

Congyu Hang

hangcong@uw.edu • 206-734-1514 • Seattle, WA • LinkedIn: cong-yu-hang-b51a15290

EDUCATION

University of Washington, Seattle

Sep 2024 – Jun 2026

Master of Science (Thesis) in Biostatistics • GPA: 3.86/4.00

Selected Coursework: Introduction to Biostatistics; Statistical Inference; Machine Learning for Health Data; Linear Models

University of Toronto, ON, Canada

Sep 2020 – Jun 2024

Bachelor of Science, Specialist of Science – Quantitative Finance • GPA: 3.96/4.00

Honors: Dean's List (2021–2023); Vincent Bladen In-course Scholarship (2023); Sotherton Wadhams In-course Scholar (2023);

Joseph Alfred Whealy In-course Scholarship (2022); A. D. Allen Memorial (In-course) Scholarship (2022)

PROGRAMMING SKILLS

Languages: R, Java, Python, SQL, MATLAB, Mathematica • **Utilities:** RStudio, HPC, Anaconda, Git, Jupyter, L^AT_EX, Tableau

PUBLICATIONS

BLNM for Allele-Specific Expression (ASE)

Hang, C.^{*}, Vogel, R., Ganapathy, K., Mohammadi, P. •

Manuscript in preparation; target: *Bioinformatics*

- Introduces an overdispersed Bayesian Binomial–Logit–Normal Mixture (BLNM) model extending ANEVA/ANEVA-DOT to robustly infer allelic-specific expression variance, model outliers via a Uniform component, and improve cis-regulatory effect estimation across GTEx tissues.

Association of Paid Family Leave Policy with Child Physical Abuse and Neglect

Georges, E., Peng, J., Ye, T., Hang, C., Brown, E. C. B., Rowhani-Rahbar, A. • Manuscript in preparation; target: *Pediatrics*

- Evaluates the association between state Paid Family Leave policies and rates of physical abuse and neglect among infants ≤ 2 years using a weighted event-study design based on the Callaway–Sant’Anna difference-in-differences framework.

Paid Family Leave and Abusive Head Trauma (AHT) / Neglect (DiD Evaluation)

Hang, C.^{*}, Ye, T. •

Manuscript in preparation

- Uses population-weighted event-study and Callaway–Sant’Anna difference-in-differences models to quantify the causal impact of Paid Family Leave on rates of abusive head trauma and neglect among one-year-olds, validated across 7 U.S. states.

RESEARCH & PROJECTS

BLNM for Allele-Specific Expression (ASE) — ANEVA/ANEVA-DOT Extensions

Research Assistant • Supervised by Prof. Pejman Mohammadi (Pejlab), Seattle Children’s

Dec 2024 – Present

- Developed the `blnm` Python package implementing a Binomial–Logit–Normal (BLN) framework for parameter estimation and likelihood calculation, supporting uni-/multi- cis-regulatory variants in imbalanced expression modeling.
- Derived the Uniform-augmented BLNM mixture to handle outliers; provided closed-form M-steps, implemented the extension, and streamlined model selection for users.
- Proposed an initialization method that visualizes alternative vs. reference read counts as scatter plots and applies clustering to choose starting parameter regions—reducing sensitivity to local extrema and improving EM stability.
- Evaluated robustness across 20+ scenarios using 10^4 Monte Carlo replicates; assessed coverage and efficiency.
- Applied to GTEx data for model selection and variance-component estimation; preparing a manuscript targeting *Bioinformatics*.

DNAnexus Applets for UK Biobank (UKBB)

Research Assistant • Supervised by Prof. Ting Ye & CHRU

Jun 2025 – Sep 2025

- Built a Proteomics applet enabling one-command analyses (linear/logistic regression, LMM, GEE, Cox), fully containerized and parameterized; auto-exports result tables and QQ plots.
- Assembled a cardiac MRI segmentation workflow (data I/O, model weights, GPU/CPU routing) inspired by `ukbb_cardiac`, aligning Docker images and script I/O for reproducibility.

Frontier LLM Evaluation for Clinical Trial Development Tasks (Thesis)

Research Assistant • Supervised by Prof. Ting Ye

Aug 2025 – Present

- Designed protocols for information retrieval and summarization in clinical trial development; implemented metrics (**retrieval AUC**, calibration) and bootstrap CIs.
- Built a reproducible pipeline to benchmark frontier LLMs on standardized trial documents.

Paid Family Leave and Abusive Head Trauma (AHT) in 1-Year-Olds

Research Assistant • Supervised by Prof. Ting Ye & Emily Georges, Seattle Children’s

Jun 2025 – Oct 2025

- Curated a 7-state panel (30G+ HCUP data); applied population-weighted event studies, Callaway–Sant’Anna (CS), Borusyak–Jaravel–Spiess (BJS) and TWFE with CR2 clustered SEs; verified pre-trends.
- Compared estimators’ efficiency and selected models based on robustness under data sparsity.

SELECTED IN-CLASS PROJECTS

Rats' TNT Preference and Rater Reliability

Course Project • STAD94 — Statistics Project

Fall 2023

- Applied **PCA + linear modeling** to denoise rater scores, revealing a **78% learning gain** in TNT preference across **32** training videos.

Canadian Stock Price Dynamics

Course Project • STAD57 — Time Series Project

Fall 2023

- Modeled market returns with **ARMA** and simulation, uncovering short-term autocorrelation and a **7.8×** gain over random-walk baselines.

Seasonal Forecasting of Canadian CPI

Course Project • STAD57 — Time Series Project

Fall 2023

- Built **SARIMA(0,1,1)(0,1,1)[12]** with iterative updating, reducing 1-step forecast error by **85%** versus 9-step predictions.

Gas Furnace Process Modeling

Course Project • STAD57 — Time Series Project

Fall 2023

- Developed **ARIMA + transfer-function** model capturing a 5-period feed lag, halving residual variance and confirming methane→CO₂ causality.

Survival Analysis of Cancer Patients

Reading Course • STAD92 — Survival Analysis

Summer 2023

- Constructed **Kaplan–Meier** and **Cox PH** models showing ulcer cohort **HR** \approx **1.7**.

Predicting Fire Severity in Toronto

Data Science Project • STAA57 — Data Science

Spring 2023

- Engineered **logistic regression with CV (AUC = 0.82)** on **6k** incidents, identifying sprinkler/alarm failures as key severity predictors.

TEACHING EXPERIENCE

Teaching Assistant (BIOST 310: Introduction to Statistical Methods)

Teaching Assistant • University of Washington, Seattle, WA

Sep 2025 – Present

- Supports a health-sciences-focused introduction to applied statistics using real-world case studies.
- Leads labs and review sessions on study design, estimation, hypothesis testing, categorical data analysis, and regression; emphasizes interpretation over computation.

Teaching Assistant

Teaching Assistant • University of Toronto Scarborough (UTSC), Toronto, ON

Sep 2022 – Apr 2024

- **Courses:** STAB57 – *Introduction to Statistics (W2023–W2024)*; STAA57 – *Introduction to Data Science (W2024)*; MATA33 – *Calculus for Management II (W2023)*; MATA32 – *Calculus for Management I (F2022)*; MATA30 – *Calculus I for Physical Sciences (F2022)*.
- Led tutorials, office hours, and exam review sessions; graded assignments and exams.
- **Statistics:** Taught likelihood, MLE, hypothesis testing, confidence intervals, and sampling theory; created proof-sketch notes linking theory and interpretation.
- **Data Science:** Instructed R workflows (**tidyverse**, **ggplot2**), data cleaning, imputation, visualization, and regression modeling; designed interactive **learnr** tutorials for auto-graded exercises.
- **Calculus:** Reviewed differentiation, optimization, integration, and Taylor series; introduced partial derivatives and constrained optimization for management applications.

WORK EXPERIENCE

Beijing Business Group of CCB (China Construction Bank) Financial Technology Co., Ltd.

Database Development Intern, Credit Card Department

Jun – Jul 2023

- Structured Excel datasets with clearly defined variable names tailored to business requirements; utilized **SQL** to create variable-level tables, improving data traceability and management.
- Analyzed relationships across **MPP databases** using table descriptions and engineered new variables to support **logistic regression models** for customer-targeted product promotion.

Tianjin Maijie Coding Technology Development Co., Ltd.

Market Analyst Intern & Management Assistant

Jan – May 2021, Apr – May 2023

- Segmented customers by purchase intention using **COUNTIF functions** and **pivot tables** to improve targeted marketing strategies.
 - Conducted statistical analyses of sales and demand data (mean, variance, standard deviation) to determine optimal inventory levels and forecast demand trends.
 - Applied **maximum demand**, **ABC analysis**, and **expected demand inventory** methods to support efficient two-bin inventory management.
-