Bayesian Structural Equation Modeling:

An Educational Study Using PISA 2022 Data and Information Criteria

Group members: Sean Zhao, Zach Zhang, Zixie Zheng

Topic: Introduce Bayesian structural equation modeling

Model Selection:

Our project will be introducing the Bayesian approach in estimating structural equation models (SEMs). The models which we cover will include a confirmatory factor analysis model and a full structural equation model. We will begin with briefly discussing the difference between maximum likelihood estimation and Bayesian estimation of SEM and the motivating benefits of conducting SEM in the Bayesian approach (e.g., more direct explanations of statistical results based on raw observations rather than the sample covariance matrix). Prior specifications for model parameters will also be discussed. We will showcase the models with simulation and real data (PISA 2022 student questionnaire data). We will also discuss the adaptations of conventional SEM model goodness-of-fit indices (e.g., Bayesian root mean square error of approximation) and information criteria (e.g., AIC and WAIC) to Bayesian SEM as well as the Bayes factor for model evaluation and comparison.

Data:

We will utilize the 2022 student questionnaire data from the Programme for International Student Assessment (PISA). This data set includes a variety of indicators related to student performance across reading, mathematics, and science, providing a rich dataset for validating our SEM models. PISA measures 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges.

Methodology

Structural Equation Modeling (SEM) is a sophisticated statistical technique that combines factor analysis and multivariate regression analysis to explore complex causal relationships between variables. This model consists of two main parts: the measurement model and the structural model. The measurement model, similar to confirmatory factor analysis, defines the relationships between latent variables and their observed indicators, while the structural model explores the causal relationships among different latent variables. SEM is widely used across various fields, including psychology, education, social sciences, and market research, to analyze issues such as the impact of personality traits on mental health, the relationship between teaching methods and student performance, how socio-economic factors influence human behavior, and the dynamics between brand image and customer satisfaction.