

-Selecione o idioma-



Associação Brasileira de Estatística
XIV EBEB - Encontro Brasileiro de Estatística
Bayesiana - Rio de Janeiro



(<http://redeabe.org.br>)

Scige



Apresentações em Pôster

Instruções para apresentadores:

- O pôster devem ser afixado entre 15h20 e 16h40 do respectivo dia da apresentação e retirado ao término da sessão pôster
- As dimensões máximas de cada pôster devem ser de 90 cm de largura e 105 cm de altura
- O pôster deve estar preferencialmente em inglês

Pôster 1 (Terça-feira)

1. Título: Bayesian Multivariate GARCH with Multiple Degrees of Freedom

Autores: Lassance, R.F.L.; Cerqueira, V.S.; Fonseca, T.C.O.

Abstract: Based on the BEKK model, presented in Engle and Kroner (1993), we propose a Bayesian process of estimation of the parameters, as well as some adaptations. This model has great merit, since it allows economists to verify the direct impact of a change in volatility in the mean vector. We include the possibility of using a fixed number of autoregressors and other covariates, considering 3 possible distributions for the data: multivariate normal, multivariate Student's t and a generalization – allowing for each univariate time series to have a different degree of freedom, while maintaining a covariance structure for the data vector. Since this model is proposed having applications in macroeconomics and finances in mind, having a different degree of freedom for each time series being analyzed is not only desired, it provides more flexibility while still using a well studied model in the literature. Through the use of simulations, we evaluate the consistency of the estimation and – following Fonseca, Ferreira and Migon (2008) – use a Jeffreys prior for the degree of freedom. Finally, we show an application using real data (exchange rates from multiple countries compared to the US Dollar (USD)), give final remarks and propose future works to further improve the model.

Keywords: Time series; Multivariate GARCH; Generalized Student t; Jeffreys prior;

2. Título: A BAYESIAN APPROACH FOR A MULTIVARIATE

DISTRIBUTION OF EXCESS

Autores: Andresson Almeida Azevedo; Valmaria Rocha da Silva Ferraz**Abstract:** The theory of extreme values has as main objective the study of the tails of probability distributions, in order to measure and quantify extreme events of maximum and minimum. In this sense, in some cases it is interesting to take into account the influence that other variables may have on the variable of interest, as in the environmental area the increase of water level of a river can influence the increase of another river nearby. In this sense, the work of Heffernan and Tawn, 2004, proposes a conditional approach to estimation of the multivariate model. Inspired by this work, we propose here a Bayesian model, considering a conditionally independent structure, where the parameters of a locality can be written as a linear function of the values of the other localities. KEY WORDS: Conditional, Bayesian, Extremes**Keywords:** *Conditional; Bayesian; Extremes;***3. Título:** A Bayesian INAR(1) model with adaptive overdispersion**Autores:** Helton Graziadei; Hedibert F. Lopes; Paulo C. Marques F.**Abstract:** Autoregressive models for discrete time series have been proposed during the past several decades, generally motivated by the integer-valued autoregressive (INAR) model. However, few Bayesian methods are available for these models. In this paper, we introduce a data augmentation scheme to compute the posterior distribution of the parameters for the INAR(1) model. We also assess the forecasting capabilities of the Geometric-INAR(1) and Poisson-INAR(1) models by a simulation study.**Keywords:** *INAR(1) model; Data augmentation; Prediction performance; Dawid-Sebastiani score;***4. Título:** A Bayesian Look On The Selection Of Disputes For Litigation**Autores:** Fernando Correa Julio Trecenti Milene Farhat**Abstract:** Priest & Klein, 1984 is one of the most influential legal publications of all time. The authors state that the distribution of lawsuits is not the same as the distribution of disputes. Under efficiency and symmetry assumptions, plaintiffs and defendants rationally choose when to prosecute, ensuring that only highly unpredictable disputes become lawsuits. Because of its empirical importance, several authors worked through the original example with many interests. Some authors worked on a mathematical formalization of the arguments (Lee & Klerman 2016, Waldfogel 1995, Shapiro 1995, Hylton & Lin 2011) and others expanded theoretical assumptions (Gelbach 2016, Bernardo et al. 2000, Shavell 1996, Friedman & Wittman 2006). This paper explores these discussions in a Bayesian flavor. We defined a natural way to consider other sources of variation to the selection process, made the assumptions explicit and provided empirical validation for our suggestions.**Keywords:** *priest and klein; selection model; jurimetrics; litigation; settlement;***5. Título:** A Bayesian Mixture Models to fit players performance in fantasy

games

Autores: Lucas da Cunha Godoy; Augusto Felix Marcolin; Rodrigo Citton Padilha dos Reis

Abstract: Fantasy games about sports are becoming popular with the technological development. Currently these games, created in the first half of the 20th century in the baseball context, have a boom with the popularization of the internet. In Brazil, the Cartola FC is the most famous game of this genre. Launched in 2005, nowadays has more than 5 million registered users. The famous product of Globo.com, has attracted many sponsors, turning it in a very profitable product. The game is based on the 1st division of Brasileirão, all users start with 100 unities of a fictional money called "cartoleta". The main idea is the users build a team picking players. Every player has a price, so the user has a limitation to set up a team. The price of these players vary with their performance on each game of the championship. The performance of each team is measured based on players scouts, i.e. kicks to the goal, defenses, right passes, and so on. Each scout add or subtract a quantity, the final score of a player team is the sum of this quantities. We collected the data of these scouts for all players for the whole championship. We intend to build a bayesian hierarchical model to adjust these data based on the model proposed in Mayring and Gonçalves (2017). The mixture model that will be considered have two components, one considering that the score of the players does not depend on their past and one considering that this score is associated with the last round that he played. We will build a Gibbs Sampler with some Metropolis-Hastings steps to sampling from the full conditionals distributions and, subsequently, make predictions.

Keywords: Fantasy Games; Cartola FC; Mixture Model;

6. Título: A Bayesian approach to model survival data with crossing survival curves

Autores: Fábio Nogueira Demarqui; Vinícius Diniz Mayrink

Abstract: Proportional hazard (PH), proportional odds (PO) and accelerated failure time (AFT) models have been widely used to model survival data in different fields of knowledge. Despite their popularity, such models are not suitable to handle survival data with crossing survival curves. Yang and Prentice (2005) proposed a semiparametric two-sample model that includes the PH and PO models as particular cases, and is suitable for survival data with crossing survival curves. Considering a general regression setting, in this work we present a novel Bayesian approach to fit the YP model by modeling the baseline survival distribution via the piecewise exponential (PE) model. As a result, the proposed model shares the flexibility of the semiparametric models and the tractability of parametric models. The usefulness of the proposed model is demonstrated through the analysis of a real data set involving patients diagnosed with gastric cancer.

Keywords: Survival analysis; Yang and Prentice model; short-term and long-term hazard ratios; piecewise exponential model;

7. Título: A Bayesian hypothesis test for separability of multivariate spatial covariance functions

Autores: Rafael Erbisti; Thaís Fonseca; Mariane Alves

Abstract: Spatial models have been increasingly applied in several areas, such as environmental science, climate science and agriculture.

Georeferenced data is usually available in space, time and possibly for

several processes. In particular, in order to model multivariate spatial data, it is necessary to specify a valid cross-covariance function, which defines the dependence between the components of a response vector for all locations in the spatial domain. It is well known that the separable covariance functions have computational advantages but are not capable of adequately expressing complex dependency structures. Nonseparable covariance functions are more realistic for modeling the dependence between processes, resulting in improved predictive performance. In this work we present Bayesian inference for a flexible nonseparable class of cross-covariance functions for multivariate spatially referenced data, for which the separable model is a special case. We propose a Bayesian hypothesis test to measure the degree of separability between space and components. We calculate a posterior probability which is a weighting ratio between separable and nonseparable structures. This probability works as a mixing measure for the separable and nonseparable models, allowing both structures to be considered in the forecast. From simulations, we note that using the mixing model we obtain better predictive results than if we consider only the separable model or only the nonseparable model. An illustration is made with weather data from Ceará, Brazil.

Keywords: *geostatistical modeling; nonseparability; Bayesian hypothesis test;*

8. Título: A Bayesian non-parametric approach for the ADCC-GJR-GARCH model via Gibbs sampling and the Hamiltonian Monte Carlo method

Autores: Rafael Paixão; Ricardo Ehlers

Abstract: In Econometrics, the family of GARCH models is used to describe the variance of time series through innovations. However, the main characteristic of financial returns is that they often have skewness and high kurtosis. Despite the existence of several models sharing such characteristic in the literature, those models are inappropriate for many real applications, mainly due to the difficulty to identify the proper innovation distribution. To fill this gap, this paper proposes a Bayesian non-parametric approach for multivariate GARCH models. This approach has been developed computationally based on the MCMC algorithm, employing Gibbs sampling and the Hamiltonian Monte Carlo method. We evaluate this approach using the ADCC-GJR-GARCH model over real data, which consist of daily indices of stock markets in Frankfurt, Paris and Tokyo. The results show that, compared to the approaches currently available in the literature, our approach is able to identify the proper innovation distribution in a wider range of scenarios.

Keywords: *GARCH models; Bayesian inference; Dirichlet process mixtures; MCMC;*

9. Título: A Bayesian regression model for the non-standardized t distribution with location, scale and degrees of freedom parameters

Autores: Margarita Marin; Edilberto Cepeda-Cuervo

Abstract: In this work we analyze situations where the variable of interest cannot be assumed to have a normal distribution, because of the presence of heavy tails. We assume that it has a non-standardized t distribution with unknown location, scale and degrees of freedom parameters, in order to study t regression models. The review of the Bayesian literature on non-standardized t models allows the identification of at least two large gaps. The first is the sparse development of joint modeling of location and scale

parameters from a univariate perspective. The second is the lack of methodological approaches to the model discrete degrees of freedom. We propose a Bayesian t regression model in which we jointly model the location and the scale parameters and estimate unknown degrees of freedom. We also propose a new transition kernel for the degrees of freedom and estimate it from discrete and continuous perspectives. Finally, we compare our proposal with different prior proposals in the literature and find that our approach performs best in terms of information criteria and acceptance of estimated values.

Keywords: Bayesian methods; t distribution; Metropolis Hastings;

10. Título: A Dynamic Approach to the Linear Degradation Model Under a Bayesian Perspective

Autores: Guilherme Augusto Veloso; Rosangela Helena Loschi

Abstract: In the analysis of real data, such as degradation tests, perfect information on the phenomenon of interest is seldom available. Even when an accurate deterministic model describing the system under study is available, there is always something that is not under our control, such as the effect of omitted variables, measurement errors, or imperfections. In degradation models, particularly in the linear model, a constant rate of degradation is attributed over time, which may not be a reasonable assumption in many practical situations. As the linear relationship is only a local approximation to the true dependence structure involving degradation and time, a model with variable parameters may be more appropriate, since the occurrence of some change in the structure of the process under investigation may justify instability in rates of degradation. In this context, this work proposes the construction of more flexible models that do not assume a regular pattern of degradation rates or a stability for the system in question, but which may include points of change or structural breaks throughout the process. As an application, the operating current database in laser emitters is studied and it was possible to improve the predictive capacity of new observations and to propose an alternative inference of the predictive distribution of the failure times of a future unit.

Keywords: Dynamic Models; Linear Degradation Model; Bayesian Inference;

11. Título: A ZERO INFLATED GENERALIZED EXTREME VALUE DISTRIBUTION

Autores: Alexandre Henrique Quadros Gramosa; Fernando Ferraz do Nascimento; Fidel Ernesto Castro Morales

Abstract: The generalized extreme value distribution (GEV) is known as the limiting result for the Author: Alexandre Henrique Quadros Gramosa (UFPI) modeling of maximum blocks of size n , which is used in the modeling of extreme events. However, it is possible for the data to present an excessive number of zeros when dealing with extreme data, making it difficult to analyze and estimate these events by using the usual GEV distribution. The Zero Inflated Distributions (ZID) is known in the literature for modeling data with inflated zeros, where the inflator parameter w is inserted. The present work aims to create a new distribution for extreme values inflated of zeros, here named as IGEV, that will be applied in data of monthly maximum precipitation, that can occur during months where there was no precipitation, being these computed as zero. An inference was made on the Bayesian paradigm, and the parameter estimation was made by numerical

approximations of the posterior distribution using MCMC. Time series were analyzed of some cities in the northeastern region of Brazil, some of them with predominance of non - rainy months. The results of these applications showed the need to use the IGEV distribution, obtaining more accurate results, and with better adjustment measures when compared to the standard distribution of extreme value analysis (GEV).

Keywords: *Extreme Value; Zero Inflated; Bayesian; MCMC; Precipitation;*

12. Título: A bayesian approach to semiparametric count data models applied to university dropout

Autores: Julian A. A. Collazos; Diana M. Galvis Soto; Lucio Rojas C.;

Abstract: In this work, we focus on semiparametric regression models for count data with excess of ones using penalized splines (P-splines). Usually the Poisson regression model is the first choice to try this kind of data, however in the presence of overdispersion of the counts, the negative binomial regression model is an alternative statistical technique, and the mixture poisson regression models can be used to model counts with heavy-tailed distributions. When there are nonlinear effects of continuous covariates related to the count mean response, we can consider to model these covariates in a nonparametric form using penalized splines, thus giving a semiparametric structure to the count models that consider covariates with linear and nonlinear effects. The motivation of this work is related to the data of university student desertion, which involve the analysis of count data that exhibit a substantially large proportion of ones, such as the last completed semester in which a student abandoned his academic studies, which is considered our response count variable. The estimation process is carried out using a bayesian approach where the MCMC chains have good properties using low-rank splines functions, and the computational implementation of this approach is performed in WinBUGS, which is a flexible statistical tool for the semiparametric estimation of the count models.

Keywords: *MCMC; Bayesian inference; Penalized splines; Semiparametric model;*

13. Título: A censored time series model for responses on the unit interval

Autores: F. Schumacher; V. Lachos; G. Ferreira

Abstract: In this paper, we propose an autoregressive model for time series in which the variable of interest lie on the unit interval being subjected to certain threshold values below or above which the measurements are not quantifiable. The model includes the independent beta regression as a special case. We provide a full Bayesian approach for the estimation of the model parameters using standard Markov Chain Monte Carlo (MCMC) methods. We discuss the construction of the proposed model and compare it with alternative models using simulated and real data sets.

Keywords: *Autoregressive processes; Bayesian inference; Beta distribution; Censored models;*

14. Título: A general class of semiparametric models for recurrent event data

Autores: Rumenick Pereira da Silva, Vinícius Diniz Mayrink, Fábio Nogueira Demarqui

Abstract: Em áreas como medicina, saúde pública, negócios, indústria,

confiabilidade, ciências sociais e seguros surgem muitas situações em que o interesse é estudar processos que geram eventos repetidamente ao longo do tempo. Esses tipos de situações são denominados processos de eventos recorrentes e os dados que eles fornecem são chamados de dados de eventos recorrentes. Nesse contexto, a modelagem proposta no presente trabalho é, fundamentalmente, um modelo de sobrevivência baseado em uma classe geral introduzida por Peña e Hollander (2004), que possui o Processo de Poisson e Processo de Renovação como casos particulares, sendo a função de risco de base (ou intensidade de base) construída sob uma ótica semiparamétrica usando o modelo exponencial por partes. O modelo proposto é flexível no sentido de não impor uma forma específica para a função de risco (ou intensidade), ter qualidades similares a dos modelos paramétricos, no que diz respeito a estimativa das funções de sobrevivência, risco (ou intensidade) e risco acumulado (ou intensidade acumulada). Além disso, incorpora o impacto do número acumulado de eventos sob a função risco (ou intensidade) individual ao longo do tempo. Inferência para o modelo proposto será realizada utilizando as abordagens Bayesiana e Clássica. A análise desenvolvida aqui apresenta resultados de um estudo simulado objetivando investigar o comportamento do modelo proposto diante de diferentes cenários e explora também dados reais de um estudo de câncer de bexiga conhecido na literatura.

Keywords: *Processo de Poisson; Processo de Renovação; Modelo exponencial por partes; Inferência Bayesiana; Inferência Clássica;*

15. **Título:** A zero-inflated multilevel beta regression model for students' performance in the Brazilian Mathematical Olympiads for Public Schools

Autores: Igor Fernandes Lopes da Silva; João Batista de Moraes Pereira; Alexandra Mello Schmidt

Abstract: We propose a zero inflated multilevel beta regression model to investigate the performance of students and schools in the Brazilian Mathematical Olympiads for Public Schools (OBMEP). The OBMEP is held annually since 2005. Typically, each edition involves about 46,000 schools and more than 19.2 million students from all over the country. Here we focus on the edition of 2013. The OBMEP is performed in two phases, involving students divided into three different educational levels. To ease the computational burden of estimating the unknowns of our models we choose to analyze a sample from this population by following stratified and cluster random sampling schemes. We have information available in the student and school's level. The multilevel structure of our model allows the mean and the precision of the beta distribution to borrow strength across schools and educational levels and a mixture component is able to account for zero scores. Inference procedure is performed under the Bayesian paradigm and uncertainty about the unknowns of the model are naturally obtained. The resultant posterior distribution does not have closed form and we make use of Markov chain Monte Carlo methods to obtain samples from the posterior distribution. In particular, the software JAGS is used to obtain samples from the posterior distribution of interest. Here we focus on the analysis of students belonging to the State of Minas Gerais.

Keywords: *beta models; multilevel models; zero-inflated models; educational data; bayesian inference;*

16. **Título:** ABC to Bernstein polynomials: A nonparametric estimation method of densities

Autores: Leandro Augusto Ferreira; Victor Fossaluza

Abstract: In recent years, many statistical inference problems have been solved by using Markov Chain Monte Carlo (MCMC) techniques. However, it is necessary to derive the analytical form for the likelihood function. Although the level of computing has increased steadily, there is a limitation caused by the difficulty or the misunderstanding of how computing the likelihood function. The Approximate Bayesian Computation (ABC) method dispenses the use of the likelihood function by simulating candidates for posterior distributions and using an algorithm to accept or reject the proposed candidates. This work presents an alternative nonparametric estimation method of smoothing empirical distributions with random Bernstein polynomials via ABC method. The Bernstein prior is obtained by rewriting the Bernstein polynomial in terms of k mixtures of beta densities and mixing weights. Three examples are used to illustrate the method proposed.

Keywords: *Approximate Bayesian Computation; Bernstein polynomials; Density estimation; Nonparametric estimation;*

17. **Título:** Accounting for intra subject variability in a zero augmented positive regression model and assessing the effect of the number of repeated measurements and sample size on its estimation

Autores: Mariana Rodrigues-Motta; Claudia Akemy Koda; Eliseu V. Junior

Abstract: A zero augmented mixed regression model of positive outcome with distribution in the class of exponential family is study in this work. Motivated by a data set where subjects have 20 repeated measures each, one study goal is to investigate model behavior by introducing intra subject random effects, thus being able to estimate intra cluster variability in zero augmented models for positive response. The other study goal is to investigate to the extent which the amount of repeated measures along with the relationship with sample size affect estimation of intra subject variation. We adopt a Bayesian hierarchical modeling approach to estimate the latent variables and parameters. We test four different prior distributions for the variance components, and executed a simulation study to assess the best one among them. With the motivating example we show how the proposed method can be applied to extract valuable information from repeated measures, helping to identify the optimum sample size of repeated measures in the context of zero augmented positive regression models.

Keywords: *random effects; usual intake; zero-augmented positive data;*

18. **Título:** Adaptative significance levels in proportion hypothesis testing

Autores: Alejandra Estefanía Patiño Hoyos; Victor Fossaluza

Abstract: The Full Bayesian Significance Test (FBST) for precise hypotheses was presented by Pereira and Stern [Entropy 1(4) (1999) 99-110] as a Bayesian alternative instead of the traditional significance test based on p-value. The FBST uses the evidence in favor of the null hypothesis (H_0) calculated as the complement of the posterior probability of the highest posterior density region, which is tangent to the set defined by H_0 . An important practical issue for the implementation of the FBST is the determination of how large the evidence must be in order to decide for its rejection. In the Classical significance tests, the most used measure for rejecting a hypothesis is p-value. It is known that p-value decreases as sample size increases, so by setting a single significance level, it usually leads H_0 rejection. In the FBST procedure, the evidence in favor of H_0 exhibits the same behavior as the p-value when the sample size increases.

This suggests that the cut-off point to define the rejection of H_0 in the FBST should be a sample size function. In this work, we focus on the case of two-sided proportion hypothesis testing and present a method to find a cut-off value for the evidence in the FBST, by minimizing the linear combination of the type I error probability and the expected type II error probability for a given sample size.

Keywords: *FBST; Significance levels; Type I and type II error probabilities; Proportion hypothesis testing;*

19. Título: Análise Bayesiana do prêmio de seguro em função da quantidade de roubos de automóveis no Brasil

Autores: Luiz Otávio de Oliveira Pala; Josiane Correia de Souza Carvalho; Manoel Vitor de Souza Veloso

Abstract: Os prêmios são valores pagos as seguradoras, para que sejam responsáveis pelos efeitos econômicos de algum sinistro inerente ao bem segurado. O valor definido nas apólices de seguro para tais prêmios variam anualmente, e tal variação é consequência de alguns fatores que estão diretamente ligados aos contratos de seguro. Dentre os seguros no ramo não-vida, o de automóveis é um dos mais comercializados no Brasil. Desta forma, o objetivo deste trabalho é analisar a variação do prêmio em função do roubo de veículos. Para tal objetivo, foi feita uma análise bayesiana para o ajuste de um modelo de regressão, utilizando o software MultiBUGS, considerando dados dos 150 principais modelos de veículos para os últimos três períodos (anos de 2008, 2009 e 2010) disponíveis no Brasil. O diagnóstico de Gelman e Rubin indicou que houve convergência das cadeias para os parâmetros estimados. Foram calculadas as estimativas a posteriori por ponto (média) e por intervalo (IC-HPD com 95%) para os parâmetros do modelo, as quais indicaram que, com o aumento em 1% no número de roubos o prêmio aumenta 0,56%, o que, atuarialmente, permite levantar a questão de solvência neste setor, sob a ótica do impacto nos prêmios dado o número de roubos. Para o período em questão, evidenciou que há, com o passar dos anos, um aumento nos prêmios, de forma que no ano de 2008 eram 12,37% menores quando comparados ao ano de 2010, e eram 8,5% menores em 2009, quando comparado com esse mesmo ano. Tais resultados conduzem a uma discussão sobre os métodos para o gerenciamento de risco que, quanto melhor ajustados, auxiliam no equilíbrio atuarial.

Keywords: *Inferência Bayesiana; Seguros; Prêmio;*

20. Título: BAYESIAN FACTOR STOCHASTIC FRONTIER MODEL APPLIED TO POSTGRADUATE PROGRAMS EVALUATION DATA

Autores: Mariana de Amorim Donin; Kelly Cristina M. Gonçalves; Helio S. Migon.

Abstract: The Coordination of Improvement of Higher Education Personnel (CAPES) plays a key role in the expansion and consolidation of strictu sensu postgraduate programs in Brazil. One of its competencies is to evaluate the postgraduate programs and set a grade for them. This work focuses on the 52 postgraduate programs of Mathematics, Probability and Statistics, evaluated by CAPES in the last four years. In this case, the grade is determined by indicators that take into account the number of professors and their production, the number of students graduate, among other variables. The aim of this work is to facilitate the achievement of this grade attributed to the programs using an extension of stochastic frontier models

for multiple products, under the Bayesian paradigm. First, the multiple indicators used by CAPES are reduced to a single aggregated output, using the factor analysis technique. The approach is based on the assumption of a linear model in which part of the variability of the random vector is explained by latent factors, and the other is attributed to a t-student error component. Then, new indicators are proposed, that measure, for example, concentration of production among professors and the number of professors with productivity grants. The aggregated output and the new indicators are related through a stochastic production function, in which a linear model with a term that measures the technical inefficiency of the programs, associates inputs to the production. When compared to the evaluation published by CAPES, the results show that the proposal is efficient.

Keywords: *Factor models; MCMC; Gini coefficient; production function Cobb-Douglas;*

21. Título: Bayesian Copula-GARCH Modelling the Dependence Structure: an Application to International Markets

Autores: Beatriz Rezzieri Marchezini; Lucas Pereira Lopes; Vicente Garibay Cancho

Abstract: The increase in global and financial flows since the 1990s has made interdependence studies among economies an extremely relevant issue for both investors and policy makers. Considering its relevance, the objective of this work is to understand the relationship of dependence between the Brazilian economy and four major world economies. Therefore, it is proposed to use the method of copulas with fixed parameters of the elliptic and archimedean families, to analyze the degree of dependence between the Brazilian economy and four major world powers: United States, Japan, Germany and England for the period 2010 to 2017, in order to identify those that are more dependent on the Brazilian economy. The implemented models were the asymmetric and symmetric GARCH model for the marginal and normal copula, t-student, Gumbel, Frank, Clayton and Joe for the bivariate distribution. In terms of inference we adopt the Bayesian perspective and computationally intensive methods based on Monte Carlo simulations via Markov Chain (MCMC). Some interesting results of the co-movement between markets are discussed.

Keywords: *Copula-GARCH models; Markov Chain Monte Carlo; Dependence, Stock Markets;*

22. Título: Bayesian Option Pricing using Copula-GARCH model

Autores: Lucas Pereira Lopes; Vicente Garibay Cancho

Abstract: An option is a financial derivative that gives the right, but not the obligation, of the investor to exercise the purchase or sale of an asset at a given time and for a certain price. There are several consolidated pricing models in the financial literature, however, they impose some restrictions, such as constant volatility and normal multivariate distribution for the underlying assets. Based on modern finance theory, such assumptions are not valid. Therefore, the objective of this work was to develop a financial options pricing model via the GARCH-copula model, where the underlying assets of the options will be modeled via GARCH and the necessary dependency structure for the pricing will be modeled by the copulas functions. In terms of inference we adopt the Bayesian perspective and computationally intensive methods based on Monte Carlo simulations via Markov Chain (MCMC). The method will be applied considering symmetric

and asymmetric GARCH, archimedean and elliptic copulas, using Brazilian stock data. The advantage of this method is that the stock volatilities are dynamically captured and copula provides all information about the joint distribution of the underlying assets of the option.

Keywords: *Copula-GARCH models; Markov Chain Monte Carlo; Option Pricing;*

23. Título: Bayesian Restoration of Audio Signals Degraded by Low-Frequency Decaying Pulses

Autores: Hugo Tremonte de Carvalho; Flávio Rainho Ávila; Ralph dos Santos Silva; Luiz Wagner Pereira Biscainho

Abstract: The study of techniques to perform audio restoration can be motivated in several ways: for example, a collection of old recordings that are physically damaged may contain important information about the musical memory of some country or region; or if some audio of a crime scene in a noisy environment is available, it will be necessary to remove the degradation in order to find out what happened at that moment.

Unfortunately there is no available algorithm that restores a recording possessing any kind of degradation, and we must tackle the possible present defects separately. In this work we consider the situation where the recording is corrupted with long pulses, that are the response of the needle-arm set of a playback device when reproducing a deeply scratched or even broken disk or cylinder recording. We propose here a parametric and non-parametric model for this degradation, the latter being described as a Gaussian Process. Following the current literature, we impose an AR model for the original signal we wish to recover, and using adequate prior distributions for the model of the pulse, we perform a bayesian estimation of the degradation, and therefore, of the original signal, via the Gibbs Sampler. A missing point of the previously proposed techniques to solve this problem was an efficient initialization procedure: since this degradation does not corrupt the whole signal, it is important to automatically locate where the defect is present. We also propose an efficient method for locating the defect, based on a time-frequency analysis.

Keywords: *audio restoration; time-frequency analysis; statistical signal processing; bayesian statistics; gaussian process;*

24. Título: Bayesian Statistical Learning Applied to Text Mining

Autores: Lobo Vianna, B.; Damiani, A.P.; Trecenti, J.; Fossaluza, V.

Abstract: Statistical learning is one of the most trendy terms today. With the growing availability of online information, text mining has gained a lot of strength in analyzing data coming from the web. Most statistical learning methods can be formalized under the context of Bayesian inference. In this study, we will analyze a text which is a sequence of collective work agreements, in order to predict the existence of health assistance. These Hypertext Markup Language (HTML) documents have a predefined structure, with a title and a description of the different types of benefits (such as food, health and retirement). After separating each agreement into subgroups and concatenating those titles into a single string, we will build a predictive model using convolutional neural networks in one dimension under the Bayesian approach to statistical learning, in order to associate each agreement with the existence indicator of health assistance (1 when the contract covers health aid, 0 when it doesn't). This method was chosen for being flexible in its hypothesis of structure and observation. It also has

shown a great power to find relations in the words and text sequences without human intervention, and there is no need to explicitly build a keyword dictionary. In this paper, we will present a solution to the problem under the Bayesian approach to statistical learning.

Keywords: *Bayesian Inference; Statistical Learning; Neural Networks; Text Mining;*

25. Título: Bayesian cross-validation of geostatistical models

Autores: Viviana das Graças Ribeiro Lobo; Thaís Cristina Oliveira da Fonseca; Fernando Antônio da Silva Moura

Abstract: The problem of validating or criticising models for georeferenced data is challenging, since the conclusions can vary significantly depending on the locations of the validation set. This work proposes the use of cross-validation techniques to assess the goodness of fit of spatial models in different regions of the spatial domain to account for uncertainty in the choice of the validation sets. An obvious problem with the basic cross-validation scheme is that it is based on selecting only a few out of sample locations to validate the model, possibly making the conclusions sensitive to which partition of the data into training and validation cases is utilized. A possible solution to this issue would be to consider all possible configurations of data divided into training and validation observations. From a Bayesian point of view, this could be computationally demanding, as estimation of parameters usually requires Monte Carlo Markov Chain methods. To deal with this problem, we propose the use of estimated discrepancy functions considering all configurations of data partition in a computationally efficient manner based on sampling importance resampling. In particular, we consider uncertainty in the locations by assigning a prior distribution to them. Furthermore, we propose a stratified cross-validation scheme to take into account spatial heterogeneity, reducing the total variance of estimated predictive discrepancy measures considered for model assessment. We illustrate the advantages of our proposal with simulated examples of homogeneous and inhomogeneous spatial processes to investigate the effects of our proposal in scenarios of preferential sampling designs. The methods are illustrated with an application to a rainfall dataset.

Keywords: *Spatial processes; Data partition; Model criticism; Discrepancy function; Importance sampling;*

26. Título: Bayesian finite mixture modeling based on scale mixtures of univariate and multivariate skew-normal distributions

Autores: Marcus Gerardus Lavagnole Nascimento; Carlos Antonio Abanto-Valle; Victor Hugo Lachos Dávila

Abstract: In this work, finite mixtures of scale mixtures of skew-normal (FM-SMSN) distributions are introduced to deal simultaneously with asymmetric behavior and heterogeneity present in some data sets. A Bayesian methodology based on the data augmentation principle is derived and an efficient Markov-chain Monte Carlo (MCMC) algorithm is developed. These procedures are discussed with emphasis on finite mixtures of skew-normal, skew-t and skew-slash distributions for both univariate as well as multivariate case. Univariate and bivariate data sets using FM-SMSN distributions are analyzed. According to the results FM-SMSN distributions support both data sets.

Keywords: *Bayesian inference; finite mixture; scale mixture of normal distributions; Markov chain Monte Carlo.;*

27. Título: Bayesian modelling and allocation of insurance risks**Autores:** Rodrigo S. Targino; Gareth W. Peters; Mario V. Wuthrich

Abstract: In this talk I will present a fully Bayesian model for actuarial claims reserving consistent with the guidelines provided by the Swiss Solvency Test, the Swiss regulatory directive. This model is, then, used to compute the company's overall actuarial reserve, which, in a second stage, must be allocated to its individual lines of business. To compute the quantities involved in the process of allocation of capital to sub-units I will present a recently developed algorithm based on (pseudo-marginal) Sequential Monte Carlo methods.

Keywords: Sequential Monte Carlo (SMC); Solvency Capital Requirement (SCR); Swiss Solvency Test (SST); Capital allocation;

28. Título: Bayesian modelling of underreported counts through auxiliary variables**Autores:** Raffaele Argiento; Marcia D'Elia Branco; Rosangela Loschi; Fabrizio Ruggeri; Guilherme Oliveira

Abstract: In poor and more socially deprived areas, economic and social data are typically underreported. As a consequence, quantities of interest for, e.g., political, social and scientific purposes, such as income, rates of death and spread of diseases, tend to be underestimated. The great challenge, in those cases, is to build models able to provide reliable estimates for such quantities, despite the poor quality of data. The usual practice to overcome the problem is to assume that data are censored. However, under such an approach some recent papers have shown to be difficult correcting underreporting bias appropriately in the absence of quite informative prior information. Motivated by that, we introduce an alternative Bayesian model for mapping risks associated to data subjected to underreporting. The proposed model considers that count data in different areas are modeled using Poisson distributions with means depending on underreporting probabilities. Prior information is used to define a structure that guides the building of an appropriate distribution for the probability of underreporting in each area. The proposed structure give more flexibility to construct prior information for the underreporting process than those needed in censored models recently proposed to deal with such a problem. To illustrate the use of the proposed model, we map the early neonatal hospital mortality in the Minas Gerais, a Brazilian state which presents heterogeneous characteristics and a relevant socio-economical inequality and findings are compared with other approaches.

Keywords: underreporting; auxiliary variables; Bayesian modelling; infant mortality;

29. Título: Bayesian quantile regression in stochastic frontier models**Autores:** Angel Arroyo Hinostroza; Ralph S. Silva; Helio S. Migon

Abstract: We present a new class of models that use the power of Bayesian quantile regression to deal with stochastic frontier models. Compared with usual models that work with regression in the conditional mean, the proposal inherits the advantages of quantile regression, such as: robustness as it does not need to assume any distribution to the data, nor assume homoscedasticity, and also a better understanding since several quantiles provides more information about the response variable. In addition, the

proposal allows to obtain statistically significant differences between technical efficiencies in lower quantiles. In order to develop the inference we use the asymmetric Laplace distribution whose normal-exponential mixture representation allows to obtain known closed-form full conditional distributions. A Monte Carlo study is presented to empirically evaluate the methodology for Bayesian quantile regression. Two applications of the proposal are presented for cross-sectional and panel data sets. We compare the results obtained from our proposal with those obtained using conditional mean regression.

Keywords: *Asymmetric Laplace; Cobb-Douglas; Gibbs sampling; Technical efficiency;*

30. **Título:** Block Nearest Neighbor Gaussian process for big spatial data

Autores: Zaida Quiroz; Marcos Prates; Dipak Dey

Abstract: This work develops a valid spatial block-Nearest Neighbor Gaussian process (block-NNGP) for estimation and prediction of location-referenced spatial datasets. The key idea behind our approach is to subdivide the spatial domain into several blocks which are dependent under some constraints. As consequence, the cross-blocks should mainly capture the large-scale spatial variation, while each block should capture the small-scale dependence. Of course, the optimal blocking depends on the sampled spatial locations, and the number of blocks represents a trade-off between computational and statistical efficiency. The block-NNGP is included as prior in the hierarchical modeling framework and efficient Markov chain Monte Carlo (MCMC) algorithms exploit the sparsity of the block precision matrix, which can be computed by distributing the operations using parallel computing. The performance of the block-NNGP is illustrated using simulation studies and applications with massive data.

Keywords: *Bayesian Hierarchical models; block-NNGP; large datasets; MCMC; parallel computing;*

31. **Título:** Censored regression models in Brazil's officials statistics

Autores: Sofia E. P. de Lima; Gustavo H. M. A. Rocha; Alinne de C. Veiga

Abstract: In this work we consider censored regression models for Brazil's officials statistics. Inference is performed under the Bayesian paradigm and we use the MCMC techniques Gibbs Sampling, Metropolis-Hastings and ARMS. We assume symmetric (normal and t-Student) and asymmetric (skew-normal and skew-t) distributions to model the behavior of the errors and compare the results. We fit all four models to analyze the 2008-2009 Consumer Expenditure Survey (POF) data collected by the Brazilian Institute of Geography and Statistics (IBGE).

Keywords: *tobit-type models; consumer expenditure survey; bayesian inference;*

32. **Título:** Comparation between MCMC algorithms for Bayesian Inference in Dichotomous IRT Models

Autores: Gabriel O. Assunção; Flávio B. Gonçalves

Abstract: This study intends to compare two MCMC algorithms already existents on the literature for Bayesian Inference in the 3 parameters normal ogive model (3PNO). Both algorithms are Gibbs Sampling algorithms and have a presence of two latent variables, Béguin and Glas (2001) proposed an algorithm based in the idea of hierarchical for the model, Gonçalves, Dias

and Soares (2017) proposed an algorithm based in the idea of mixture for the model. It was done simulated studies for two scenarios, first: 2000 individuals and 40 items; second: 10000 individuals and 60 items; for both were calculated the time for an iteration and the effective sample size for several iterations and for a period. Results shows that Gonçalves, Dias and Soares algorithm had an effective sample size for several iterations and for a period greater than Béguin and Glas algorithm in both scenarios, but in the second scenario the results were quite similar. The time spent for an iteration was similar for both algorithms. The Béguin and Glas algorithm had more problems with the most difficult items in the simulated tests. In conclusion, with the adopted criteria the Gonçalves, Dias and Soares algorithm showed better results and less penalization for difficult items.

Keywords: *Item Response Theory; MCMC; Bayesian Inference; Algorithms;*

33. Título: Dynamic Calibration of Numerical Models for Forecasting Wind Speed in the State of Minas Gerais

Autores: Luiz Eduardo Silva Gomes; Thais Cristina Oliveira Fonseca; Kelly Cristina Mota Gonçalves

Abstract: Numerical predictions of climate variables are based on mathematical models that make deterministic predictions based on current atmospheric conditions. Based on the fluids dynamics theory, these models can be seen as a system of differential equations that have no analytical solution and numerical integration is used to simulate physical, dynamic and thermodynamic atmospherical processes depending on their current conditions, being possible to solve the system for any posterior time instant and therefore are widely used in the medium and long term climate forecast. Uncertainties in the representation of local and microscale phenomena from surface homogenization and imperfections in the formulation of the system of physical/dynamic equations of the mathematical model can be repaired using statistical post-processing techniques. These procedures improve the quality of numerical predictions, generating calibrated and accurate probabilistic forecasts of climate variables. Due to systematic errors from different sources of uncertainties which the numerical predictions for forecasting climatic variables are subject, the present work aims at the development of improved extensions of the main statistical postprocessing procedures that generate probabilistic forecasts and take the form of probability distributions (Gneiting and Raftery, 2005; Raftery et al., 2005) based on the vast class of Bayesian dynamical models (Harrison and West, 1997) and the technique of data augmentation (Tanner and Wong, 1987), producing a decrease in the forecasts errors of climate variables in meteorological fields. In particular, there is a greater interest in developing methods that support observable climatic variables from a threshold, censored or non-negative.

Keywords: *Dynamic bayesian models; Data augmentation; Spatio-temporal models; Calibration; Wind speed forecast;*

34. Título: Exact Bayesian inference in Level-set Cox processes

Autores: Bárbara da Costa Campos Dias; Flávio Bambirra Gonçalves.

Abstract: The most common point process in the literature is the Poisson process and an interesting extension is the Cox process, which is an inhomogeneous Poisson process where the intensity function (IF) evolves stochastically. This work proposes an exact inference methodology for spatial Cox processes when the intensity function is based on the Bayesian

Level-set model proposed by Dunlop et al. (2016). It is believed that in practice there may be cases where the intensity function is constant in pieces in space. Constructing an exact methodology that fits well with this type of problem generates more simplicity to the model and consequently one expects a greater precision in the results. A level set function, driven by a Gaussian random field, will flexibly determine regions of space that have constant intensities. Despite the intractability of the likelihood function and the infinite dimension of the problem, a methodology based on an MCMC (Markov chain Monte Carlo) algorithm with no discretization error involved, exploring recent stochastic simulation techniques such as pseudo-marginal Metropolis (Andrieu and Roberts, 2009), e Poisson estimator (Beskos et al., 2006).

Keywords: Cox processes; Bayesian inference; Markov chain Monte Carlo sampling;

35. **Título:** Dynamic Generalized Linear Models via Information Geometry

Autores: Raíra Marotta Bastos Vieira; Mariane Branco Alves; Hélio dos Santos Migon

Abstract: Dynamic generalized linear models are an extension for dynamic linear models (in the sense of considering non-Gaussian responses) and for generalized linear models, which consider responses in the exponential family, but presume fixed effects over time. One of the Bayesian inference methods for this class of models was proposed by West et al. (1985) and applies Linear Bayes to obtain estimates of the model, since the canonical parameter and predictive distributions have a closed analytic form, but the state parameters, which control the structural effects in the predictor, do not. The predictor is deterministically related to the canonical parameter. If a prior distribution is assigned to the states, it implies a prior on the canonical parameter, which must be compatible to the conjugate prior. Thus, there are two prior distributions for the natural parameter of the exponential family: one induced by the state vector and the conjugate prior in the exponential family. The solution suggested by West et al. was to equate the first and second moments of such prior distributions and the convenience of the closed analytic form for the posterior distribution of the canonical parameter and predictive distribution is preserved. We propose a new form of pooling prior distributions. This approach uses concepts of Information Geometry such as Projection Theorem and Bregman's Divergence. The idea is to project the prior induced by the vector of states in the geodesic of the conjugated prior distribution and then combine them. Once the priors are combined, the update structure follows the one proposed by W.H.M (1985).

Keywords: Dynamic linear models; Information Geometry; pooling distributions; Linear Bayes;

36. **Título:** Dynamic modeling of future longevity

Autores: Victhor S. Sartório; Thais C. O. Fonseca

Abstract: The fall of mortality rates throughout the last few decades has been a great achievement for society in general, being reflected through greater life expectancy. However, these changes in mortality rates also carry big economic consequences that are of particular interest for the insurance industry and government bodies. As such, being able to estimate future values of mortality and life expectancy adequately are of great importance. These changes can be modeled stochastically, and when doing so, the Lee-Carter model (1992) has been the most popular choice in recent years, because it manages to estimate easily interpretable parameters and latent

states. This work presents an extension of the Bayesian Lee-Carter model which includes an autoregressive parameter in the temporal evolution of latent parameters. This extension improves considerably the prediction of future mortalities and life expectancies, in particular in terms of precision and narrower credible intervals. Furthermore, an R package to perform Bayesian estimation of the parameters for the proposed model is developed. The package also includes methods for performing forecasts on both the mortality rates and life expectancies and implementations regarding the estimation of some extensions of the usual Bayesian Lee-Carter model. The usefulness of this proposal is illustrated with the analysis of mortality data from several countries. The results are compared with the ones obtained for the traditional models used in this context.

Keywords: *Mortality; Dinamic Linear Models; Forecasting;*

37. **Título:** Estimating a prevalence using information from an unobserved social network

Autores: Yasmin Cavaliere; Mariana Albi; Leo Bastos

Abstract: In order to promote public health policies, it is of utmost importance to get to know the size of hard-to-count populations such as sex workers, illicit drug users, men who have sex with other men, among others. These populations are often stigmatized and even criminalized. However, they have an important role in the dynamics of several transmissible diseases. The Network Scale-Up Method (NSUM) uses indirect information about the personal networks of respondents to make population size estimates. In this paper, we propose simulation studies to validate the NSUM method under fully controlled conditions and, then, discussed the real data biases and problems to data from a study of HIV/AIDS prevalence in Curitiba, Brazil. Network populations were simulated using a random network model. For each generated node, characteristics were randomly assigned; this way, the characteristics of each individual as well as their contacts will be known. Random samples were selected from these populations. Later, estimates for their size were calculated through Bayesian models, giving evidence that the Network Scale-up Method has proved to be efficient.

Keywords: *Network Scale-up; Hidden populations; Network sampling; Epidemiology; Public health;*

38. **Título:** Estimação da Intensidade de Tráfego de Filas Markovianas Multiservidoras via SIR

Autores: Sandy P. Alves; Frederico R. B. Cruz; Roberto C. Quinino

Abstract: Filas multiservidoras com chegadas Poisson e tempos de serviço exponenciais são o foco deste trabalho, mais especificamente, o problema de estimação bayesiana da sua intensidade de tráfego, definida como a razão entre a taxa de chegada e a taxa de serviço. Tais filas são importantes como modelos aproximados de diversas situações práticas, tais como em redes de computadores e de telecomunicações, sistemas de manufatura, de serviços e de saúde, entre outros problemas similares. É investigado aqui o uso método de amostragem/reamostragem de importância (ou SIR, do inglês sampling/importance resampling) e sua implementação em R. A determinação da intensidade de tráfego é ponto de partida para obtenção de outras medidas de desempenho importantes, tais como, o tamanho médio da fila, o número esperado de usuários no sistema, a probabilidade de sistema vazio, dentre outras. Resultados numéricos são

apresentados, para demonstrar a eficácia e a eficiência do SIR, diante de amostras de tamanho finito.

Keywords: *filas markovianas; multiservidor; inferência em filas; SIR;*

Patrocínio	Organização	Apoio
 CAPES Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (http://www.capes.gov.br/)	 IBGE Instituto Brasileiro de Geografia e Estatística  ISBrA ISBrA - Seção Brasileira da International Society for Bayesian Analysis (http://www.ime.usp.br/~isbra/)	 ISBA International Society for Bayesian Analysis (http://bayesian.org/)
		 impa Instituto de Matemática Pura e Aplicada  IME Associação Brasileira de Estatística (http://www.redeabe.org.br/site/)
		 IM-UFRJ Instituto de Matemática - UFRJ (http://www.im.ufrj.br)

© GwD - Sistemas Gerenciais e Analíticos 2019