

YUCHEN WANG

+1 (573) 239-5737
ycwang0712@gmail.com

30 Gardner Road, Apartment 4F
Brookline, Massachusetts, 02445

EDUCATION

Master of Arts in Statistics, University of Missouri **January 2014 - July 2015**

- GPA 3.9 with A+ in Probability Theory, Advanced Linear Models and Bayesian Statistics.
- Took research classes in spatial statistics; developed *sppmix* (GitHub) for fitting mixture models to spatial point patterns.

Bachelor of Science in Statistics, East China Normal University **September 2009 - June 2013**

- Nominated for best Bachelor's thesis: *Residual Control Charts based on Semi-parametric Regression Models*.

Minor (36 credits) in Computer Science, Shanghai Jiao Tong University **February 2011 - June 2013**

RELEVANT EXPERIENCE

Analyst, Marketing Science, In4mation Insights, Needham, Massachusetts **September 2015 – Present**

- Analyzed consumer behavior data; developed a scoring system (method and software) by borrowing ideas from spatial statistics to define product superiority.
- Participated in projects involving market structure, segmentation, and choice modeling; Used nested logit to improve a choice model, where the choices are from two sub-markets.
- Developed an *R* interface framework for a C++ Bayesian Gibbs sampling library.

Research Assistant, University of Missouri, Columbia, Missouri **August 2014 – May 2015**

- Worked with Ph.D. students on Bayesian computation problems. By caching candidate samples in a low-acceptance MCMC algorithm and vectorization (*Fortran*), I sped up the algorithm by 30x and developed it into the *MethyBayes* package (R-CRAN).
- Parallelized a large-scale Bayes factor computation program (*Fortran*) using both *OpenMP* and *MPI*; Tuned for distributed computing on different kinds of high-performance computing clusters to achieve the best performance.

Research Assistant, East China Normal University, Shanghai, China **June 2012 - July 2014**

- Participated in a Statistical Process Control project for tobacco manufacturing. Developed quality control models for high-frequency data (GBs/hour) based on a semi-parametric regression model.
 - Led a group of 4 members, designed an interactive web application for a sequential experimental design procedure (*R/Shiny*).
-

RELATED SKILLS

Statistical Computing Expert in *R*; Developer of 15+ *R* packages (GitHub: @wangyuchen); Familiar with most quality control, MCMC and spatial statistics packages; experienced with lower-level language integration (C++/*RCpp* and *Fortran*), *shiny* web applications, and *ggplot2* and *knitr* for graphics and reports. Proficient in *SAS* procedures for statistical models like *glm*, *nlin*, *genmod*, and *mixed*.

Databases Familiar with relational databases and SQL concepts. Everyday user of *dplyr* and *tidyr*, fluent in *data.table*.