HUNYONG CHO

hunycho@live.unc.edu ◇ https://tarheels.live/hunyongcho

RESEARCH INTERESTS

Statistical methods and theory—Precision medicine, machine learning, survival analysis, empirical processes, causal inference, semiparametric inference, categorical data analysis; **Machine learning methods and applications**—random forest, deep learning, mHealth, dynamic treatment regime; **Omics**—scRNA-seq data, microbiomics, oral health research

EDUCATION

Ph.D. in Biostatistics, **University of North Carolina, Chapel Hill** — August 2021
M.S. in Applied Statistics, **Loyola University Chicago** — December 2016
B.A. in Business, **Yonsei University, Korea** — August 2005

PROFESSIONAL EXPERIENCE IN STATISTICS

Postdoctoral fellow, **University of North Carolina, Chapel Hill** — Sep 2021–present
Graduate research assistant, **University of North Carolina, Chapel Hill** — Sep 2017–Sep 2021

SELECTED HONORS AND AWARDS

Inducted into Delta Omega Honor Society in Public Health — 2021

JSM Paper Award — 2021
Biometrics Section, Joint Statistical Meeting 2021, American Statistical Association

Public Health Student Scholarship — 2020-2021
Gillings School of Global Public Health, University of North Carolina, Chapel Hill

Special Commendation for Doctoral Comprehensive Exam (Theory) — September 2017
Department of Biostatistics, University of North Carolina, Chapel Hill

Winner of LUC Hackathon — October 2014
Department of Computer Science and Department of Mathematics and Statistics, Loyola University Chicago

Bronze medal, National Mathematics Competition — May 2000
Korea University, Seoul, Korea

RESEARCH EXPERIENCE

University of North Carolina, Chapel Hill — 2021 - Present
Postdoctoral research (Advisor: Dr. Michael R. Kosorok) — *Chapel Hill, NC*

- Change-plane limiting distribution theory and non-convex optimization
- Outcome weighted learning weak convergence theory
- Infinite horizon reinforcement learning method development for type-I diabetes patients

University of North Carolina, Chapel Hill — 2017 - 2021
Doctoral research (Advisor: Dr. Michael R. Kosorok) — *Chapel Hill, NC*

- Development of the dynamic treatment regime estimator for survival outcomes
- Development of the random survival forest estimator for interval censored data

· Development of the semiparametric G-estimation method using random forests and kernel ridge regression
· Analysis of asymptotic behavior of the influence functions of the treatment regime values

University of North Carolina, Chapel Hill 2017 - 2021
Doctoral research (Advisor: Dr. Di Wu) *Chapel Hill, NC*

· Development of gene set testing methods based on bivariate zero-inflated negative binomial models
· Integrative analysis of metagenomics and metatranscriptomics on early childhood dental caries

Loyola University Chicago August 2014 - December 2015
Research Assistant (Advisor: Dr. Gregory J. Matthews) *Chicago, IL*

· Analysis of confidence interval methods of the area under the receiver operating characteristic curve (AUC) in the presence of missing data
· Model building of the online advertisement click-through rates based on random forests and ridge regression

TEACHING EXPERIENCE

University of North Carolina, Chapel Hill Fall 2020
Teaching assistant (Instructor: Dr. Michael R. Kosorok) *Chapel Hill, NC*

· BIOS760 Advanced Probability and Statistical Inference I
· Weekly office hours and grading

University of North Carolina, Chapel Hill Fall 2018, 2019
Tutoring service for Department of Biostatistics *Chapel Hill, NC*

· Supported graduate students preparing for their comprehensive exams

University of North Carolina, Chapel Hill April 2019
A guest lecture in Quantitative Methods in Clinical Research (DPET831) *Chapel Hill, NC*

· Introduction to machine learning: statistical learning, deep learning, and application to pharmaceutical research

University of North Carolina, Chapel Hill Fall 2018
Teaching assistant (Instructor: Dr. Matthew Biggs) *Chapel Hill, NC*

· Introduction to Data Science (BIOS 611)
· Participated in designing/planning the first course offering.
· Led lab sessions and in-class discussion.
· Made and graded homework.

University of North Carolina, Chapel Hill Fall 2019
Grader *Chapel Hill, NC*

· BIOS667 Longitudinal data analysis

PUBLICATIONS

1. Methods and Theory

Kang, C., **Cho, H.**, Song, R., Banerjee, M., Laber, E.B., Kosorok, M.R. Inference for Change-Plane Regression. (2022) *arXiv*.

Cho, H.*, Holloway, S.T., Couper, D. J., Kosorok, M.R. Multi-stage optimal dynamic treatment regimes for survival outcomes with dependent censoring. (2022) *arXiv*.
R pacakge: dtrSurv on CRAN.

Honvoh, G.D., **Cho, H.**, Kosorok, M.R. Model Selection for Survival Individualized Treatment Rules Using the Jackknife Estimator. *(submitted)*.

Cho, H., Jewell, N.P., Kosorok, M.R. Interval Censored Recursive Forests. (2021) *Journal of Computational and Graphical Statistics*.
R package: icrf on CRAN.

Cho, H.*, Liu, C., Preisser, J.S., Wu, D.* A bivariate zero-inflated negative binomial model and its applications to biomedical settings. R package: bzinb on CRAN. (2022) *bioRxiv*.

* corresponding author; † co-first author

2. Review and discussion

Cho, H., Liu, C., Tang, B., Lin, B.M., Roach, J., Ribeiro, A., Love, M., Divaris, K., Wu, D. Distribution-based comprehensive evaluation of differential expression analysis for metatranscriptome data. (2021) *bioRxiv*.

Kim, S., **Cho, H.**, Bang, D., De Marchi, D., El Zaatari, H., Shah, K.S., Valancius, M., Zikry, T., Kosorok, M.R. (2021). Discussion of 'Estimating time-varying causal excursion effect in mobile health with binary outcomes' *Biometrika*.

Wu, D., Karhade, D., Pillai, M., Jiang, M., Huang, L., Li, G., **Cho, H.**, Roach, J., Li, Y., Divaris, K. (2021). Machine learning and deep learning in genetics and genomics. *Machine Learning in Dentistry*

Cho, H., Zitkovsky, J., Li, X., Lu, M., Shah, K., Sperger, J., Tsilimigras M. C. B., Kosorok, M.R. (2020). Comment: Diagnostics and Kernel-based Extensions for Linear Mixed Effects Models with Endogenous Covariates *Statistical Science*.

Lawson, M. T., **Cho, H.**, Choudhury, A., Cui, Y., Jiang, X., Pokaprakarn, T., Kosorok, M.R. (2019). Discussion of Laber et al. "Optimal treatment allocations in space and time for on-line control of an emerging infectious disease." *Journal of the Royal Statistical Society Series C*.

Cho, H., Matthews, G. J., & Harel, O. (2019). Confidence intervals for the area under the receiver operating characteristic curve in the presence of ignorable missing data. *International Statistical Review*.

3. Application and Collaborative

Cho, H.†, Zhi, R.†, Divaris, K., Roach, J., Lin, B.M., Liu, C., Azcarate-Peril, M.A., Simancas-Pallares, M.A., Shrestha, P., Orlenko, A., Ginnis, J., North, K.E., Zandona, A.G.F., Ribeiro, A.A., Wu, D., and Koo, H. (2022) Pathobiont-mediated spatial structuring enhances biofilm virulence in childhood oral disease. *(In Review at Nature Communications)*.

Simancas-Pallares, M., Gormley, A., Shrestha, P., Gu, Y., **Cho, H.**, Spangler, H.D., Burk, Z., Smith, M. Dashper, S., Burgner, D., Zandoná, A.G.F., Ginnis, J., Vann W.F., Esberg. A., Roach, J., Wu, D., Silva, M.J. Holgerson, P.L., Haworth, S., Johansson, I., North, K.E., and Divaris, K. (submitted) Evidence for clinical subtypes of early childhood caries.

Xie, J., **Cho, H.**, Lin, B.M., Pillai, M., Roach, J., Heimisdottir, L.H., Zou, F., Divaris, K., Wu, D. (2021). Improved Metabolite Prediction Using Microbiome Data-Based Elastic Net Models. *Frontiers in Cellular and Infection Microbiology*.

Rosa, T., Neves, A.A., Azcarate-Preil, M.A., Divaris, K., Wu, D., **Cho, H.**, Moss, K., Paster, B.J., Chen, T., Freitas-Fernandes, L.B., Fidalgo, T.K.S., Lopes, R.T., Valente, A.P., Arnold, R.R., Riberio, A.A. (2021). The Bacterial Microbiome and Metabolome in Caries Progression and Arrest. *Journal of Oral Microbiology*

Heimisdottir, L., Lin, B., **Cho, H.**, Orlenko, A., Ribeiro, A., Simon-Soro, A., Roach, J., Shungin, D., Ginnis, J., Simancas-Pallares, M., Spangler, H., Zandona, A., Wright, J.T., Ramamoorthy, S., Moore, J., Ku, H., Wu, D., Divaris, K. (2021). Metabolomics insights in early childhood caries. *Journal of Dental Research*

Karhade, D.S., Roach, J., Shrestha, P., Simancas-Pallares, M., Ginnis, J., Ribeiro, A.A., **Cho, H.**, Wu, D., Divaris, K. (2021). An Automated Machine Learning Classifier for Early Childhood Caries Prediction. *Pediatric Dentistry.*

Divaris. K., Slade, G.D., Zandona A.G.F., Preisser, J.S., Ginnis, J., Simancas-Pallares, M.A., Agler, C.S., Shrestha, P., Karhade, D.S., Ribeiro, A.A., **Cho, H.**, Gu, B.Y., Meyer, B.D., Joshi, A.R., Azcarate-Peril, M.A., Basta, P.V., Wu, D., North, K.E. (2020). Cohort Profile: ZOE 2.0—A Community-Based, Genetic Epidemiologic Study of Early Childhood Oral Health. *International Journal of Environmental Research and Public Health.*

Divaris, K., Shungin, D., Rodríguez-Cortés, A., Basta, P. V., Roach, J., **Cho, H.**, Wu, D., Ferreira-Zandoná, A. G., Ginnis, J., Ramamoorthy, S., Kwintkiewicz, J., Butz, J., Azcarate-Peril, M. A. (2019). The supragingival biofilm in early childhood caries: clinical and laboratory protocols and bioinformatics pipelines supporting oral metagenomics, metatranscriptomics and metabolomics studies of the oral microbiome. *Methods in Molecular Biology.*

PRESENTATIONS

1. Talks

Non-parametric dynamic treatment regimes for survival outcomes, A Biometrics Section Paper Award, Joint Statistical Meetings, August, 2021.

Interval censored recursive forests, Joint Statistical Meetings, August, 2020.

Bivariate zero-inflated negative binomial model For single cell RNAseq data, 2019 Spring Meeting, Eastern North American Region International Biometric Society, March, 2019.

Bivariate zero-inflated negative binomial model For single cell RNAseq data and gene set testing, GenSTAT meeting, University of North Carolina, December, 2018.

2. Poster presentations

Biofilm metagenomics and metatranscriptomics in early childhood caries, Adams School of Dentistry Research Day 2020, February, 2020.

Biofilm metagenomics and metatranscriptomics in early childhood caries, JCVI/AADR Fall Focused Symposium, November, 2019.

COMPUTER LITERACY

High proficiency in R, C, C++, SAS (IML), Linux, and VBA
(R packages based on lower level languages have been developed: `bzinb`, `icrf`, `dtrSurv`)
Intermediate proficiency of Python, MATLAB, Nextflow, and Docker
Simulation design, complex data management, and efficient programming

JOURNAL REFEREE SERVICE

Journal of American Statistical Association	2018-2022
Annals of Statistics	2019-2020
Journal of the Royal Statistical Society, Series B	2019
Frontiers in Physiology	2022

PROFESSIONAL EXPERIENCE IN FINANCE

Hyundai Commercial Inc. January 2013 - June 2014
Credit Risk Manager *Seoul, Korea*

· Risk measurement tool development for financial products

PricewaterhouseCoopers October 2005 - August 2007, January 2011 - January 2013
Certified Public Accountant *Seoul, Korea*

· Financial advisory services (Due diligence, Feasibility study, Valuation)

Republic of Korea Air Force September 2007 - December 2010
Finance officer (Captain) *Suwon, Korea*

· Expenditure control and compilation of financial statements

PROFESSIONAL AFFILIATION

Member, American Statistical Association 2015 - 2017, 2020 - Present

Chartered Financial Analyst, CFA Institute October 2013 - October 2014

Certified Public Accountant, Korea Institute of CPA September 2003 - Present