Newsworthy Research Highlights

from JSM 2023

The 2023 Joint Statistical Meetings will bring together statisticians and data scientists from around the world from Saturday, August 5, through Thursday, August 10. This year, JSM will be held in Toronto. This tip sheet highlights interesting presentations from the conference. Complementary press registration is open, courtesy of the ASA. Email *edoffice@amstat.org* for more information.

FEATURED RESEARCH (SYNOPSES BELOW)

MONDAY HIGHLIGHTS

1. It's Who Is Missing That Matters: Can a Nonignorable Missingness Mechanism Explain Bias in Estimates of COVID-19 Vaccine Uptake?

2. How Much Does Home Field Advantage Matter in Soccer Games? A Causal Inference Approach for English Premier League Analysis

3. What Really Matters for Fairness in Machine Learning: Delayed Impact and Other Desiderata

TUESDAY HIGHLIGHTS

4. Fundamentals of Interpretable Machine Learning

5. Harnessing Socio-Cultural Similarities Between Diverse Populations to Identify Determinants of Cancer Screening Use in Under-Studied Populations

6. Efficient Detection of Rare Events in Autonomous Vehicles Driving Data

WEDNESDAY HIGHLIGHTS

7. Building a New Measure of Food Insecurity in America

8. An Algorithm for Forensic Toolmark Comparisons

9. Rethinking Race-Ethnicity: Introducing Novel Survey-Based Measures of Lifetime Experience of Discrimination and Stress

THURSDAY HIGHLIGHTS

10. Cryptocurrency in a Global Context: Intentions and Motivations of Consumers in 20 Markets around the World

11. Genomic Risk Prediction: Algorithms, Fairness, and Applications

12. Optimal When-to-Treat Policies Under Dynamic Resource Constraints

MONDAY SYNOPSES

1. It's Who Is Missing That Matters: Can a Nonignorable Missingness Mechanism Explain Bias in Estimates of COVID-19 Vaccine Uptake?

Recently, attention was drawn to the 'failure' of two very large internet-based probability surveys (Delphi-Facebook CTIS, Census Household Pulse) to correctly estimate COVID-19 vaccine uptake in the US in early 2021. These surveys overestimated vaccine uptake substantially (14 and 17 points in May 2021) compared to retroactively available CDC benchmark data. Though large, these surveys had very low response rates, thus nonignorable nonresponse could have substantially affected estimates. Specifically, it is plausible that 'anti-vaccine' individuals were less likely to complete a survey about COVID-19 or an official government-sponsored survey.

This presentation considers models to retrospectively estimate the proportion of adults who received at least one vaccine dose, using data from these two surveys, under a nonignorable nonresponse assumption. These estimates are compared to the true benchmark uptake numbers, enabling assessment of whether nonignorable nonresponse is a plausible explanation for the biased estimates.

2) How Much Does Home Field Advantage Matter in Soccer Games? A Causal Inference Approach for English Premier League Analysis

In many sports, it is commonly believed that the home team has an advantage over the visiting team, known as the home field advantage, yet its causal effect on team performance is largely unknown. This presentation considers a novel causal inference approach to studying the causal effect of home field advantage

in the English Premier League. The method was implemented on the 2020–2021 English Premier League data and assesses the causal effect of home advantage on 11 summary statistics that measure the offensive and defensive performance and referee bias. Findings indicate the home field advantage resides more heavily in offensive statistics than it does in defensive or referee statistics and that teams with lower rankings retain a higher home field advantage.

3) What Really Matters for Fairness in Machine Learning: Delayed Impact and Other Desiderata

From education to lending, consequential decisions in society increasingly rely on data-driven algorithms. Yet the long-term impact of algorithmic decision-making is largely ill understood and there exist serious challenges to ensuring equitable benefits, in theory and practice. While the subject of algorithmic fairness has received much attention, algorithmic fairness criteria have significant limitations as tools for promoting equitable benefits.

This presentation reviews various fairness desiderata in machine learning and when they may be at odds with learning. This research demonstrates that several statistical criteria for fair machine learning, if applied as a constraint to decision-making, can result in harm to the welfare of a disadvantaged population. The presentation will conclude by considering future directions for fairness in machine learning that evince a holistic and interdisciplinary approach.

TUESDAY SYNOPSES

4) Fundamentals of Interpretable Machine Learning

Interpretability in machine learning is crucial for high-stakes decisions and troubleshooting. Interpretable machine learning started as far back as the 1970s but has gained momentum as a subfield only very recently.

This presentation will touch on recent research in the area, provide fundamental principles for interpretable machine learning, and dispel common misunderstandings that dilute the importance of this crucial topic. Ten technical challenge areas in interpretable machine learning will be discussed.

5) Harnessing Socio-Cultural Similarities Between Diverse Populations to Identify Determinants of Cancer Screening Use in Under-Studied Populations

Screening is effective when it detects cancer early. It is easier to treat, and the chances of survival are better. However, there is considerable racial disparity in the use of cancer screening. Standard approaches to identify factors associated with these disparities begin by stratifying the data on race and then fitting regression models. However, strata based on discrete labeling of race such as non-Hispanic White, non-Hispanic Black, and so on are likely to be oversimplified when attempting to interpret screening use in diverse populations. Further, the sample sizes of under-studied populations such as non-Hispanic Black, Hispanic, and Asian subgroups are often small. To address these challenges, this research considers a novel approach to stratify individuals according to their socio-cultural similarities, regardless of their race/ethnicity. The approach will be illustrated using multiple years of data from the National Health Interview Survey to examine disparities in mammogram screening use in diverse populations.

6) Efficient Detection of Rare Events in Autonomous Vehicles Driving Data

Autonomous vehicles have made great strides and continue to be an active area of investment and research. Widespread scaling of autonomous vehicles requires a robust evaluation process for the safety and quality of the driving, prior to releasing autonomous driving systems on public roads. Evaluating the performance of autonomous vehicles typically involves identifying an event of interest and estimating the frequency at which that event occurs. This presentation will outline the role of human labeling in this estimation process, the operational limitations and resulting properties of the resulting labeled data sets, and the statistical challenges that result.

WEDNESDAY SYNOPSES

7) Building a New Measure of Food Insecurity in America

The food insecurity measure in the US is currently estimated using responses from the Current Population Survey. This research presents a new measure of food insecurity at the census tract level for the National Capital Area (14 counties and cities in the Washington, DC, Area). The model is based on a financial evaluation of living costs—including food, housing, child care, health care, transportation, taxes, and other expenses—using different publicly available information. The analysis focuses on three key determinants: household size and composition; household income; and geographic cost-of-living. A procedure was used to estimate joint distributions of the population at the census tract level and categorize households as food insecure, at risk, and food secure. Including the geographic cost-of-living made prediction of the changing needs for

interventions after variations in cost categories possible. The estimates were benchmarked with SNAP administrative data and income at the census tract level. Unlike non-survey food insecurity measures, variability estimates are provided to give a realistic range of scenarios around interventions.

8) An Algorithm for Forensic Toolmark Comparisons

Forensic practitioners determine whether two marks were generated by the same tool by observing the 2D images of the marks using a comparison microscope and deciding whether the surface contours of two toolmarks are in sufficient agreement' based on the examiner's subjective opinion that another tool could not have made the marks. This study proposes a novel algorithm (an objective method) for forensic toolmark comparisons that can compare marks made at different angles and directions and provide a measure of uncertainty.

9) Rethinking Race-Ethnicity: Introducing Novel Survey-Based Measures of Lifetime Experience of Discrimination and Stress

Though race-ethnicity is not a biological variable, race-ethnicity is included in nearly every medical study and often very statistically and meaningfully significant. New measures are critically needed that will allow biomedical researchers to disentangle race-ethnicity from the true individual, interpersonal, and structural causes of health disparities. This presentation introduces novel measures using simple survey items to capture self-reported experiences of discrimination and stress spanning the life course, agnostic to source. Early results from an employee survey will be presented. In the future, by implementing tools such as these within electronic medical records, clinicians may better identify patients at risk of accelerated aging.

THURSDAY SYNOPSES

10) Cryptocurrency in a Global Context: Intentions and Motivations of Consumers in 20 Markets Around the World

Who intends to use Bitcoin or other cryptocurrencies, and for what purposes are cryptocurrencies being used? Leveraging global research in 20 markets around the world, this presentation examines the prevalence of cryptocurrency usage intent and the motivations for doing so. Some intend to use crypto as a speculative long-term investment. Others are likely to day trade. Still others intend to use crypto as a way of securing their privacy, to shelter their income from taxation, to send money across borders, or simply to buy goods or services.

The data reveals that the intention to use crypto is greater in emerging markets and among certain demographics of the population. This large survey of more than n = 14,000 internet users also allows for stratification by market, gender, age, and income.

11) Genomic Risk Prediction: Algorithms, Fairness, and Applications

This overview lecture is given by a leading expert in disease risk prediction, especially pertaining to cancer disease. He will introduce the concepts of genetic risk prediction in the context of large-scale, genome-wide association studies, with significant focus on theoretical properties, implementation, and illustration to breast cancer risk ascertainment. This will be a self-contained lecture focused on cutting-edge methods of polygenic-risk scores and discussion of its clinical utility in disease discrimination.

12) Optimal When-to-Treat Policies Under Dynamic Resource Constraints

Precision medicine uses data to leverage patient heterogeneity to learn optimal tailored treatments. While many strategies have been proposed for learning optimal tailored treatment policies, only recently have methods been developed for dealing with the question of treatment timing and, separately, attempts have been made to incorporate resource constraints. Using the when-to-treat (WTT) policy class introduced by Nie, Brunskill, and Wager, we propose a policy learning algorithm to find an optimal WTT policy under a dynamic resource constraint in an indefinite time horizon setting. Crucially, our work accounts for the impact of allocating resources to patients today on the health and disease progression for patients treated tomorrow. We present preliminary performance guarantees using empirical processes and assess performance in both simulated and real-world resource-constrained data.