Colorado – Wyoming Chapter of the American Statistical Association

Fall 2025 Meeting

Date: Friday, November 14, 2025
Time: 11:30–4:30 p.m.
Location: Plaza Building, Room M204, 955 Lawrence Way #122, Denver, CO 80204
The plaza building is on a plaza. Use the East Door on the plaza.
Parking Information: Auraria Campus Visitor Parking
RTD Light Rail: Close stations are Ball Arena, Auraria West Campus, and Colfax at Auraria
Questions? Contact Jason Bernstein at `jstatbernstein@gmail.com`.

Program Schedule

11:30 – 12:00
Room becomes available at 11:30, folks welcome to hang out, chat, etc.

12:00 – 12:50
Lunch.
Box lunches provided by Biscuits and Berries.
Delivery window is 11:30 am - 12:15 pm.

12:50 – 1:00
Opening remarks:

- Jason Bernstein, President of CO-WY Chapter, 2025 – 2026
- Kit Kennedy, President of Stat Science Club of MSU Denver
- Prof. Shahar Boneh, MSU Denver

1:00 – 1:20
Illuminating Marching Band Competition Scores
Matt Pocernich

Under the theme of "Things you're not interested in until your kids are," I've developed a new interest in how marching band competitions are scored. Typically, results appear in cryptic spreadsheets that obscure more than they reveal. The scoring system itself has an interesting hierarchy built around three broad categories: music performance, visual performance, and general effect. Each of these branches into multiple sub-categories, judging both individual performers and the full ensemble, and ultimately rolling up into a single, tidy score on a 0–100 scale.

Conveniently, many years of results are publicly available online. This presentation highlights an effort to build an interactive dashboard that clarifies these complex scoring structures, reveals

trends and insights hidden within the data, and explores which factors do—and don't—impact competition outcomes.

1:20 – 1:40
Discussion: The Role of the ASA and Advocacy in Statistics
Kristin Broms

This session will start a conversation about the purpose of the ASA and the role of advocacy within the organization. Topics include: the balance between professional support and non-partisanship, the meaning of advocacy, and what topics the ASA should champion. The goal is to encourage open dialogue and debate rather than definitive answers.

1:40 – 2:00
Sparse two-block dimension reduction
Sven Serneels, Aspentech and University of Antwerp

A method is introduced to perform simultaneous sparse dimension reduction on two blocks of variables. Beyond dimension reduction, it also yields an estimator for multivariate regression with the capability to intrinsically deselect uninformative variables in both independent and dependent blocks. An algorithm is provided that leads to a straightforward implementation of the method. The benefits of simultaneous sparse dimension reduction are shown to carry through to enhanced capability to predict a set of multivariate dependent variables jointly. Both in a simulation study and in two chemometric applications, the new method outperforms its dense counterpart, as well as multivariate partial least squares.

2:00 – 2:15
Break

2:15 – 2:35
Spatiotemporal Multi-Fidelity Emulation
Tristan Contant, Ander Wilson, Yawen Guan, Colorado State University

Multi-fidelity (MF) emulation is a modeling approach that combines data from simulations of varying fidelity to construct an efficient and accurate surrogate model. Typically, these data sources are a small number of accurate but computationally expensive high-fidelity (HF) runs and a larger number of cheaper, less accurate low-fidelity (LF) runs. We develop an MF emulator using Tucker decomposition for dimensionality reduction, Gaussian process (GP) priors for flexible function approximation, and additive discrepancy modeling to systematically account for differences between LF and HF outputs. This approach enables scalable emulation not only across high-dimensional input parameter spaces but also for large spatiotemporal output fields, encompassing many locations and time points. It preserves strong predictive accuracy and uncertainty quantification while effectively capturing complex spatiotemporal trends. By leveraging the complementary strengths of LF and HF data, the emulator greatly reduces the computational cost associated with MPAS-Seaice, a complex and resource-intensive sea ice model, making it particularly suitable for prediction at un-

tried inputs and for data assimilation tasks in climate science. This is especially important for sea ice research, where understanding ice dynamics under climate change requires both computational efficiency and robust uncertainty quantification across a wide range of scenarios.

2:35 – 2:55

A Bayesian Functional Concurrent Zero-Inflated Dirichlet-Multinomial Regression Model with Application to Infant Microbiome

Brody Erlandson, Ander Wilson, Matthew D. Koslovsky, Colorado State University

The infant microbiome undergoes rapid changes in composition over time and is associated with long-term risks of conditions such as immune strength, allergy, asthma, and other health outcomes. Modeling the associations between exposures or treatments and microbial composition over time is essential for understanding the factors that drive these changes. Estimating these temporal dynamics has several challenges including: repeated measures, overdispersion, compositionality, high-dimensional parameter spaces, and zero-inflation. Many longitudinal regression models used in human microbiome research assume constant effects over time that cannot capture time-varying or functional effects of exposures, ignore the compositional structure of the data by modeling each taxon separately, and are not equipped to handle potential zero-inflation. Dirichlet-multinomial (DM) regression models inherently accommodate overdispersion and the compositional structure of the data and have been extended to account for excess zeros. However, existing DM-based regression models are unable to additionally handle repeated measures designs. To fill this gap, we propose a functional concurrent zero-inflated Dirichlet-multinomial (FunC-ZIDM) regression model which is designed to model time-varying relations between observed covariates and microbial taxa while accounting for zero-inflation, compositionality, and repeated measures. Through simulation, we demonstrate that the model can accurately estimate the underlying functional relations and scale to large compositional spaces. We apply our model to investigate time-varying associations between infant microbiome composition and observed covariates during the 11-week postnatal period. We found that α-diversity is positively associated with a higher gestational age and proportion of breast milk in the diet.

2:55 – 3:15

ggdims: toward a viz- and code-first hitchhiker's guide to dimensionality reduction

Gina Reynolds

ggplot2 lets you intuitively translate variables to visual representation. You specify how variables (e.g., sex, age, employment status) are to be communicated via visual channels (x and y axis position, color, transparency, etc). However, in ggplot2 these specifications are individual-variable-to-individual-visual-channel which does not lend itself easily to visualizations in the world of dimension reduction (e.g. PCA, tsne, umap). The usual one-var-to-one-aesthetic requirement means that it may not feel obvious how to extend ggplot2 for dimensionality reduction visualization, which deals with characterizing many variables. So while using ggplot2 under-the-hood is common in the dim-red space, it feels like there may be less consistency across dim-red APIs. For users of these APIs, getting quickly acquainted with techniques (students) or doing comparative work (practitioners) may be more challenging than it needs to be. The ggdims package explores a new

dims() and dims_expand() utility that could help with greater consistency across dim-red APIs, with standard ggplots, and within the ggplot2 extension ecosystem. We demonstrate the fluency that these new utilities afford with a replication of plots created in 'How to Use t-SNE Effectively (https://distill.pub/2016/misread-tsne/).

3:15 – 3:30
Break

3:30 – 3:50
Did We Cover That in Class? Going from Student to Alleged Expert
Catherine Durso, Office of Research and Sponsored Programs, University of Denver

In 2012, the faculty of the University of Colorado, Denver, officially let it be known that I had learned enough of the material in the Master of Science in Applied Statistics program to graduate. In 2014, the University of Denver hired me to provide statistical support to academic researchers on campus. The transition was interesting. Some specifics from the degree program have been surprisingly useful. I've lost track of the number of times I've explained what the log odds ratios from a logistic regression model do and don't mean. The deeper lessons in how to learn new material have proved essential. I frantically studied up on Generalized Estimating Equations, MCMC Methods for Multi-Response Generalized Linear Mixed Models, Random Forests, Raking, Bayesian Item Response Theory, Aster Models, Gaussian Mixture Models, Latent Class Analysis, and more. I learned to punt SEM questions to others. Also, I learned the importance of keeping clear records of the analysis process, the importance of listening to consulting clients, the importance of listening skeptically to consulting clients, and of asking what the client wants and when. Examples will be provided.

3:50 – 4:10
Is It Empathy or Something Else? A Case Study on Developing Data-informed Indicators of Item Misresponse
Wendy Christensen, Ph.D., Rachael Tan, Ph.D., and Tai Lockspeiser, M.D., MHPE.
University of Colorado School of Medicine Office of Assessment, Evaluation, and Outcomes

Empathy is a vital component of patient care, and medical schools increasingly incorporate training and assessment related to empathetic communication. At the University of Colorado School of Medicine, we use the Toronto Empathy Questionnaire (TEQ) to measure empathy in medical students. Our work with thousands of authentic responses has shown that the TEQ, which includes equal numbers of positively- and negative-coded items, is sensitive to item misresponse due to mixed coding. This issue complicates both the interpretation of individual scores and the scale's factor structure.

Although measurement research has described how and why mixed coding can lead to misresponse, there is little guidance on how to address this issue in established scales. Manual review of responses is impractical at scale, so we developed data-informed indicator variables to flag individual responses with a high likelihood of misresponse. In this presentation, I will share how we developed our initial

set of indicators and outline plans for refining the process based on early findings. While our results are specific to the TEQ, the broader aim of this presentation is to demonstrate a scalable approach to identifying and addressing misresponse that can be applied across different scales and areas of research.

4:10 – 4:30
SAS Econometrics: A Whirlwind Tour
Matt Simpson, SAS

SAS Econometrics is a powerful tool for econometricians and researchers, offering a range of features and capabilities to analyze complex economic data. In this talk, I will highlight several of its capabilities with an emphasis on cutting edge techniques and interesting projects. Topics covered will include high dimensional time series, combining causal inference and machine learning, and interesting applications.

After Meeting
No-host social
Following the meeting, we will have a no-host social at Brooklyn's, which is about 0.3 miles or a 7 minute walk from the Plaza Building. The address is 901 Auraria Pkwy, Denver, CO 80204