PROGRAM
WORKSHOP ON DATA SCIENCE AND EDUCATION
INTERNATIONAL CONFERENCE ON DATA SCIENCE
ICDS 2023 CHILE

CHAIR: JORGE BAZÁN, UNIVERSITY OF SÃO PAULO, BRAZIL
ORGANIZERS: PAULA FARIÑA, ALBA MARTÍNEZ RUIZ, UDP, CHILE

NOVEMBER 7, 2023
AUDITORIUM FACULTY OF SOCIAL SCIENCES AND HISTORY[1]
UNIVERSIDAD DIEGO PORTALES

LOCAL TIME: SANTIAGO – CHILE, UTC -3

08.20 – 08.30	Welcome words. Chair Jorge Luis Bazán
08.30 – 09.15	Jorge Luis Bazán, University of São Paulo, Brazil
09.15 – 09.30	Questions and discussion
09.30 – 10.15	Alina A. von Davier, Duolingo, USA
10.15 – 10.30	Questions and discussion
10.30 – 11.00	Coffee break
11.00 – 11.30	Paula Fariña, Universidad Diego Portales, Chile
11.30 – 11.45	Questions and discussion
11.45 – 12.15	Inés Varas, Pontificia Universidad Católica de Chile
12.15 – 12.30	Questions and discusión
12.30 – 13.15	Walter L. Leite, University of Florida, USA
13.15 – 13.30	Questions and discussion

[1] Faculty of Social Sciences and History, UDP, Ejercito Street 333, Floor -1, Santiago, Chile

November 7, 2023

08.20 Welcome words. Chair: Jorge Luis Bazán, University of São Paulo, Brazil

08.30 Classification in educational data: Cognitive diagnostic models using different R packages

Jorge Luis Bazán
Department of Applied Mathematics and Statistics
University of São Paulo, Brazil

Abstract. In recent years the Cognitive Diagnosis Models (CDMs) have gain considerable space in literature. Different methods were already considered, taking also in account diverse scoring methodologies. CDMs are useful psychometric tools for identifying test-takers' profile or level of possession of a set of latent attributes underlying a latent variable; the latent variable may be a cognitive skill (say, mathematics achievement), a psychological trait, or an attitude. In this workshop we will talk about the use of Classical and Bayesian approach to the estimation of parameters of the Cognitive Diagnostic models (CDM) using different R packages. Specifically, we showed the codes to reproduce an application from the paper da Silva, de Oliveira, Davier and Bazán (2018) and give some comments about the use of this type of models in the Educational Assessment.

09.30 AI-Driven content generation for educational assessment: Implications for teaching, testing, and the future of education

Alina A. von Davier, Chief of Assessment Duolingo, USA
Honorary Research Fellow University of Oxford
Senior Research Fellow Carnegie Mellon University

Abstract. As artificial intelligence (AI) continues to advance, its applications in the realm of educational assessment are becoming increasingly significant. This presentation explores the potential of AI-driven content generation in educational assessment and its implications for teaching practices and the future of education. By leveraging the power of natural language processing and deep learning algorithms, large language models (LLMs) and other large computational models are now capable of generating contextually relevant, diverse, and high-quality content for educational assessments (text, images, animation, voice, etc). This revolutionizes the way educators and developers design, administer, and evaluate assessments, allowing for greater efficiency and a more personalized learning and testing experience for students. The implementation of AI-driven content generation in testing presents numerous opportunities for teachers, students, and test developers from the assessment industry.

For teachers, it offers the potential to streamline classroom longitudinal quiz creation, reduce bias, and improve the validity and reliability of these evaluative data. For students, it promises a more engaging and adaptive assessment experience, tailored to their individual learning needs and preferences. For test developers it offers an efficient way to scale up the number of items needed to protect the security of the test. However, the integration of AI in educational assessment also raises several concerns and challenges. These include issues of construct relevance, cheating, data privacy and security, the potential for perpetuating existing inequalities in education, and the ethical considerations surrounding the use of AI-generated content. In this presentation I will provide an analysis of the current state of AI-driven content generation in educational assessment, discuss its potential impact on teaching practices, and present a vision for the future of education in light of these advances. I will illustrate the application of LLMs for generating test questions within the theoretical ecosystem of the digital-first assessments such as the Duolingo English Test (DET) and discuss the newly developed DET Responsible AI Standards. Ultimately, I hope to contribute to a meaningful dialogue on how AI can be harnessed to revolutionize educational assessment and teaching practices while addressing the associated ethical and societal concerns.

11.00 Bayesian networks in computerized adaptive test for statistical learning

Paula Fariña
School of Industrial Engineering, Faculty of Engineering and Sciences
Universidad Diego Portales, Chile

Abstract. Bayesian Networks are a powerful tool for modeling complex relationships between variables in various fields, including education. In particular, they are increasingly being used in Computerized Adaptive Learning (CAL) to personalize the learning experience for students. By incorporating Bayesian Networks in CAL, the system can adapt to the student's needs, abilities, and learning preferences, providing a more effective and efficient learning experience. In this workshop we will explore the use of Bayesian Networks, including how they can be used to model student knowledge, track learning progress, and provide personalized feedback and recommendations. A CAL App designed for Statistical Learning is also presented as an example.

11.45 Latent models for linking measurements

Inés Varas
Department of Statistics
Pontificia Universidad Católica de Chile, Chile

Abstract. Equating is the most popular linking method used to adjust scores on different test forms so that scores can be used interchangeably. These methods map the scores of test form X into their equivalents on the scale of test form Y by using scores distributions. Equating methods tackle differences in distributions attributed to differences in the difficulty of the forms. To overcome differences in the score distributions attributed to differences in the ability of test takers different data collection designs are considered. Although test score scales are usually subsets of integer numbers, in the equating literature the mapping estimation is based on continuous approximations of score distributions. Thus, equated scores are no longer discrete values. Varas et al. (2019, 2020) proposed the latent equating method to obtain discrete equated measurements based on a latent representation of scores distributions and a Bayesian nonparametric model for it. An extension of the latent method is proposed to be applied on different sampling designs. It is included the non-equivalent anchor test design (NEAT) where common items are used to link scores of test takers sampled from different populations. Several methods are discussed to evaluate the performance of the extension applied to simulated and real datasets.

12.30 Treatments as latent variables: Combining machine learning with latent variable modeling to estimate average treatment effects

Walter L. Leite
Professor & Program Coordinator
Research and Evaluation Methodology (REM) Program
School of Human Development and Organizational Studies in Education
College of Education
University of Florida, USA

Abstract. Machine learning methods have been increasingly used to improve causal inference in experimental and quasi-experimental designs. The objective of this talk is to demonstrate some applications of machine learning to estimate the effect of a treatment defined as a latent variable. In the examples provided, latent variable models are combined with machine learning to measure individual exposure to a treatment conceptualized as either a categorical or continuous latent variable. Then, machine learning methods combined with propensity score methods are used to remove selection bias and estimate the average treatment effect.

SUPPORTING INSTITUTIONS - ENDORSEMENTS