

# Beomjo Park, PhD

✉ [beomjop@gmail.com](mailto:beomjop@gmail.com) | [in beomjop](#) | [G beomjopark](#) | [Website](#) | 📍 Mountain View, CA

## PROFESSIONAL SUMMARY

Research-focused Data Scientist developing and implementing quantitative methodologies, from experiment and measurement design to statistical (causal) inference, for optimization and analysis of large-scale data in diverse applications.

## EXPERIENCES

- Data Scientist, Research, Google, Mountain View, CA** Aug 2023 – Present
- Developed measurement and optimization frameworks for YouTube Ads video creative generation and enhancement, achieving a balance between revenue uplift and serving resource costs.
  - Design end-to-end diagnostic methods to identify and pinpoint systematic inefficiency at the most granular level within the YouTube Ads serving funnel.
- Data Scientist Intern, Google, Mountain View, CA** May 2022 – Aug 2022
- Quantified the causal mediation effect of Ad quality on user engagement by analyzing YouTube session data.
  - Designed an improved counter metric accounting for the causal structure of user engagement.
  - Selected as an outstanding intern presentation in YouTube Ads QUADS team.
- Graduate Researcher, Carnegie Mellon University, Pittsburgh, PA** Jan 2019 – Aug 2023
- Developed robust universal inference methods for constructing batch and sequential confidence sets accounting for model misspecification, lack of regularity, and data contamination [6][7]. Demonstrated the practical application of the methods in causal discovery using causal linear structural equation model.
  - Constructed spatio-temporal heat transport field of global oceans from large-scale autonomous profiling float observations that are partially missing, heterogeneous, and sparsely distributed [1].
  - Delivered a scientific insight into climatological phenomena by collaborating with domain scientists.
- Graduate Researcher, Korea University, Seoul, Korea** Sep 2016 – Jul 2019
- Researched hierarchical Bayesian model representations and nonparametric mixture processes.
  - Tailored above methods to a meta-analysis in medical studies [2] and functional data analysis.
  - Implemented and assessed model selection criteria for scalable Variational inference [3][5].
  - Enhanced and reviewed the end-user application and built discipline-specific worked examples [4].
- Research Assistant, NCSOFT (NLP lab), Korea** Jul 2016 – Dec 2016
- Extracted key features and importance affecting individual players' seasonal performance by analyzing Korea Baseball Championship historical data with a hierarchical Bayesian latent model.

## HONORS AND AWARDS

- Student paper award by Statistical Learning and Data Science Section, *American Statistical Association*. Aug 2023
- Outstanding intern presentation by YouTube Ads QUADS team, *Google*. Aug 2022
- PhD TA of the year (2021-2022) by Dept. of Statistics, *Carnegie Mellon University*. May 2022
- SG graduate student paper presentation award (3<sup>rd</sup> place) by the *Korean Statistical Society*. Nov 2017
- National Science Scholarship by *Korea Student Aid Foundation*. Fall 2015
- High Honors (with scholarship) by *Korea University*. Mar 2011 – Aug 2016

## EDUCATION

- Ph.D. in Statistics & Data Science, *Carnegie Mellon University*, Pittsburgh, PA Aug 2018 – Aug 2023
- M.Sc. in Statistics, *Korea University*, Seoul, Korea Sep 2016 – Aug 2018
- B.Sc. in Industrial Engineering & B.Ec. in Statistics, *Korea University*, Seoul, Korea Mar 2010 – Aug 2016

## PUBLICATIONS

---

- [1] **Park, B.**, Kuusela, M., Giglio, D. & Gray, A. (2022) Spatio-Temporal Local Interpolation of Global Ocean Heat Transport using Argo Floats: A Debiased Latent Gaussian Process Approach. *Annals of Applied Statistics*
- [2] Jo, S., **Park, B.**, Chung, Y., Kim, J., Lee, E. & Choi, T. (2021) Bayesian semiparametric mixed effects models for meta-analysis of literature data: An application to cadmium toxicity studies. *Statistics In Medicine*.
- [3] Lim, D., **Park, B.**, Nott, D. J., Choi, T., & Xueue, W. (2020) Sparse signal shrinkage and outlier detection in high-dimensional quantile regression with variational Bayes. *Statistics and Its Interface*.
- [4] Jo, S., Choi, T., **Park, B.**, & Lenk, P.J. (2019) bsamGP: An R Package for Bayesian Spectral Analysis Models using Gaussian Process Priors. *Journal of Statistical Software*.
- [5] Ong, V. M., Mensah, K. M., Nott, D. J., Jo, S., **Park, B.**, & Choi, T. (2017) A variational Bayes approach to a semiparametric regression using Gaussian process priors. *Electric Journal of Statistics*.

## PREPRINTS

---

- [6] **Park, B.**, Balakrishnan, S. & Wasserman, L. (2023+) Robust Universal Inference For Misspecified Models. *Under Revision, Biometrika*. [arXiv preprint](#)  
► Winner of ASA Statistical Learning and Data Science Student Paper Award.
- [7] **Park, B.**, Balakrishnan, S. & Wasserman, L. (2023) Nonparametric Functional Estimation under Contamination.

## CONFERENCE PRESENTATIONS

---

- Park, B.**, Balakrishnan, S. & Wasserman, L. (Aug. 2023) Robust Universal Inference. *JSM*, Toronto, Canada.
- Park, B.**, & Kuusela, M. (Aug. 2020) Spatio-Temporal Local Interpolation for Quantifying Global Ocean Heat Transport from Autonomous Observations. (Contributed Talk) *JSM*, virtual.
- Park, B.**, & Choi, T. (Jul. 2018) Bayesian Hierarchical Varying-coefficient Mixed Model. (Poster session) *The third East Asia Chapter of ISBA Conference*, Seoul, Korea.
- Park, B.**, & Choi, T. (Nov. 2017) Bayesian Multivariate Hierarchical Semiparametric Mixed Model with Gaussian Process Priors. *The Korean Statistical Society Autumn Conference*, Seoul, Korea.  
► 3<sup>rd</sup> place on SG Graduate Student Paper Presentation Award.

## TEACHING EXPERIENCES

---

- Teaching Assistant, Carnegie Mellon University**
- AUG 2018 – JULY 2023 Carnegie Mellon Undergrad Research Experience program (Sports Analytics & Optum Camp), Introduction to Statistical Inference (head TA), Advanced Methods for Data Analysis (head TA), Probability and Mathematical Statistics (head TA), Statistical Graphics and Visualization, Statistical Computing.
- Teaching Assistant, Korea University**
- SEP 2016 – AUG 2017 Mathematical Statistics, Research Methods II, Statistical Computing Methods, Elementary Computational Statistics.

## LANGUAGES

---

- Languages: English (Proficient), Korean (Native)
- Programming: R<sup>†</sup>, Python, MATLAB, C++  
† Current maintainer of bsamGP package on [CRAN](#).