Yuki Ohnishi

PhD Student · Statistics

242 S.Salisbury St. Apt. 7, West Lafayette, IN, USA, 47906

+1 765-400-6824 | yohnishi@purdue.edu | https://yukiohnishi.github.io

Research Interest

Causal Inference, Bayesian Analysis, Differential Privacy, Machine Learning, Clinical Trials, Digital Marketing

Education

Purdue University *IN, USA*
PhD Statistics *2018 - present*
- Advisor: Dr. Jordan Awan and Dr. Arman Sabbaghi

The University of Tokyo *Tokyo, Japan*
MS Industrial Engineering, Operations Research *2011 - 2014*

The University of Tokyo *Tokyo, Japan*
BS Industrial Engineering, Operations Research *2007 - 2011*

Professional Experience

2023 **Data Science Summer Intern**, Boehringer Ingelheim, CT, USA
2021-2023 **Graduate Research Assistant**, Purdue University, IN, USA
2018-2022 **Graduate Teaching Assistant**, Purdue University, IN, USA
2015-2018 **Data Scientist and Machine Learning Engineer**, BizReach, inc., Tokyo, Japan

Publications

Published/Accepted

Ohnishi, Y and A. Sabbaghi. 2022. A Bayesian Analysis of Two-Stage Randomized Experiments in the Presence of Interference, Treatment Nonadherence, and Missing Outcomes. *Bayesian Analysis*, 1–30., doi.org/10.1214/22-BA1347.

Ohnishi, Y and Jean Honorio. 2021. Novel change of measure inequalities with applications to PAC-Bayesian bounds and Monte Carlo estimation. *International Conference on Artificial Intelligence and Statistics*, 1711-1719.

In Review

Ohnishi, Y, B. Karmakar and A. Sabbaghi. Degree of Interference: A General Framework for Causal Inference under Interference

Ohnishi, Y and J. Awan. Locally Private Causal Inference

Ohnishi, Y, B. Karmakar and W. Kar. Inferring Causal Effect of a Digital Communication Strategy under a Latent Sequential Ignorability Assumption and Treatment Noncompliance

In Preparation

Targeted Bayesian Causal Inference via Loss Function

Locally Private Causal Inference for Observational Data

Transportable and Calibrated Causal Inference in Sequential Digital Campaigns

Presentations

Ohnishi, Y. 2023. A Bayesian Analysis of Two-Stage Randomized Experiments in the Presence of Interference, Treatment Nonadherence, and Missing Outcomes. International Conference on Design of Experiments (ICODOE 2023). The University of Memphis, Memphis, TN.

Ohnishi, Y. 2023. Locally Private Causal Inference. Midwest Machine Learning Symposium (MMLS 2023), Chicago, IL.

Sabbaghi, A. and **Y. Ohnishi**. 2022. A Bayesian Analysis of Two-Stage Randomized Experiments in the Presence of Interference, Treatment Nonadherence, and Missing Outcomes. INFORMS Workshop on Quality, Statistics, and Reliability. Indianapolis, IN.

Ohnishi, Y. 2021. Novel Change of Measure Inequalities with Applications to PAC-Bayesian Bounds and Monte Carlo Estimation. The 24th International Conference on Artificial Intelligence and Statistics (AISTAT 2021), Virtual.

Ohnishi, Y. 2019. Applying Bayesian Hierarchical Probit Model to Interview Grade Evaluation. KDD'19 INTERNATIONAL WORKSHOP ON TALENT AND MANAGEMENT COMPUTING, Anchorage, AK.

Ohnishi, Y. 2018. Bayesian Hierarchical Bernoulli-Weibull Mixture Model for Extremely Rare Events. INFORMS Business Analytics Conference, Baltimore, MD.

Teaching Experience

Fall 2022	**STAT 598 Differential Privacy**, Grader
Spring 2021	**STAT 538 Probability Theory I**, Teaching Assistant
Fall 2021	**STAT 539 Probability Theory II**, Grader
Fall 2021	**STAT 350 Introduction To Statistics**, Teaching Assistant
2018-2020	**STAT 190 Topics In Statistics For Undergraduates**, Grader
2018-2020	**STAT 113 Statistics And Society**, Teaching Assistant

Research Experience

Completed Projects

Advisor: Dr. Arman Sabbaghi and Dr. Bikram Karmakar *Jan. 2021 - present*

- **Causal Inference under Violated Assumptions**
  Causal inference is performed under several core assumptions. These assumptions include:
  - *No Interference*: An individual's outcomes remain unaffected by another's treatment status.
  - *Treatment Compliance*: All units consistently adhere to their designated treatments.
  - *Positivity*: Both treatment and control groups have representative units.

  However, in modern, complex experiments and observation studies, these assumptions often go unmet. My research projects focus on developing methodologies for performing statistically valid causal inference in the face of these violations, addressing both design and analysis angles.

Advisor: Dr. Jordan Awan *Aug. 2022 - Present*

- **Causal Inference for Differentially Privatized Data**
  Local differential privacy (LDP) is a differential privacy (DP) paradigm in which individuals first apply a DP mechanism to their data (often by adding noise) before transmitting the result to a curator. In this project, we develop methodologies to infer causal effects from locally privatized data under the Rubin Causal Model framework. First, we present frequentist estimators under various privacy scenarios with their variance estimators and plug-in confidence intervals. We show that using a plug-in estimator results in inferior MSE compared to minimax lower bounds. We also develop a Bayesian nonparametric methodology along with a blocked Gibbs sampling algorithm, which can be applied to any of our proposed privacy mechanisms. We then present simulation studies to evaluate the performance of our proposed frequentist and Bayesian methodologies for various privacy budgets, resulting in useful suggestions for performing causal inference for privatized data.

On-going Projects

May. 2023 - Present

- **Targeted Bayesian Causal Inference via Loss Functions**
  Bayesian inference is one of the modes of inference within the potential outcome framework. The standard Bayesian causal inference starts with the joint distribution of all quantities, including missing potential outcomes, to derive the posterior distributions of parameters of interest and missing quantities. In this procedure, the propensity score, the treatment assignment mechanism, has a paradoxical role when used as the weighting score, as it is often assumed to be *ignorable*, i.e., it drops out of the likelihood. We provide a foundation to naturally incorporate the propensity score as the weighting score into the Bayesian analysis based on the general Bayesian updating scheme (Bissiri et al., 2016). Our framework also provides a unified view of the coherent inference under model misspecifications and the population shift, where the observed population differs from our inferential target population in their distribution.

Academic services

Reviewer

Journal of Applied Statistics

Skills & Languages

Skills Julia, Python, R, Java, Stan, SQL, Git, Linux, TensorFlow/Keras, GCP/AWS

Languages Japanese, English, Chinese