30th Young Statisticians Meeting

YSM30

Program and Abstracts

October 2–4, 2026

Mönichwald, Austria

Supported by

Austrian Statistical Society
Graz University of Technology, Institute of Statistics
Land Steiermark
TU Wien, Computational Statistics

International Program Committee

Paola Berchialla (University of Torino, Italy)
Ksenija Dumičić (University of Zagreb, Croatia)
Peter Filzmoser (TU Wien, Austria)
Luka Kronegger (University of Ljubljana, Slovenia)
David Simon (University of Budapest, Hungary)
Nenad Suvak̆ (University of Osijek, Croatia)

Local Organizer

Peter Filzmoser (TU Wien, Austria)

PROGRAM

Friday, October 2nd, 2026

17.00–19.00 Registration at Seegasthof Breineder, Mönichwald

Saturday, October 3rd, 2026

8:50–9:00 Welcome and Opening
Peter Filzmoser, TU Wien, Austria

Session 1: **Chair: Peter Filzmoser**

9.00–9.30 Robust and explainable outlier detection for random surfaces
Leopold Micheler, TU Wien, Austria

9:30–10:00 A randomized convex optimization method for highly constrained statistical learning
Balázs Miavecz, Hungarian Research Network and Eötvös Loránd University, Hungary

10:00–10:30 Locating the best-subset model within the multiverse: a quantile-based diagnostic for the robustness of automatic variable selection
Anna Sordo, University of Padua, Italy

10.30–11.00 Coffee Break

Session 2 **Chair: Ksenija Dumičić**

11.00–11.30 One reference rate, many pass-throughs: An ARDL analysis of monetary policy transmission in the Euro area
Maksimilijan Balatinec, University of Zagreb, Croatia

11.30–12.00 A new methodology for the meta-analysis of discrete choice experiments
Matteo Bracco, University of Turin, Italy

12.00–13.30 Lunch

Session 3 **Chair: Luka Kronegger**

13.30–14.00 Monte Carlo evaluation of methods for estimating regression coefficients involving latent variables
Katja Škoda, University of Ljubljana, Slovenia

14.00–14.30 Efficient Bayesian inference for polytomous item response models
Stefan Haan, University of Klagenfurt, Austria

14.30–15.00 Model misfit and parameter estimate bias in structural equation modelling
Jošt Bartol, University of Ljubljana, Slovenia

15.00–15.30 Coffee Break

Session 4 **Chair: David Simon**

15.30–16.00 Reading the economy through the News: AI-assisted narrative and sentiment analysis of media text
Luka Šikić, Croatian Catholic University, Croatia

16.00–16.30 AI and physics-based weather forecasting: A comparison across raw and post-processed operational ensembles
Mátyás Kocsis, University of Debrecen, Hungary

19.00 Reception at Seegasthof Breineder, Mönichwald

Sunday, October 4th, 2026

Session 5 **Chair: Paola Berchialla**

9.00–9.30 Long-term risk prediction from short-term data – a microsimulation approach
Moritz Pamminger, Medical University of Vienna, Austria

9.30–10.00 Multifractal analysis of EEG signals for predicting neurodevelopmental outcomes after pediatric cerebral malaria
Aleksandra Petrović, J. J. Strossmayer University of Osijek, Croatia

10.00–10.30 Lead time estimation in cancer screening programmes: extending the MOCCI method
Ema Požek, University of Ljubljana, Slovenia

10.30–11.00 Coffee Break

Session 6 **Chair: Nenad Suvaǩ**

11.00–11.30 Predicting Formula 1 overtaking outcomes: Development and comparison of Machine Learning models for decision support
Balázs Tobak, Eötvös Loránd University, Hungary

11.30–12.00 A two-phase Bayesian model for longitudinal risk estimation of anterior knee pain in semi-professional basketball players
Marco Vitturini, University of Turin, Italy

12.00 Closing

ABSTRACTS

(sorted in alphabetic order of the presenter's family name)

One reference rate, many pass-throughs: An ARDL analysis of monetary policy transmission in the Euro area

Maksimilijan Balatinec

Faculty of Economics and Business, University of Zagreb, Croatia

A common monetary policy, such as the one operating in the euro area, cannot respond to country-specific shocks; it reacts instead to conditions in the currency area as a whole. This can lead to lasting differences in inflation rates across member states. One of the main routes by which this policy reaches national economies is the interest-rate channel. By changing its key reference rates, the European Central Bank (ECB) sets off movements in money market and interbank rates, which then feed through to the rates that commercial banks set on loans to, and deposits from, households and non-financial corporations. This mechanism is known as interest rate pass-through. A complete pass-through would mean the interest rate channel works fully and evenly, while an incomplete, sluggish, or uneven pass-through would mean the common stance is not passed on in the same way across all member states. This is a second source of divergence: even the single stance the ECB sets is not taken up evenly across national banking systems. The outcome is a kind of monetary policy that is identical on paper yet unequal in fact: during a tightening cycle, a country with lower and slower pass-through tends to be left with higher inflation than one where transmission is stronger, even though they both face the same ECB reference rates.

This study examines the uniformity of interest rate pass-through, drawing on a sample of monthly data that runs from the 2022 inflation surge to the present (a period that includes the steepest monetary tightening cycle in the history of the euro) and covering countries that were euro area members throughout the sample. Using an ARDL (autoregressive distributed lag) modelling framework, the study measures the degree and speed of pass-through across different categories of bank loan and deposit rates and describes how this varies across member states, since this variation indicates how uniform the interest rate channel is. It further examines whether pass-through is symmetric across the recent restrictive and expansionary policy regimes, or whether it instead exhibits a rockets-and-feathers pattern, with bank rates adjusting more readily in one direction than the other. Pronounced differences or asymmetry in pass-through would show where the common stance reaches member states evenly and where national transmission, rather than policy itself, accounts for divergence.

Model misfit and parameter estimate bias in structural equation modelling

Jošt Bartol

University of Ljubljana, Faculty of Social Sciences, Slovenia

Structural equation modelling is a widely used data analysis method. Its key strength lies in its flexibility to simultaneously assess the validity and reliability of measures and test substantive hypotheses. In the social sciences, a typical use case is validating the survey scales used to measure theoretical constructs and subsequently examining the associations among these constructs. However, to ensure that the results obtained from such analyses are valid, researchers must evaluate the fit of the model to the data. Only when fit is adequate may the results be interpreted. This is done by examining the discrepancy between the sample covariance matrix and the model-implied covariance matrix. If this discrepancy is large, the model should be rejected or modified. Typically, researchers use fit indices (e.g. CFI, RMSEA, SRMR) to assess model fit. These indices, however, indicate only the average or global level of (mis)fit, failing to capture what is considered local (mis)fit, or the localised mismatch between the two matrices. For large models, an overall good average fit can mask localised misfit, leading analysts to accept suboptimal models with potentially biased estimates of the parameters of interest. However, studies have rarely systematically examined the relationship between global and local fit, or the extent to which localised misfit biases estimates of parameters of interest. Therefore, the aim of this research was to conduct a simulation study investigating the amount of relative bias in parameter estimates produced by models with different degrees of misfit. We also evaluated whether global fit indices would reject misspecified models. The results suggest that global fit indices cannot identify models with biasing levels of misfit. In such cases, the relative bias for certain parameter estimates averaged around 50%. Based on these results, we discuss approaches to assessing model fit in structural equation modelling.

A new methodology for the meta-analysis of discrete choice experiments

Matteo Bracco[1], Emanuele Pietropaolo[1], Martina di Blasio[1], Ileana Baldi[2], Alessia Visconti[1], Paola Berchialla[1]

[1] *Center for Biostatistics, Epidemiology, and Public Health, Department of Clinical and Biological Science, University of Turin, Italy*
[2] *Department of Cardiac Thoracic Vascular Sciences and Public Health, University of Padua, Padua, Italy*

Introduction: Understanding patients' priorities and preferences is essential to inform decision-making process along the whole medical products life cycle. Discrete Choice Experiments (DCEs) represent one of the most robust and wide-spread tools used to elicit patient preferences across healthcare literature. DCEs are able to estimate patient utility functions defined over a set of fixed treatment related attributes (e.g. efficacy, administration route, side effect risk), each one declined with multiple levels (e.g. administration route: oral, injection). Utilities functions are determined by estimating a set of coefficients called Preference Weights (PWs). PWs are used to reconstruct complex preference patterns such as benefit-risk trade-offs, willingness-to-pay and the relative importance of treatment attributes. Developing DCEs, however, remains expensive and time-consuming, highlighting the need for evidence synthesis via meta-analysis.
Objectives: As PWs cannot be directly compared across DCEs, as they are estimated on different latent scales, the current meta-analytic approach works instead with scale-free trade-off parameters. Yet this requires the identification of at least two common attributes for each pooled study to derive the trade-off parameter of interest, hence limiting the possible therapeutic areas of application as well as the DCEs pool's size. We propose a novel methodology stemming from network theory to re-scale PWs allowing their comparison across DCEs and hence enabling evidence synthesis on the higher level of PWs.
Methods: We define a network of DCEs where links are constructed based on common attributes' levels and their corresponding PWs found between two nodes. By sampling spanning trees via a Markov-Chain Monte Carlo (MCMC) procedure, we are able to create samples of re-scaled PWs sets where we can conduct aggregation and evaluate distributions for trade-off parameters of attribute importance rankings. This flexible approach allows to include multiple DCEs in the meta-analytic pool and to evaluate different trade-off parameters of interest from the PWs synthesis, including non-linear ones.
Conclusion: Our framework represents a key step towards more robust and resource-efficient modelling of patients' preferences. The production of multiple aggregated preference parameters outputs defines a comprehensive picture of patient preferences and priorities across different therapeutic areas, guiding the development of new pharmaceutical products and regulatory decision-making processes which account for patient real needs.

Efficient Bayesian inference for polytomous item response models

Stefan Haan[1,2], Gregor Kastner[1], Heather Foran[2], Michael Radloff[2]

[1] *Department of Statistics, University of Klagenfurt, Austria*
[2] *Health Psychology Unit, University of Klagenfurt, Austria*

In psychological and medical research, multi-item questionnaires are often used to assess various outcomes such as well-being, depression or family functioning. Such data is frequently analyzed using sum-scores of ordinal values, implicitly assuming equal intervals between levels and equal measurement precision of all items. Liddell and Kruschke [3] show that treating data generated from the ordered probit model as metric can lead to Type I and II errors as well as systematic inversion true effect sign and its estimate. This can happen even when the thresholds of the ordered probit model are evenly spaced or responses are averaged over multiple items.

To analyze ordinal item responses we build on the graded response model (see [2]): We assume each person has a unidimensional person parameter that is the mean of underlying continuous responses. The support of the continuous responses is divided via cutpoints into ordinal item responses. The variances of the continuous responses as well as the cutpoints are allowed to vary by item.

We extend the graded response model by two features: (1) To estimate the treatment effect of an intervention, where item responses are collected pre- and post-treatment, we add a latent regression for the difference of the person parameter to our model. (2) We try to detect guessed answers by introducing guessing proneness parameters for persons and items. If an answer was guessed, we assume an uniform likelihood for the observation, instead of one driven by the person parameter.

We estimate all model parameters jointly in a fully Bayesian approach, which is advantageous in multiple ways: (1) We do not use a two-step approach and thus carry over parameter uncertainty in a principled manner. For example, if we suspect an answer was guessed, MCMC will incorporate this suspicion coherently into the uncertainty of the treatment effect. (2) For maximum likelihood methods, joint estimation of both person and item parameters is not recommended since estimates are not consistent [4]. Instead, person parameters are treated as nuisance parameters and integrated out. Group contrasts can still be computed by introducing virtual items, but inclusion of arbitrary (pre-treatment) person-level covariates is inherently limited.

While Bayesian estimation via Gibbs sampling might seem straightforward, some carefulness is required to achieve efficient mixing. We follow the advice of Cowles [1] and integrate out the continuous responses when sampling the cutpoints, as failing to do so results in very slow convergence. We also reparameterize the cutpoint parameters following the ideas outlined in Nandram and Chen [5].

Finally, we apply our model to a large randomized trial dataset (FLOURISH[1]), which evaluates the effectiveness of a family-based intervention on improving adolescents' and parent's mental health and well-being in two European countries.

[1] `https://www.flourish-study.org`. This research received financial support from the European Union Horizon Europe (grant agreement number 101095528).

References

[1] M. K. Cowles (1996). Accelerating Monte Carlo Markov chain convergence for cumulative-link generalized linear models. *Statistics and Computing*, **6**(2), 101–111.

[2] J.-P. Fox (2010). *Bayesian Item Response Modeling: Theory and Applications*. Springer.

[3] T. M. Liddell and J. K. Kruschke (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, **79**, 328–348.

[4] P. Mair, R. Hatzinger and M. J. Maier (2025). Extended Rasch Modeling: The R Package eRm. Vignette included in R package eRm, version 1.0-10. `https://CRAN.R-project.org/package=eRm`.

[5] B. Nandram and M.-H. Chen (1996). Reparameterizing the generalized linear model to accelerate Gibbs sampler convergence. *Journal of Statistical Computation and Simulation*, **54**(1-3), 129–144.

AI and physics-based weather forecasting: A comparison across raw and post-processed operational ensembles

Mátyás Kocsis[1,2], Sándor Baran[2]

[1] *Doctoral School of Informatics, University of Debrecen, Hungary*
[2] *Faculty of Informatics, University of Debrecen, Hungary*

Numerical weather prediction models describe the dynamical and physical behaviour of the atmosphere and are the primary tools for providing weather forecasts. A prominent example is the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF) which is considered the gold standard in weather forecasting. The state-of-the-art method is to run the models multiple times with different initialisation/parametrisation to obtain an ensemble of forecasts, allowing for the assessment of forecast uncertainty and hence opening the door for probabilistic forecasting. Despite their numerous advantages over single forecasts, raw ensembles are usually under/overdispersive and possess systematic biases, necessitating some form of post-processing. This need led to the development of a broad range of statistical methods, making statistical post-processing a crucial part of atmospheric sciences and applied statistics. In recent years, machine learning-based weather forecasting models have taken center stage in research due to their impressive efficiency. The ECMWF has placed itself at the forefront of AI-based weather forecasting as it was first in the world to launch an operational AI model, the so-called Artificial Intelligence/Integrated Forecasting System (AIFS). The deterministic and probabilistic counterparts of this model, AIFS-Single and AIFS-CRPS have been operational since 2025 February 25 and 2025 July 1, respectively. The talk aims to present a systematic comparison of the ECMWF IFS and AIFS-CRPS systems in the light of raw and post-processed 10 m wind speed ensemble predictions at more than 9000 observation stations across the globe. In the post-processed case, a parametric Ensemble Model Output Statistics method based on the truncated normal distribution has been used and a nonparametric quantile regression approach has also been applied to enhance the calibration of the ensembles. According to various verification scores, raw IFS forecasts perform significantly better at all lead times than the corresponding raw AIFS predictions, while post-processing paints a different picture as in this case IFS maintains its significant advantage only at short forecast horizons [1].

References

[1] M. Kocsis and S. Baran (2026). AI and physics-based weather forecasting: A comparative study. *arXiv:2606.02508*.

A randomized convex optimization method for highly constrained statistical learning

Balázs Miavecz[1,2], Balázs Csanád Csáji[1,2]

[1] *Institute for Computer Science and Control (SZTAKI), Hungarian Research Network (HUN-REN), Budapest, Hungary*
[2] *Department of Probability Theory and Statistics, Institute of Mathematics, Eötvös Loránd University (ELTE), Budapest, Hungary*

This work addresses the computational difficulties of solving highly constrained convex optimization problems in statistical learning, where the number of constraints vastly exceeds the number of decision variables. Such problems arise naturally in domains such as support vector machines (SVMs), reinforcement learning (RL), and stochastic model predictive control (MPC). We introduce Batch Optimization via Softmax Selection (BOSS), a novel cutting-plane method designed specifically for this setting. Unlike traditional approaches that process the entire constraint set simultaneously, BOSS iteratively constructs the support constraint set using geometric pruning heuristics, probabilistic sampling of constraint violations via the Boltzmann-Gibbs distribution, and a dynamically expanding bounded truncation domain for subproblems. Crucially, while BOSS relies on randomization, it is a Las Vegas algorithm that guarantees termination with an exact optimal solution. Additionally, we develop a distributed implementation of BOSS to further improve its scalability. Numerical experiments on random Linear Programs (LPs), Minimum Volume Enclosing Ellipsoids (MVEE), and soft-margin SVMs with over a million constraints demonstrate that, for these types of problems, BOSS can significantly outperform state-of-the-art general-purpose solvers, including Gurobi, Clarabel, and HiGHS.
This research was supported by NKFIH via the ADVANCED project 153390.

Robust and explainable outlier detection for random surfaces

Leopold Micheler, Marcus Mayrhofer, Una Radojicic, Peter Filzmoser

Institute of Statistics and Mathematical Methods in Economics, TU Wien, Austria

This contribution introduces a new distance measure and a robust method for covariance estimation of random surfaces with a separable covariance structure. The approach links separable random surfaces to the matrix-variate distribution of their basis function representations, which provides a convenient and principled framework for estimation.
Building on this connection, we develop a robust procedure based on the Matrix Minimum Covariance Determinant (MMCD) [1] estimator, combined with a truncated functional Mahalanobis semi-distance to estimate mean and covariance functions. This formulation is designed to remain stable under contamination and to provide reliable distance-based inference for functional data.
To improve interpretability, we extend the Shapley value [2] methodology to the functional setting. This allows us to decompose the proposed distance measure and thus functional outlyingness into contributions from specific spatial and temporal regions. We further introduce a novel formulation for computing these functional Shapley values while preserving their fundamental axiomatic properties.
The resulting framework integrates matrix-variate modeling, robust distance measures, and interpretable decomposition techniques into a unified approach for analyzing random surfaces. We provide theoretical justification alongside empirical results on real-world datasets, demonstrating strong performance in classification and robust outlier detection tasks.

References

[1] M. Mayrhofer, U. Radojičić and P. Filzmoser (2025). Robust covariance estimation and explainable outlier detection for matrix-valued data. *Technometrics*, **67**(3), 516–530.

[2] L. S. Shapley (1953). A Value for n-Person Games. *Contributions to the Theory of Games, Volume II*, 307–318.

Long-term risk prediction from short-term data – a microsimulation approach

Moritz Pamminger

Institute of Clinical Biometrics, Center for Medical Data Science, Medical University of Vienna, Austria

Long-term survival prediction is particularly relevant in areas such as cardiovascular disease, where age is a highly important prognostic factor. Young individuals may face a low short-term (5-10 years), but high long-term (20-30 years) risk. Aiming to build models capable of making such long-term predictions, researchers must rely on the availability of long-term observational data. However, such data is rare and necessarily old. I aim to describe, implement and evaluate a method that facilitates long-term survival predictions using short-term contemporary big data.

I assume a data set including information on time-to-event and repeated measurements of prognostic factors (e.g., cholesterol, BMI, blood pressure) with a limited follow-up period (e.g., 5 years), such as data collected from routine health screenings. In principle, one could use joint models for survival and longitudinal markers with age as time scale to obtain long-term predictions from such data. Joint models are, however, computationally demanding, in particular with multiple longitudinal markers and age as time scale. The aim of my project is to develop and evaluate a microsimulation approach to long-term risk prediction. Specifically, the approach consists of fitting short-term generative prediction models and of applying them iteratively to generate long-term trajectories of prognostic factors and survival episodes, ultimately leading to long-term predictions of survival. To account for prediction uncertainty that increases at each iteration, for each individual multiple trajectories are simulated, that can then be appropriately analysed and graphically illustrated.

The proposed methodology has been evaluated in an initial simulation study based on synthetic data. This study considers different scenarios, including varying sample sizes and alternative specifications of the short-term models, and provides a first assessment of the accuracy and robustness of the method.

The method is currently being implemented in an R-package that can perform microsimulation semi-automatically. The package will support various model classes, including linear models, generalized linear models, and ordinal regression models. Flexible specifications of these models with nonlinear functional forms specified as splines and accommodating interaction terms between predictors will allow users to tailor the framework to their specific application.

This project is the subject of my master's thesis and is part of the ongoing research project RIPOSTE-BD, which has received support from the Austrian Science Fund (FWF), grant P-36727-N.

Multifractal analysis of EEG signals for predicting neurodevelopmental outcomes after pediatric cerebral malaria

Aleksandra Petrović

School of Applied Mathematics and Informatics, J. J. Strossmayer University of Osijek, Croatia

Electroencephalographic (EEG) recordings provide an important approach for investigating brain dysfunction in neurological disorders. In this study, multifractal detrended fluctuation analysis (MF-DFA) is applied to EEG recordings from Ugandan children in a coma due to cerebral malaria to extract quantitative descriptors of EEG dynamics. The resulting features, including the Hurst exponent, dominant singularity strength, multifractal spectrum width, and asymmetry parameter, are evaluated together with clinical, laboratory, sociodemographic, and baseline neurodevelopmental predictors. The results suggest that selected MF-DFA-derived EEG features may provide complementary information for predicting six-month neurodevelopmental outcomes after pediatric cerebral malaria.

Lead time estimation in cancer screening programmes: extending the MOCCI method

Ema Požek, Bor Vratanar

Institute for Biostatistics and Medical Informatics, Faculty of Medicine, University of Ljubljana, Slovenia

Breast cancer screening programmes aim to detect disease at an early stage, when treatment options may be broader and chances of recovery greater. However, evaluating the benefits of screening programmes is not straightforward, as naïve comparisons of survival of those invited and those not invited to the programme are biased. Apparent improvements in survival among screen-detected cases may reflect earlier diagnosis (lead time bias), or detection of cases that would never have been detected in the absence of screening (overdetection).

The MOCCI method was developed to jointly estimate lead time and overdiagnosis, enabling better understanding of the disease and improving the survival-based evaluation of screening. Based on comparison of cancer incidences between invited and non-invited groups, the method aims to identify lead time and overdiagnosis distributions that best explain observed differences using maximum likelihood principles.

In this work, we extend the MOCCI method by integrating the biological tumour growth model. Alongside age and calendar time at diagnosis, tumour size is introduced as a third modelling dimension, shifting the focus toward estimating latent processes (e.g. tumour growth rate), from which lead time and overdiagnosis can be derived. We outline the extended MOCCI estimation procedure and present results of simulation study assessing the feasibility of the proposed approach.

References

[1] B. Vratanar and M. Pohar Perme (2023). Evaluating cancer screening programs using survival analysis. *Biometrical Journal*, **65**(7), e2200344.

[2] B. Vratanar and M. Pohar Perme. Estimating lead time and overdiagnosis in cancer screening programmes: the MOCCI method. Under review.

[3] G. Ishedén and K. Humphreys (2019). Modelling breast cancer tumour growth for a stable disease population. *Statistical Methods in Medical Research*, **28**(3), 681–702.

Reading the economy through the News: AI-assisted narrative and sentiment analysis of media text

Luka Šikić

Croatian Catholic University, Zagreb, Croatia

Economic decision-making is increasingly shaped by how events are framed in the media, yet the resulting text data is high-dimensional, unstructured, and hard to quantify at scale. This talk presents a reproducible, AI-assisted pipeline that turns large volumes of news text into structured indicators of narrative and sentiment, and links them to economic outcomes. Drawing on recent work on media framing during Croatia's euro adoption and on native advertising and clickbait in digital content, I illustrate how modern natural language processing (NLP) and machine-learning tools can be combined with classical time-series and network methods to extract, validate, and forecast with text-based signals. The emphasis is on a transparent, end-to-end workflow – from data collection to model and result – that keeps large-language-model tooling reproducible and auditable for empirical economic research.

Monte Carlo evaluation of methods for estimating regression coefficients involving latent variables

Katja Škoda

University of Ljubljana, Ljubljana, Slovenia

Structural equation modeling (SEM) is the standard approach for estimating regression relationships between latent variables because it simultaneously models measurement error and structural relations. However, alternative approaches are often used in applied research due to their simplicity or because SEM may encounter convergence problems, particularly in small samples. Common alternatives include regression based on estimated latent variables, which can produce biased regression coefficient estimates due to measurement error. A recently proposed alternative is Structural-After-Measurement (SAM), a two-step approach that accounts for measurement error while potentially reducing nonconvergence and improper solutions. Despite its potential advantages, SAM remains relatively unexplored, particularly for ordinal indicators commonly used in social science research.

This study presents a Monte Carlo simulation comparing four approaches for estimating regression coefficients involving latent variables: regression using standardized mean scores, regression using Empirical Bayes Modal (EBM) factor scores, SAM, and SEM. The simulation systematically varies sample size, factor loadings, number of indicators, indicator scale type (continuous vs. ordinal), indicator distribution, and the role of the latent variable in the regression model. Performance is evaluated in terms of the bias, and root mean squared error (RMSE) of the standardized regression coefficient, as well as the frequency of nonconvergence and improper solutions. Preliminary results suggest that estimation accuracy is primarily affected by sample size, factor loadings, and the number of indicators. Overall, SEM and SAM outperform regression approaches based on estimated latent variables, producing less biased and more accurate estimates of the standardized regression coefficient.

References

[1] N. Bhaktha and C. M. Lechner (2021). To score or not to score? A simulation study on the performance of test scores, plausible values, and SEM, in regression with socio-emotional skill or personality scales as predictors. *Frontiers in Psychology*, **12**, 679481.

[2] R. B. Kline (2023). *Principles and Practice of Structural Equation Modeling*. Guilford Publications.

[3] Y. Rosseel and W. W. Loh (2024). A structural after measurement approach to structural equation modeling. *Psychological Methods*, **29**(3), 561.

[4] A. Skrondal and P. Laake (2001). Regression among factor scores. *Psychometrika*, **66**(4), 563–575.

[5] A. Skrondal and S. Rabe-Hesketh (2004). *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. Chapman and Hall/CRC.

[6] R. Van de Schoot and M. Miočević (2020). *Small Sample Size Solutions: A Guide for Applied Researchers and Practitioners*. Routledge.

Locating the best-subset model within the multiverse: a quantile-based diagnostic for the robustness of automatic variable selection

Anna Sordo

Department of Cardiac, Thoracic, Vascular Sciences and Public Health, University of Padua, Italy

Background. Variable selection is a decisive and under-scrutinized step in model building. Best subset selection returns the single covariate combination optimizing a chosen criterion (AIC, BIC, adjusted R^2), but this optimum is a point summary of an objective surface: because the criterion is itself estimated from finite data, small perturbations of the sample can shift the local minimum to a different subset. The reported model may therefore be unstable even when its in-sample fit is not, and the uncertainty induced by the selection step is not reflected in the inference conditional on the selected model.
Methods. Multiverse analysis enumerates all 2^p subsets of the p candidate covariates and retains, for each specification, the quantity of inferential interest, e.g. the estimate and Wald p-value of the focal coefficient, replacing a single number with an empirical distribution over the specification space. We use this distribution not to arbitrate which specification is correct, but as a reference against which the best-subset result can be positioned. Concretely, we treat the selected model's estimate (and p-value) as one draw from the multiverse and compute its empirical quantile within the specification curve. A result near the median identifies the selected model as a typical specification, largely interchangeable with its neighbors; a result in the tails flags it as an outlier whose apparent optimality may be an artefact of the selection step, warranting explicit caution before it is reported as a conclusion.
We characterize this quantile position through a factorial Monte Carlo study with a known data-generating mechanism, crossing the number of candidate covariates p, which controls both the cardinality of the multiverse and the degree of collinearity among specifications, with the sample size n. For each scenario we report the distribution of the quantile of the best-subset result, the proportion of replicates in which that result falls in the tails of the specification curve, and the associated sign-flip and significance-flip rates.
Results and discussion. The design quantifies how the location of the best-subset model within the specification distribution degrades as dimensionality grows and stabilizes as n increases. The resulting diagnostic is inexpensive, requires no assumption beyond those already made by the selection procedure, and can be reported alongside any automatically selected model. It yields practical, scenario-specific guidance on when a best-subset result can be taken at face value and when the multiverse should be reported in its place.

Predicting Formula 1 overtaking outcomes: Development and comparison of Machine Learning models for decision support

Balázs Tobak

Eötvös Loránd University (ELTE), Budapest, Hungary

The study predicts Formula 1 overtaking outcomes using race data from the 2018–2025 seasons. It compares logistic regression, XGBoost, and neural networks with entity embeddings with respect to predictive accuracy, model transparency, and their suitability for real-time decision support. Current results indicate that XGBoost achieves the best predictive performance and stability among the evaluated models. Ongoing work focuses on extending the framework with architectures capable of capturing temporal dynamics in race data. The research also emphasizes Explainable AI to improve model transparency by identifying the key factors driving overtaking outcomes.

A two-phase Bayesian model for longitudinal risk estimation of anterior knee pain in semi-professional basketball players

Marco Vitturini

Centre for Biostatistics, Epidemiology and Public Health, Department of Clinical and Biological Sciences, University of Turin, Italy

Anterior knee pain (AKP) is a frequent overuse condition in jumping-intensive sports, and basketball's repeated jump-landing demands make it a persistent threat to training continuity and career longevity. Its burden is easily underestimated, since standard sport-stop registers only capture episodes severe enough to halt play, missing the sub-clinical loading that precedes them. Turning surveillance into athlete-level prevention is also hard: club populations are small, risk combines physical, psychosocial and behavioural factors, and existing screening tools remain scarce and cross-sectional.

We applied a two-phase Bayesian model in which cross-sectional associations informed prior distributions for exploratory individualized longitudinal risk estimation in a small pilot study. In phase one, a Bayesian logistic model related AKP to prespecified demographic, sporting and behavioural covariates in 371 male semi-professional players aged 18–35. In phase two, a 24-athlete observational pilot was followed bi-monthly across four waves. Longitudinal priors reused the complete phase-one posterior draws, perturbed with small Student-t terms to let effects evolve over time. Incident AKP among event-free players was modelled through a discrete-time hazard with wave-specific baselines and time-varying covariates, so the model updates next-interval risk from an athlete's latest measurements and projects risk over any chosen horizon within that season. We compared posterior sampling efficiency under priors informed by phase one and under weakly informative priors using the posterior effective sample size calculated from 60,000 draws. The small number of at-risk person-periods and incident events resulted in very low posterior sampling efficiency under weakly informative priors. When phase-one posterior information was used to construct the priors, the posterior effective sample size for the predictor coefficients increased from 3 to 1063 out of 60,000 draws, indicating improved posterior sampling efficiency. Training load and its change, smoking status and playing experience showed the strongest posterior associations with risk, and the model produced individualized, updateable season-risk projections for each athlete.

In this pilot analysis, incorporating phase-one posterior information facilitated the fitting of a full-covariate longitudinal Bayesian model and the estimation of individualized AKP risk trajectories. These findings are exploratory and require validation in larger longitudinal cohorts, together with a formal assessment of sample-size requirements and predictive performance.