Applied Statistics Workshops

Spring 2024

The **Center for Statistical Computing** (CSC) invites all graduate students, staff, and faculty to participate in our applied statistics workshops. These sessions are conducted on **Zoom** and/or in **Green Labs** on the upper level (UL) of Healey Library. These workshops encompass recently developed statistical methods, utilizing tools such as SPSS, SAS, Stata, R, AMOS, Mplus, and WinBUGS. Descriptions for each workshop, along with the schedule and Zoom links, are provided below.

Workshop Descriptions:

COVID-19 Data Analysis Using R: This workshop will involve downloading COVID-19 data for states and Massachusetts from the Center for Systems Science and Engineering of Johns Hopkins University and the Department of Public Health (DPH) Massachusetts. We will employ time series and spatial regression models to analyze the COVID-19 data, utilizing R packages such as *forecast*, *tseries*, *spdep*, *maptools*, and *ggplot2*. Additionally, this workshop will demonstrate how to use R to generate reports for COVID data.

Sample Size Estimation and Power Calculations (SAS): This workshop covers sample size determinations and power estimation for various statistical comparisons and tests using the PROC POWER procedure in SAS.

Introduction to HLM (Mixed Models) (SPSS): This workshop provides an overview of the fundamental principles of multilevel/hierarchical linear models. Topics include the necessity for appropriate methods to model dependencies (e.g., clustering of students within schools), formulating and interpreting two-level multilevel models and their relevant parameters, and using SPSS to estimate model parameters.

Missing Data Analysis (SAS & Stata): This workshop covers the mechanisms of missing data, analysis of non-random selection bias, and methods of single and multiple imputation (MI) using SAS and Stata. Missing data is a common issue in various datasets. Most statistical software packages automatically eliminate entire cases with missing data from analysis, potentially leading to low sample sizes and biased results.

Structural Equation Modeling I (AMOS & R): This workshop introduces techniques for structural equation modeling (SEM). SEM is employed to test complex relationships between observed (measured) and unobserved (latent) variables. Topics covered include fundamentals underlying SEM, SEM notation, path diagrams, data preparation, mediation analysis, path analysis, parameter estimation, and assessment of model fit. AMOS and R are used to demonstrate examples.

Structural Equation Modeling II (AMOS & R): The second SEM workshop delves into advanced topics including measurement error, latent variables analysis, exploratory factor analysis (EFA), confirmatory factor analysis (CFA), development of structural equation models with estimation, and model testing. Additionally, this workshop introduces latent growth models for longitudinal data. R

program and AMOS are utilized to demonstrate model structures, parameter estimation, and model modification.

Introduction to Statistical Learning (R): This is an introductory workshop in statistical learning focusing on the important elements of modern data analysis such as regression and classification methods. Topics covered include linear and logistic regression, linear discriminant analysis, cross-validation, principal components, and clustering. Data analysis examples in this workshop are demonstrated using R.

Spatial Regression (R): This workshop introduces spatial modeling, exploring tools such as R's *maptools* and *spdep* packages. The workshop covers essential topics including spatial data visualization in R, understanding spatial autocorrelation, statistical methods for spatial dependence, creating spatial weights, and building spatial regression models.

Introduction to Time Series Analysis (R): This workshop emphasizes the practical aspects of time series analysis. Methods are hierarchically introduced, starting with terminology and exploratory graphics, moving to descriptive statistics, and ending with practical modeling procedures including how to choose an appropriate time series forecasting method, fit a model, evaluate its performance, and use it for forecasting. It focuses on the most popular business forecasting methods: regression models, smoothing methods including Moving Average (MA) and Exponential Smoothing, and Autoregressive (AR) models. Practical implementation in R is illustrated at each stage of the workshop.

Event History Analysis (Survival Models) in SPSS: This workshop using SPSS presents statistical methods of survival analysis, specifically focusing on studies where the outcome is a time-to-event variable. It covers the estimation of survival time using the life table and Kaplan-Meier Methods, as well as modeling survival risk. It also assesses the relationship of risk factors and survival times using the Cox regression model. SPSS 28.0 will be used for data analysis.

Event-Study Regression using R: Event-study is a causal inference research design method for analyzing the impact of a specific event on a particular outcome or variable of interest over a defined time period. The event can be considered as the treatment in a Difference-in-Difference (DiD) analysis, and the dynamics of the impact can be assessed by comparing the changes in outcomes over the time between the treated and control groups. This workshop will make use of a variety of R packages, specifically, *fixest*, *plm*, and *did* for event-study regression. Topics covered include data preparation, DiD analysis, dynamic DiD model, and the graphic display of the dynamic event effects.

Applied Statistics Workshops
Spring 2024

Spring 2024 Applied Statistical Workshop Schedule:

Topic	Date	Day	Time	Registration
Sample Size Estimation & Power Calculations (SAS)	Feb. 29	Thursday	10:00-12:00 P.M.	Virtual Register
Event History Analysis (Survival Models) in SPSS	Mar. 05	Tuesday	10:00-12:00 P.M.	In-Person Register --- Virtual Register
Intro to HLM (Mixed Models) (SPSS)	Mar. 18	Monday	10:00-12:00 P.M.	Virtual Register
Structural Equation Modeling I (AMOS & Mplus)	Mar. 19	Tuesday	10:00-12:00 P.M.	Virtual Register
COVID-19 Data Analysis Using R	Mar. 25	Monday	10:00-12:00 P.M.	In-Person Register --- Virtual Register
Intro to Statistical Learning (R)	Mar. 26	Tuesday	10:00-12:00 P.M.	In-Person Register --- Virtual Register
Structural Equation Modeling II (AMOS & Mplus)	Mar. 28	Thursday	10:00-12:00 P.M.	Virtual Register
Missing Data Analysis (SAS & Stata)	Apr. 03	Wednesday	10:00-12:00 P.M.	In-Person Register --- Virtual Register
Spatial Regression (R)	Apr. 04	Thursday	10:00-12:00 P.M.	Virtual Register
Intro to Time Series Analysis (R)	Apr. 09	Tuesday	10:00-12:00 P.M.	In-Person Register --- Virtual Register
Event-Study Regression using R	Apr. 11	Thursday	10:00-12:00 P.M.	Virtual Register

Registration Procedures:

Seats and handouts are limited. Please register in advance.

1. Click the 'In-person Register' or 'Virtual Register' under Registration.
2. Fill up all information request and submit your registration.
3. Join the workshops via Zoom link in the confirmation email or attend in-person sessions for hybrid workshops.

All Hybrid workshops will be held in Green Lab on the upper level (UL) of Healey Library.

Please contact Mr. Zihan Li at zihan.li001@umb.edu for any questions regarding to the workshops.

Web: https://www.umb.edu/academics/graduate/info_for_graduate_students/center_for_statistical_computing

Location: **Healey Library, Green Lab.** (From the main elevators in Healey Library, take the Upper level (UL). Turn right out of elevator, and you'll find Green Lab on the left in the hallway)