HYEONJU KIM

3900 NCTR Rd., Jefferson, AR 72079
+1(520) 465-0109 Mobile
hyeonjukm01@gmail.com

CURRENT POSITION

Visiting Scientist (Mathematical Statistician), August 2022 – Present
National Center for Toxicological Research (NCTR)-US Food & Drug Administration (FDA), Jefferson, AR

EDUCATION

Ph.D. in Statistics (Minor in Finance)**,** May 2015
Dissertation: *Probabilities of Ruin in Economics and Insurance under Light- and Heavy-Tailed Distributions.*
University of Arizona, Tucson, AZ
Adviser: Dr. Rabi N. Bhattacharya

M.S. in Applied Mathematics, December 2007
University of New Mexico, Albuquerque, NM
Thesis: *Pricing the European Call Option by the Explicit Finite Difference Method.*
Adviser: Dr. Stanly Steinberg

B.S. in Mathematics, February 2003
Dong-A University, Busan, Korea

PUBLICATIONS

1. R. N. Bhattacharya, **H. Kim**, and M. Majumdar. Sustainability in the stochastic Ramsey model. *Journal of Quantitative Economics*, 13(2):169-184, 2015. doi:10.1007/s40953-015-0020-5.
   - Developed a new method for an intractable survival probability in the economic model using multiple Chebyshev inequalities, compared with the true survival distribution by simulation in R
2. R. A. Krukowski, **H. Kim**, M. Stansbury, Q. Li, Ś. Sen, G. Farage, and D. S. West. Importance of multiple reinforcing comments and areas for change in optimizing dietary and exercise self-monitoring feedback in behavioral weight loss programs: Factorial design. *J Med Internet Res,* 22(10): e18104, 2020. doi:10.2196/18104.
   - Analyzed data using mixed effects ordinal logistic regression in STATA and wrote the statistical method section in manuscript
3. C. Trotter, **H. Kim**, G. Farage, P. Prins, R. W. Williams, K. W. Broman, and S. Sen. Speeding up eQTL scans in the BXD population using GPUs. *G3*, 2021. doi:10.1093/g3journal/jkab254.
   - Wrote down linear algebra formulas for implementation, performed data wrangling and cleaning, provided mathematical guidance to lead author
4. R. A. Krukowski, B. Johnson, **H. Kim**, Ś. Sen, and R. Homsi. A pragmatic intervention utilizing financial incentives for pregnancy weight management: A feasibility randomized controlled trial. *JMIR Formative Research,* 5(12): e30578, 2021. doi:10.2196/30578.
   - Analyzed data using linear regression, Poisson and quasi-Poisson regressions for a 2x2x2 factorial design in R and wrote the statistical method section in manuscript
5. **H. Kim**, G. Farage, J. T. Lovell, J. K. Mckay, T. E. Juenger, and Ś. Sen. Flexible multivariate linear mixed models for structured multiple traits. *bioRxiv.* doi:10.1101/2020.03.27.012690.
   - Derived mathematical formulas for algorithm, implemented Julia software, performed simulation studies, analyzed example data, and write the first draft of paper

HONORS AND AWARDS

Travel Award, Rutgers Equilibrium Theory Summer School and Workshop, June 10-13, 2019.
Travel Award, Workshop on Risk Measurement and Regulatory Issues in Business, September 11-14, 2017.
Travel Award, Institute of Mathematical Statistics (IMS) New Researchers Conference, July 27-29, 2017.
Diversity Scholarship, JuliaCon 2017, June 20-24, 2017.
Travel Award, 47th John H. Barrett Memorial Lectures, May 1-3, 2017.
Workshop Scholarship, Jackson Laboratory, October 16-22, 2016.
Thesis/Dissertation Tuition Scholarship, University of Arizona, January 2015.
Graduate College Dean's Tuition Award, University of Arizona, August 2011, August 2012, 2013 – 2014.
Graduate College Fellowship, University of Arizona, August 2011, August 2012, January 2015.
Research Assistantship, University of Arizona, 2009 - 2011, January - July 2012, 2013 – 2014.
Teaching Assistantship, University of New Mexico, 2005 – 2007.
Kappa Mu Epsilon (Mathematics Honor Society), New Mexico Alpha Chapter, 2006 – Present.

RESEARCH EXPERIENCE

Postdoctoral Research Scientist   August 2021 - Present
Gertrude H. Sergievsky Center, Columbia University Irving Medical Center
Mentor: Dr. Gao Wang

- Developing a fine mapping algorithm for high-dimensional human genetics/genomics data of extending the Bayesian linear model, Sum of Single Effects (SuSiE) developed by Dr. Gao, to Bayesian generalized linear mixed model for binary traits (Alzheimer's and diabetes) and the Julia software package.
- Validated case-only study data analysis mathematically derived from case-control study analysis for a staff biostatistician to be able to verify the equivalence by simulation for grant proposal.
- Attended the Advanced Gene Mapping Course to learn analysis of sequence and other omics data: data quality control, single variant and aggregate rare variant association analysis of whole-genome data, meta-analysis, analysis of RNA-Seq data, etc. using various available software, e.g., FasT-LMM, GCTA, REGENIE, etc. and acted as a teaching assistant for technical support for students to work on course exercises smoothly.

Postdoctoral Fellow   September 2016 – July 2021
Biostatistics, Department of Preventive Medicine, UTHSC, Memphis, TN
Mentor: Dr. Śaunak Sen

- Systems Analytics and Modeling (SAM) Project National Institute on Drug Abuse (NIDA) P30 Center: Developed a fast algorithm for Quantitative Trait Loci (QTL) analysis to detect genetic associations with phenotypes under a multivariate linear mixed model where phenotypes across high-dimensional environments are correlated. Methods employed were the Expectation Conditional Maximization (ECM) embedded in Speed restarting Nesterov's accelerated gradient method using Julia.
- Implemented extensive simulations to rigorously test the model and algorithm, optimized code to improve performance time and memory allocation, and applied to multi-traits, multi-environments, and time-valued traits for plant, mouse data, etc. on JupyterLab or REPL.
- Developed a Julia software package (CPU parallelization) for QTL analysis containing 1D, 2D genome scans, permutation test, stepwise model selection approach (forward selection and backward elimination) for finalizing multiple QTL, visualization, and relevant utility functions, etc.
- Large-scaled BXD mouse data analysis: Collaborated with a biologist in Preventive Medicine department at the UTHSC. Computed genotype probabilities and genetic relatedness matrix (GRM) and performed QTL analysis adjusting age and diet covariates using R-qtl2 library in R.

- Bloom Research Study: Collaborated with a behavioral scientist in the Department of Preventive Medicine at the UTHSC. As a consulting project, generated a descriptive table for key baseline of sociographic characteristics for participants at gestational age and histograms for 4 variables: physical activity, frequency of self-weighing, weight goal, and the number of months to meet weight goal. Perform linear regression, Poisson, and quasi-Poisson regressions for a 2x2x2 factorial design in R and computed all factorial effects with standard errors to examine the impact of frequent chances to win small, proximal incentives for meeting short-term goals as well as larger incentives for meeting long-term goals. Prepared technical reports by markdown and explained findings and suggestions for the non-quantitative collaborator in meetings. Wrote a statistical method in a manuscript and provided feedback.
- Self-monitoring Feedback in Behavioral Weight Loss Programs: Collaborated with a behavioral scientist in the Department of Preventive Medicine at the UTHSC. Analyzed archived email feedback on diet and exercise self-monitoring for participants in the iREACH3 behavioral weight control study using mixed effects ordinal logistic regression in R and STATA to assess individuals' perception of the feedback emails. Prepared technical reports by markdown and explained findings and suggestions for the non-quantitative collaborator in meetings. Wrote a statistical method in a manuscript and review the manuscript for resubmission.
- Time-valued Diversity Outbred /Heterogeneous Stock (DO/HS) Mouse Data Analysis: Collaborated with quantitative biologists at Jackson Laboratory. Calculated genotype, allele probabilities using R-qtl2 library, and performed QTL analysis of multivariate longitudinal traits using the algorithm in Advanced Computing Facility (ACF).
- Interventions Testing Program (ITP) Aging Project: Collaborated with biologists in the Department of Genetics, Genomics and Informatics at the UTHSC. Conducted ITP data analysis for identifying polymorphic alleles from GeneSeek data for 2400 genotyped mice predicting lifespan in a four-way cross using R-qtl package. Generated automated functions for data cleaning, prepared automated technical reports using markdown in a timely manner, kept reports by version control, and presented the findings for biologists.
- Serviced statistical consultations for medical researchers in the Biostats, Epidemiology, and Research Design (BERD) clinic, supervised a master-level staff statistician to write a scope of work and to complete analyses in a timely manner.

Graduate Research Assistant   January 2012 - May 2015
Department of Mathematics, University of Arizona, Tucson, AZ
Adviser: Dr. Rabi N. Bhattacharya
- Conducted theoretical and empirical research on ruin problems in economics and the general renewal model of insurance (the Sparre Andersen model) under light- and heavy-tailed distributions in finite and infinite time horizons.
- Developed a new method of piecing together in an optimal manner estimates of an intractable survival probability in the economic model (the stochastic Ramsey model) by using multiple Chebyshev inequalities in order to provide conservative lower estimates of the survival probability in both finite and infinite time horizons.
- Compared the estimates with the true survival distribution by Monte Carlo simulations in R.
- Computed finite and infinite time horizon ruin probabilities under the Cramér–Lundberg model of insurance by Monte Carlo simulations using R to check the validity of using the asymptotic theory for moderate initial assets.

Graduate Research Assistant   August 2010 - May 2011
Department of Mathematics, University of Arizona, Tucson, AZ
Supervisor: Dr. Daniel Madden
- Collaborated with 3 mathematicians in designing mathematics problem sets used for final projects targeted for middle school mathematics teachers with elementary certification pursuing master's degrees

in Middle School Mathematics Teaching Leadership in the program of Arizona Teacher Initiative funded by a 5-year, $4.8M National Science Foundation Math and Science Partnership grant.

- Collected data sets from open databases, identified sources of information applicable to developing mathematics questions, and synthesized relevant data sets for data analysis.
- Analyzed the processed data using appropriate graphing and statistical techniques and tested the statistical validity of candidate questions in Excel or R.
- Discussed the findings for further refinements of the problem sets in regular group meetings.

Graduate Research Associate   August 2009 - May 2010
Statistics Consulting Laboratory, BIO5 Institute, University of Arizona, Tucson, AZ
Supervisor: Dr. Dean Billheimer

- Developed an R statistical software package for data cleansing using mid-sized data in order to facilitate data analysis.
- Conducted statistical analysis using appropriate graphing, statistical techniques and testing procedures in R for client needs.
- Prepared automated technical reports on results for non-quantitative clients with R Sweave integrated into LaTeX documents in a timely manner.

TEACHING EXPERIENCE

Postdoctoral Fellow   Spring 2017 – Summer 2021
Biostatistics, Department of Preventive Medicine, UTHSC, Memphis, TN
Mentoring:

- Guided a summer intern (Ye Eun Bae, PhD student at Florida State University) for daily checkup, implementing QTL analysis using a R/qtl package, and resolving technical issues arising from network, clusters, etc. (Summer 2021)
- Taught a scientific programmer (MS. Chelsea Trotter) statistical model, animal data structure to write GPU computation code for univariate genome scan, permutation test for large-scale mouse data and to build up an interface between GeneNetwork (a combined database and open-source bioinformatics data analysis software for systems genetics) and Julia, R (Spring 2018 – Spring 2021).
- Helped a staff statistician (MS. Jiajing Wang) understand theoretical statistics and probability for PhD research and advised R-code for debugging for an hour per week (Spring 2017).

Teaching:

- Translated SAS-code to R-code for Survival Analysis (BIOE 868, Supervisor: Dr. Fridtjof Thomas) and generate R-markdown tutorials to help students improve coding skills (Spring 2019).

Teaching Assistant   August 2014 – December 2014
University of Arizona, Department of Mathematics, Tucson, AZ

- Taught College Algebra (Math 112, 1 section, 3 hrs. /wk.) to 30 undergraduates.
- Prepared lectures, designed midterm and final questions, held office hours and tutoring table on a weekly basis, graded homework, and attended a weekly TA training seminar.
- Developed different teaching skills to explain abstract concepts concretely, motivated students to do group discussions for engagement in class and responded appropriately to students' feedback for improving teaching performance in class.

Graduate Associate   January 2014 - May 2014
University of Arizona, Department of Marketing, Tucson, AZ
Supervisor: Dr. Shshila Umashankar

- Taught and tutored Analytical Methods for Business in laboratory sessions (BNAD 277, 7 labs, 10 hrs. /wk.) to 25-40 undergraduates/lab.

- Provided graders with solutions of homework and specific instructions on homework grading, prepared midterms and a final exam, and monitored a large class during the exams.

Teaching Assistant   August 2005 - May 2007
University of New Mexico, Department of Mathematics and Statistics, Albuquerque, NM
- Taught College Algebra (Math 121, 1 section, 3 hrs. /wk.), Elements of Calculus I (Math 180, 1 section, 3 hrs. /wk.), and Calculus I recitation classes (Math 162, 2 sections, 3 hrs. /wk.) to 30-35 undergraduates.
- Prepared lectures, weekly quizzes, and midterms held office hours, and weekly Algebra Table or Calculus Tutoring Table, graded homework, quizzes and exams, and attended weekly TA training seminars.
- Graded Linear Algebra with Applications (Math 314, 3 sections, 20 hrs. /wk.).

CONFERENCES/WORKSHOPS PRESENTED/ATTENDED

- Advanced Gene Mapping Course, Rockefeller University, New York, NY, January 10-14, 2022.
- Contributed talk: "A General Multivariate Linear Mixed Model for Detecting Gene by Environment Interactions", Joint Statistical Meetings 2019, Denver, CO, July 27-August 1, 2019.
- Rutgers Equilibrium Theory Summer School and Workshop, Rutgers University, New Brunswick, NJ, June 10-13, 2019.
- Poster: "Multivariate Linear Mixed Models for Detecting GxE", 16th Annual Meeting of Complex Trait Community in collaboration with the Rat Genomics Community (CTC-RG2018), University of Glasgow, Glasgow, UK, June 20-22, 2018.
- Invited talk: "Multivariate Linear Mixed Models for Detecting GxE", International Chinese Statistical Association (ICSA) 2018 Applied Statistics Symposium, New Brunswick, NJ, June 14-17, 2018.
- Poster: "A Multivariate Linear Mixed Model for Detecting GxE", Population, Evolutionary and Quantitative Genetics Conference, Madison, WI, May 13-16, 2018.
- Poster: "Sustainability in the Stochastic Ramsey Model", Workshop on Risk Measurement and Regulatory Issues in Business, Centre de Recherché Mathématiques (CRM), Université de Montréal, Montréal, QC, Canada, September 11-14, 2017.
- Poster: "Sustainability in the Stochastic Ramsey Model", Workshop on Environmental Risk Modeling and Extreme Events, Centre de Recherché Mathématiques (CRM), Université de Montréal, Montréal, QC, Canada, August 28-31, 2017.
- Poster: "QTL analysis for GxE using a matrix-variate linear mixed model", Institute of Mathematical Statistics (IMS) New Researchers Conference, Johns Hopkins University, Baltimore, MD, July 27-29, 2017.
- JuliaCon 2017, University of California, Berkeley, CA, June 20-24, 2017.
- 15th Annual Meeting of the Complex Trait Community in collaboration with the Rat Genomics Community (CTC-RG2017), FedEx Institute of Technology, Memphis, TN, June 13-17, 2017.
- 47th John H. Barrett Memorial Lectures, University of Tennessee, Knoxville, TN, May 1-3, 2017.
- Workshop on the Interface of Statistics and Optimization (WISO), Duke University, Durham, NC, February 8-10, 2017.
- Workshop: Short Course on Systems Genetics, Jackson Laboratory, Bar Harbor, ME, October 16-22, 2016.
- News and Finance Conference: How the Media Affects Markets organized by the Program for Financial Studies, Columbia Business School, New York, NY, March 11, 2016.
- Allied Social Science Associations (ASSA) 2016 Annual Meetings, San Francisco, CA, January 3-5, 2016.
- SQA Half Day Conference: Insights for Investors in 2015, New York, NY, November 11, 2015.

PROFESSIONAL ASSOCIATIONS

American Statistics Association (ASA)	2010 - Present
Society for Industrial and Applied Mathematics (SIAM)	2014 - Present
Institute of Mathematical Statistics (IMS)	2015 - Present
International Association for Quantitative Finance (IAQF)	2015 - Present
American Economic Association (AEA)	2015 - Present
Association for Women in Mathematics (AWM)	2018 - 2020
Genetics Society of America (GSA)	2018 - 2021
American Risk and Insurance Association (ARIA)	2015 - 2017
Society of Quantitative Analysts (SQA)	2015 - 2017

SKILLS

- **Statistics & Applied Mathematics**: Stochastic Process/Modeling, Stochastic Differential Equation, (Generalized) Linear Mixed Models, Optimization, Empirical Bayes, Survival Analysis, Machine Learning, General Linear Models, Theoretical/Computational Statistics, Extreme Value Theory, Econometric Modeling.
- **Finance**: Dynamic Asset Pricing Theory, Fixed Income: Markets, Instruments & Strategies, Interest Rate Models, Advanced Risk Management & Derivatives.
- **Technical:** Julia, R, Octave, MATLAB, STATA, Python, SAS, LaTeX, Markdown, Sweave, MS Office, Mercurial, Git.

SOFTWARE REPOSITORY

- FlxQTL.jl: https://github.com/hkim89/FlxQTL.jl
- SuSiEGLMM.jl: https://github.com/hkim89/SuSiEGLMM.jl

LANGUAGES

Native in Korean. Professional working proficiency in English.

REFERENCES

Provided upon request.