 2011- 2012.
- Member, Colloquium Committee, Department of Mathematics and Statistics, University of Saskatchewan, 2011 - 2012.
- Member, Subsearch Committee for a faculty position, Department of Mathematics and Statistics, University of Saskatchewan, 2010 - 2011.
- Member, Subsearch Committee for a faculty position, Department of Mathematics and Statistics, University of Saskatchewan, 2009 – 2010.