Universidade Federal do Rio Grande do Sul
Instituto de Matemática e Estatística
Departamento de Estatística

Anais de Resumos

XI SEMANÍSTICA

XI Semana Acadêmica do Departamento de Estatística

da UFRGS

http://www.ufrgs.br/semanistica

Porto Alegre - 27, 28 e 29 de outubro de 2021

Sottile et al. (2020) use this approach to perform LASSO variable selection using information on all quantiles simultaneously. In this work, we propose a global coefficient estimation and variable selection method based on the estimator presented in Sottile et al. (2020), introducing the group LASSO penalty, suggested in Yuan and Lin (2006), and applied in a Quantile Autoregressive Distributed Lag (QADL) model. Furthermore, since we are in a time series context, we also evaluate the variable selection penalization applying higher penalties to higher lags, as proposed in Konzen and Ziegelmann (2016). The results suggest that a weighted penalized approach can provide better results in selecting the variables as well as in estimating the coefficients. In particular, both LASSO and group LASSO penalization with higher weights for higher lags were the ones that had lower mean squared error to estimate most of the tested scenarios and set the zero coefficients correctly more often.

Mini Palestra 3:

Correlation Selection in phylogenetic multivariate probit models

Felipe Grillo Pinheiro, Taiane Schaedler Prass e Gabirela Bettella Cybis

Resumo: The multivariate phylogenetic latent liability model, first proposed by Cybis et al. (2015), and the recent phylogenetic multivariate probit model (PMPM), developed by Zhang et al. (2021), are important tools for investigating the association structure between mixed-type biological traits controlling for the shared evolutionary history of related organisms. We model these associations through the correlation matrix, R, of the latent Gaussian variables in a multivariate Brownian diffusion process along a phylogenetic tree informed by molecular sequences. However, besides the well-known limitation of parameter identifiability in probit models, another difficulty lies in the arbitrary criteria used to determine significance of these associations. Correlations have been considered significant if a chosen percentage highest posterior density (HPD) interval does not contain zero. Estimating sparse correlation matrices provides both, a systematic solution for elimination of spurious correlations and parameter reduction, which is a major gain since the number of parameters scales quadratically in trait dimension. However, due to model assumptions or identifiability reasons, when the covariance matrix is assumed to be a correlation matrix, as in probit models, the options for prior distributions on R are limited, especially if one requires sparsity in addition. To bypass this limitation, Bayesian inference for probit models is usually performed using the data augmentation representation of Chib and Greenberg (1998), where the binary traits, that require unit variance to be identifiable, are rescaled. Consequently, the correlation matrix is expanded to a covariance matrix that can be modelled using standard conjugate priors

(on the covariance or on its inverse, the precision matrix K) before being projected back to a correlation matrix. We propose a Bayesian approach for inference of sparse inverse correlation matrices in PMPM via the parameter expansion data augmentation strategy. We model the precision matrix (associated with the expanded correlation matrix) with a G-Wishart conjugate prior in the context of decomposable graphs. This prior choice allows us to explore conditional independence between traits which results in a sparse K. We obtain a final sparse R through a decomposition on each sampled K.

Mini Palestra 4:

Distribuição função erro complementar unitária

Miguel Peña Ramírez, Gladys Choque Ulloa

Resumo: No trabalho se apresenta um novo modelo uniparamétrico útil para modelar dados com suporte no intervalo unitário. No modelo proposto, são analisadas algumas das suas propriedades matemáticas mais importantes tais como função de distribuição acumulada, função quantílica, função taxa de falha, momentos, momentos incompletos e função geradora de momentos. É apresentada a estimação do parâmetro do modelo pelos métodos de máxima verossimilhança e dos momentos. Também, um estudo de simulação de Monte Carlo para avaliar o desempenho do estimador do parâmetro em amostras finitas e realizado.

Mini Palestra 5:

Smoothing quantile regressions with time series data

Miguel Jandrey Natal, Eduardo de Oliveira Horta

Resumo: Quantile regression (QR) fits quantiles of the response variable and brings the concept of a quantile into the framework of general linear models. Although QR was first introduced more than 40 years ago, only recently it became practible for large data, due to computational advances. As the objective function that the standard QR estimator aims to minimize is not smooth, statistical inference is not straightforward. Fernandes, Guerre and Horta (2021, Journal of Business & Economic Statistics) propose to smooth its objective function, thus presenting an alternative estimator: the convolution-type kernel QR estimator. Based on this alternative approach for quantile regression modeling, this work aims to implement the convolution-type kernel QR estimator in a time series data context.