

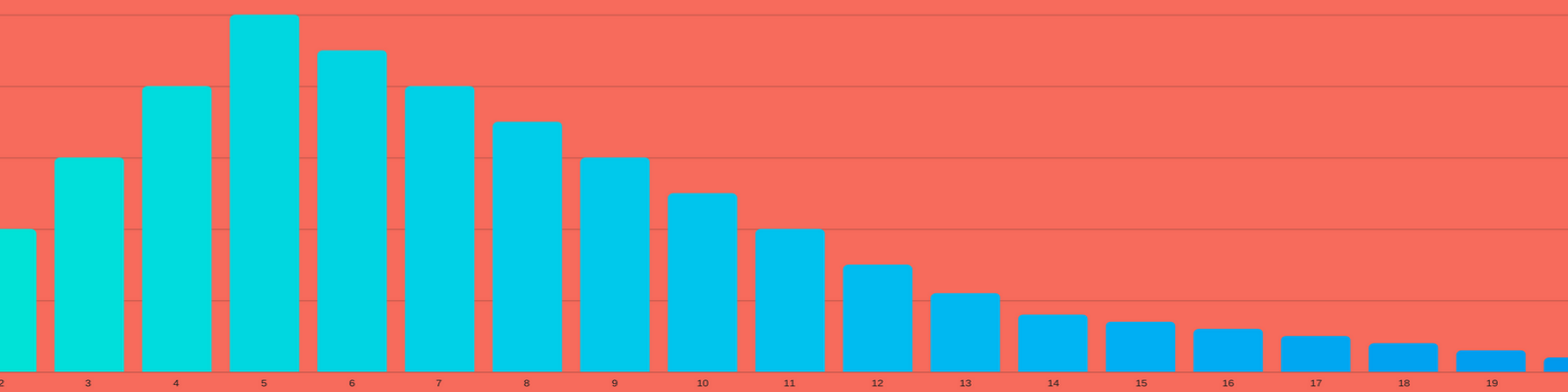
Meeting organized by the  
Joint Graduate Program in Statistics  
UFSCar/USP - PIPGEs

---

# VI Workshop on Probabilistic and Statistical Methods **PROGRAM**

---

February 5, 6, 7, 2018 - UFSCar, São Carlos/SP - Brazil



# 6th Workshop on Probabilistic and Statistical Methods

February 05–07, 2018

UFSCar, São Carlos, SP, Brazil

## PROGRAM

ICMC/USP - DEs/UFSCar

# About the 6WPSM

UFSCar, São Carlos, February 05–07, 2018

The Workshop on Probabilistic and Statistical Methods is an activity of the Joint Graduate Program in Statistics UFSCar/USP (PIPGEs), which brings together the research groups of probability and statistics working at ICMC-USP and UFSCar, in São Carlos/SP, Brazil.

The meeting intends to discuss new developments in statistics, probability and their applications. Activities include conferences and invited speaker sessions, contributed talks, poster sessions and a short course devoted to undergraduate/graduate students. The presentations of this new edition are related to probability and stochastic processes, statistical inference, regression models, survival analysis and related topics.

## Organizing Committee

Daiane Zuanetti - UFSCar  
Rafael Stern - UFSCar  
Ricardo Ehlers - USP (Chair)  
Pablo Rodriguez - USP  
Vera Tomazella - UFSCar (Chair)

## Scientific Committee

Adriano Polpo - UFSCar  
Carlos Alberto de Bragança Pereira - USP  
Enrico Colosimo - UFMG  
Francisco Louzada Neto - USP  
Josemar Rodrigues - USP

## Invited Speakers

Bruno Santos - UFBA  
Christian Galarza - UNICAMP  
Daniele Granzotto - UEM  
Edward George - University of Pennsylvania  
Fernando Quintana - Pontificia Universidad Católica de Chile  
Jean-Yves Dauxois - Université de Toulouse  
Jeremias Leão - UFAM  
Marcelo Bourguignon Pereira - UFRN  
Marina Paez - UFRJ  
Nancy Garcia - UNICAMP  
Paulo Justiniano Ribeiro Junior - UFPR (Short Course)  
Silvia Lopes de Paula Ferrari - USP

## Special Sessions

### *Latent variable modeling:*

Caio Lucidius N. Azevedo - UNICAMP  
Hedibert Freitas Lopes - Insper  
Jorge Luis Bazán Guzmán - USP

### *Probability:*

Carolina Bueno - USP  
Cristian Coletti - UFABC  
Mary Luz Rodiño - Universidad de Antioquia  
Miguel Abadi - USP

### *Survival Analysis:*

Francisco Louzada Neto - USP  
Jean-Yves Dauxois - Université de Toulouse  
Manoel dos Santos Neto - UFSCar and UFCG  
Vinicius Calsavara - A. C. Camargo Cancer Center

# 6th Workshop on Probabilistic and Statistical Methods

February 05–07, 2018

UFSCar, São Carlos, SP, Brazil

## SCHEDULE

ICMC/USP - DEs/UFSCar

## MONDAY 5 FEV

**8h00 - 8h30** : Registration

**8h30 - 9h00** : Opening

**9h00 - 10h00** : Edward George - University of Pennsylvania

**10h00 - 10h30** : Coffee Break

**10h30 - 11h30** : Bruno Santos - UFBA

**11h30 - 12h00** : Jeremias Leão - UFAM

**12h00 - 14h00** : Lunch

**14h00 - 15h00** : Silvia Lopes de Paula Ferrari - USP

**15h00 - 16h00** : Oral Communications

- 15h00 - 15h20: Clécio da Silva Ferreira - UFJF
- 15h20 - 15h40: Elizabeth Mie Hashimoto - UTFPR
- 15h40 - 16h00: Alex Sousa - UNICAMP

**16h00 - 17h00** : Coffee Break / Poster Session 1

**17h00 - 18h00** : Marcelo Bourguignon Pereira - UFRN

## TUESDAY 6 FEV

**8h00 - 10h00** : Short Course by Paulo Justiniano Ribeiro Junior - UFPR

**10h00 - 10h30** : Coffee Break

**10h30 - 11h30** : Nancy Garcia - UNICAMP

**11h30 - 12h00** : Daniele Granzotto - UEM

**12h00 - 14h00** : Lunch

**14h00 - 15h00** : Jean-Yves Dauxois - Université de Toulouse

**15h00 - 16h00** : Oral Communications

- 15h00 - 15h20: Lorena Cáceres Tomaya - UFSCar/USP
- 15h20 - 15h40: Amanda Morales Eudes D'Andrea - UFSCar/USP
- 15h40 - 16h00: Demerson André Polli - UnB and UFSCar/USP

**16h00 - 17h00** : Coffee Break / Poster Session 2

**17h00 - 18h00** : Marina Paez - UFRJ

## WEDNESDAY 7 FEV

**8h00 - 10h00** : Short Course by Paulo Justiniano Ribeiro Junior - UFPR

**10h00 - 10h30** : Coffee Break

**10h30 - 11h00** : Christian Galarza - UNICAMP

**11h00 - 12h00** : Fernando Quintana - Pontificia Universidad Católica de Chile

**12h00 - 14h30** : Lunch

**14h30 - 16h45** : Special Sessions

- *Latent Variable Modeling:*

- 14h30 - 15h15: Caio Lucidius N. Azevedo - UNICAMP
- 15h15 - 16h00: Hedibert Freitas Lopes - Insper
- 16h00 - 16h45: Jorge Bazán - USP

- *Probability*

- 14h30 - 15h00: Carolina Bueno - USP
- 15h00 - 15h30: Cristian Coletti - UFABC
- 15h30 - 16h00: Miguel Abadi - USP
- 16h00 - 16h30: Mary Luz Rodiño - Universidad de Antioquia

- *Survival Analysis:*

- 14h30 - 15h00: Francisco Louzada Neto - USP
- 15h00 - 15h30: Jean-Yves Dauxois - Université de Toulouse
- 15h30 - 16h00: Manoel dos Santos Neto - UFSCar and UFCG
- 16h00 - 16h30: Vinicius Calsavara - A. C. Camargo Cancer Center

**16h45 - 17h00** : Coffee Break



6th Workshop on Probabilistic and Statistical Methods

February 05–07, 2018

UFSCar, São Carlos, SP, Brazil

## ABSTRACTS

ICMC/USP - DEs/UFSCar

## Conferences

### **Bruno Santos - UFBA.**

On Bayesian quantile regression.

*Abstract: In this work we discuss the progress of Bayesian quantile regression models since their first proposal and we discuss the importance of all parameters involved in the inference process. Using a representation of the asymmetric Laplace distribution as a mixture of a normal and an exponential distribution, we discuss the relevance of the presence of a scale parameter to control for the variance in the model. Besides that we consider the posterior distribution of the latent variable present in the mixture representation to showcase outlying observations given the Bayesian quantile regression fits, where we compare the posterior distribution for each latent variable with the others. We illustrate these results with simulation studies and also data about Gini indexes in Brazilian states from years with census information.*

### **Edward George - University of Pennsylvania.**

Mortality Rate Estimation and Standardization for Public Reporting: Medicare's Hospital Compare.

*Abstract: Bayesian models are increasingly fit to large administrative data sets and then used to make individualized recommendations. In particular, Medicare's Hospital Compare webpage provides information to patients about specific hospital mortality rates for a heart attack or Acute Myocardial Infarction (AMI). Hospital Compare's current recommendations are based on a random-effects logit model with a random hospital indicator and patient risk factors. Except for the largest hospitals, these individual recommendations or predictions are not checkable against data, because data from smaller hospitals are too limited to provide a meaningful check. Before individualized Bayesian recommendations, people derived general advice from empirical studies of many hospitals; e.g., prefer hospitals of type 1 to type 2 because the risk is lower at type 1 hospitals. Here we calibrate these Bayesian recommendation systems by checking, out of sample, whether their predictions aggregate to give correct general advice derived from another sample. This process of calibrating individualized predictions against general empirical advice leads to substantial revisions in the Hospital Compare model for AMI mortality. In order to make appropriately calibrated predictions, our revised models incorporate information about hospital volume, nursing staff, medical residents, and the hospital's ability to perform cardiovascular procedures. For the ultimate purpose of comparisons, hospital mortality rates must be standardized to adjust for patient mix variation across hospitals. We find that indirect standardization, as currently used by Hospital Compare, fails to adequately control for differences in patient risk factors and systematically underestimates mortality rates*

at the low volume hospitals. To provide good control and correctly calibrated rates, we propose direct standardization instead. This is joint research with Veronika Rockova, Paul Rosenbaum, Ville Satopaa and Jeffrey Silber.

**Fernando Quintana - Pontificia U. Católica de Chile.**

Repulsion in Bayesian mixture models: how and why.

*Abstract: Bayesian mixture models have become very popular for applications such as density estimation and clustering. A typical assumption in this context is that component-specific parameters are modeled as independent quantities. An undesired consequence of this assumption in the context of clustering is the presence of very small or singleton clusters, which are very hard or impossible to interpret. This talk presents some approaches that have been recently proposed for solving this problem. The basic idea is to introduce the notion of repulsion of component-specific location parameters, that is, a probability model that encourages separation of these quantities. Specific constructions and applications to density estimation and clustering will be discussed.*

**Jean-Yves Dauxois - Université de Toulouse.**

Statistical inference in models of imperfect maintenance with geometric or arithmetic reduction of intensity.

*Abstract: In this talk I will introduce and study two new models of Imperfect Maintenance in Reliability: a model of Geometric Reduction of Intensity and another of Arithmetic Reduction of Intensity on the inter-arrival times of failures on a system subject to recurrent failures. Based on the observation of the recurrent failures of a single repairable system and assuming that a perfect repair is operated after  $N$  failures, we introduce estimators of the parameters (euclidean and functional) in this two semi-parametric models and we prove their asymptotic normality. Then a simulation study is carried out to learn the behavior of these estimators on samples of small or moderate size. We will end with applications on a real dataset.*

**Marcelo Bourguignon Pereira - UFRN.**

An INAR(1) process for modeling count time series with equidispersion, underdispersion and overdispersion.

*Abstract: We present a novel first-order non-negative integer-valued autoregressive model for stationary count data processes with Bernoulli-geometric marginals based on a new type of generalized thinning operator. It can be used for modeling time series of counts with equidispersion, underdispersion and overdispersion. The main properties of the model are derived, such as probability generating function, moments, transition probabilities and zero probability. The maximum likelihood method is used for estimating the model parameters. The proposed model is fitted to time series of*

*counts of iceberg orders and of cases of family violence illustrating its capabilities in challenging cases of overdispersed and equidispersed count data.*

*Joint work with: Christian H. Weiß- Department of Mathematics and Statistics, Helmut Schmidt University, Hamburg, Germany.*

### **Marina Paez - UFRJ.**

Hierarchical stochastic block model for community detection in multiplex networks.

*Abstract: Multiplex networks have become increasingly more prevalent in many fields, and emerged as a very powerful tool for modeling the complexity of real networks. There is a critical need for developing statistical models for inference in multiplex networks that can take into account potential dependency across different layers. There is in particular a demand for models for community detection. We fill this gap by proposing a novel and efficient Bayesian model for community detection in multiplex networks that take into account the dependency within and across different layers. A random partition prior is imposed for partitions across different layers of the multiplex network, under which a stochastic block model (SBM) is assumed. We also assume that the structure of the partitions is somewhat similar by imposing a hierarchical stochastic block model (HSBM) to the multiplex network. One of the key features of our model is that it allows the communities at different layers of the network to vary, which differs from many of existing methods for modeling multiplex networks, which assume that the communities are the same or fixed for all the layers.*

*Efficient MCMC algorithms were developed for sampling the posterior of ‘communities’, or the partition structure, as well as the link probabilities between nodes or communities. The developed algorithms were applied to extensive simulation studies and data examples which demonstrated the good performance of the models and algorithms.*

### **Nancy Garcia - UNICAMP.**

Modeling textile images with hidden Gibbs random fields.

*Abstract: When a new textile dyeing technology is developed, evaluating the quality of these techniques involves measuring the resulting color homogeneity using digital images. The presence of a texture caused by the fabric creates a sophisticated dependence structure in pixels coloring that is not accommodated by the available probabilistic models. Due to several factors, the random field that generates the texture can be seen as a mixture of colors and the mixture is given by a hidden Gibbs process with complex interactions. Joint work with Victor Freguglia Souza.*

### **Silvia Lopes de Paula Ferrari - USP.**

Box-Cox t random intercept model for estimating usual nutrient intake distributions.

*Abstract: The issue of estimating usual nutrient intake distributions and preva-*

lence of inadequate nutrient intakes is of interest in nutrition studies. Box-Cox transformations coupled with the normal distribution are usually employed for modeling nutrient intake data. When the data present highly asymmetric distribution or include outliers, this approach may lead to implausible estimates. Additionally, it does not allow interpretation of the parameters in terms of characteristics of the original data and requires back transformation of the transformed data to the original scale. We propose an alternative approach for estimating usual nutrient intake distributions and prevalence of inadequate nutrient intakes through a Box-Cox  $t$  model with random intercept. The proposed model is flexible enough for modeling highly asymmetric data even when outliers are present. Unlike the usual approach, the proposed model does not require a transformation of the data. A simulation study suggests that the Box-Cox  $t$  model with random intercept estimates the usual intake distribution satisfactorily, and that it should be preferable to the usual approach particularly in cases of highly asymmetric heavy-tailed data. In applications to data sets on intake of 19 micronutrients, the Box-Cox  $t$  models provided better fit than its competitors in most of the cases. Joint work with Giovana Fumes and José Eduardo Corrente.

## Mini-Conferences

### Christian Galarza - UNICAMP.

On moments of truncated multivariate Student- $t$  distribution: a recurrence approach.

*Abstract: Recurrence relations for integrals that involve the density of multivariate Student- $t$  distributions are developed. These recursions allow fast computation of the moments of folded and truncated multivariate normal and Student- $t$  distributions. Besides being numerically efficient, the proposed recursions also allow us to obtain explicit expressions of low order moments of folded and truncated multivariate Student- $t$  distributions. The newly methods are implemented in the new R package MoMtt. Joint work with: Victor Hugo Lachos Davila - UNICAMP, Tsung-I Lin - National Chung Hsing University (Taiwan) and Wan-Lun Wang - Feng Chia University (Taiwan).*

### Daniele Granzotto - UEM.

Gompertz-log-logistic distributions and minimum quadratic distance estimation.

*Abstract: The objective of this study is twofold: first to introduce two new parametric families of distributions with support on the positive real line that contain both the Gompertz and the log-logistic distributions as particular cases. The new families are flexible as they contain increasing, decreasing and bathtub or inverted bathtub hazard rate functions.*

*The construction of the new families is based on the cumulative hazard function*

and this leads to the second contribution of the research, a new method of parameter estimation based on minimizing a quadratic distance from the cumulative hazard of the model to the Nelson- Aalen estimator. We illustrate the use of the new estimation method on the proposed families and performed a simulation study to illustrate the finite sample properties of the derived estimators to show some of the robustness properties of the minimum quadratic distance estimators as compared to maximum likelihood.

Joint work with: Karim Anaya-Izquierdo - University of Bath and Francisco Louzada of USP.

### **Jeremias Leão, UFAM.**

Birnbaum-Saunders frailty regression models: diagnostics and application to medical data.

*Abstract: In survival models, some covariates affecting the lifetime could not be observed or measured. These covariates may correspond to environmental or genetic factors and be considered as a random effect related to a frailty of the individuals explaining their survival times. We propose a methodology based on a Birnbaum-Saunders frailty regression model, which can be applied to censored or uncensored data. Maximum-likelihood methods are used to estimate the model parameters and to derive local influence techniques. Diagnostic tools are important in regression to detect anomalies, as departures from error assumptions and presence of outliers and influential cases. Normal curvatures for local influence under different perturbations are computed and two types of residuals are introduced. Two examples with uncensored and censored real-world data illustrate the proposed methodology. Comparison with classical frailty models is carried out in these examples, which shows the superiority of the proposed model.*

## **Minicurso**

### **Paulo Justiniano Ribeiro Junior - UFPR.**

Métodos computacionais para inferência estatística.

*Abstract: Métodos numéricos e/ou computacionais são parte essencial do ferramental para tratamento de problemas de inferência estatística. Ambientes de prototipação e linguagens como o R ([www.r-project.org](http://www.r-project.org)) favorecem a implementação de métodos visando flexibilidade na especificação de modelos estatísticos. Este curso visa revisar a discutir diversos métodos aplicáveis em modelagem estatística. Serão abordados procedimentos numéricos como algoritmos para otimização/maximização, solução de sistemas, dentre outros, bem como métodos estocásticos para inferência via métodos computacionalmente intensivos. Exemplos de implementação são fornecidos em linguagem R.*

## Special Session: Latent Variable Modeling

**Caio Lucidius N. Azevedo - UNICAMP.**

Time series and multilevel modeling for longitudinal item response theory data.

*Abstract: Longitudinal Item Response Theory (IRT) data occurs when experimental units are submitted to measurement instruments (e.g., cognitive test, psychiatric questionnaires, biological essays among others) along different assessment conditions, as different time points. Very often, in this kind of study, we are interested in the so-called latent variables (or latent traits) and their behavior along these conditions, including the modeling of their inter-dependency structure. In this work we use some stationary and nonstationary time series and multilevel models to represent longitudinal IRT data. More specifically, we consider first order autoregressive (AR(1)), first order moving average (MA(1)), first order autoregressive-moving average (ARMA(1,1)) time series models as well as the Uniform and Hankel dependency structures, induced by appropriate multilevel models. These structures are studied under a time-homocedastic and time-heteroscedastic fashions. We developed a Bayesian inference framework, which includes parameter estimation, model fit assessment and model comparison, through MCMC algorithms. Simulation studies are conducted in order to measure the parameter recovery and model comparison tools. A real data analysis, concerning a longitudinal cognitive study for Mathematics achievement, conducted by the Federal Brazilian government, is performed. All computational implementations are made through the WinBUGS program, using the R2WinBUGS package, from R program. Joint work with Dalton F. Andrade.*

**Hedibert Freitas Lopes - Insper.**

Dynamic sparsity on dynamic regression models.

*Abstract: We consider variable selection and shrinkage for Gaussian Dynamic Linear Models (DLM) within a Bayesian framework. In particular, we propose a novel method that accommodates time-varying sparsity, based on an extension of spike-and-slab priors for dynamic models. This is done by assigning appropriate priors for the time-varying coefficients? variances, extending the previous work of Ishwaran and Rao (2005). Our approach is similar to the Normal Gamma Autoregressive (NGAR) process of Kalli and Griffin (2014), nevertheless, we assume a Markov switching structure for the process variances instead of a Gamma Autoregressive (GAR) process. Furthermore, we investigate different priors, including the common Inverted gamma prior for the process variances, and other mixture prior distributions such as Gamma priors for both the spike and the slab, which leads to a mixture of Normal-Gammas priors (Brown and Griffin, 2010) for the coefficients and also different distributions for the spike and the slab. In this sense, our prior can be view as a dynamic variable selection prior which induces either smoothness (through the slab) or shrinkage*

towards zero (through the spike) at each time point. The MCMC method used for posterior computation uses Markov latent variables that can assume binary regimes at each time point to generate the coefficients' variances. In that way, our model is a dynamic mixture model, thus, we could use the algorithm of Gerlach et al. (2000) to generate the latent processes without conditioning on the states. Finally, our approach is exemplified through simulated examples and a real data application. This is joint work with Paloma Uribe.

*Keywords:*

*Cholesky decomposition, dynamic models, Normal-Gamma prior, spike-and-slab priors, high-dimensional data, scale mixture of Normals.*

*References:*

*Brown and Griffin (2010) Inference with normal-gamma prior distributions in regression problems. Bayesian Analysis, 5(1), 171-188.*

*Gelarch, Carter and Kohn (2000) Efficient bayesian inference for dynamic mixture models. Journal of the American Statistical Association, 95(451), 819-828.*

*Ishwaran and Rao (2005) Spike and slab variable selection: frequentist and Bayesian strategies. Annals of Statistics, 33(2), 730-773.*

*Kalli and Griffin (2014) Time-varying sparsity in dynamic regression models. Journal of Econometrics, 178(2), 779-793.*

**Jorge Luis Bazán Guzmán - USP.**

Estimating the DINA Model Parameters Using the No-U-Turn Sampler.

*Abstract: The deterministic inputs, noisy "and" gate (DINA) model is a popular Cognitive Diagnosis Model (CDM) in psychology and psychometrics used to identify test takers' profiles with respect to a set of latent attributes or skills. In this work we propose an estimation method for the DINA model with the No-U-Turn Sampler (NUTS) algorithm, an extension to Hamiltonian Monte Carlo (HMC) method. We conduct a simulation study in order to evaluate the parameter recovery and efficiency of this new Markov chain Monte Carlo method and to compare it with two other Bayesian methods, the Metropolis Hastings and Gibbs sampling algorithms, and with a frequentist method, using the Expectation-Maximization algorithm. The results indicated that NUTS algorithm employed in the DINA model properly recovers all parameters and is more accurate than the other known methods used in the comparison. We apply this methodology in the mental health area in order to develop a new method of classification for respondents to the Beck Depression Inventory. The implementation of this method for the DINA model applied to other psychological tests has the potential to improve the medical diagnostic process.*

*Conjoint work with Marcelo Andrade da Silva, Eduardo Schneider Bueno de Oliveira and Alina A. von Davier.*



## Special Session: Probability

### Carolina Bueno - USP.

The fitness of the strongest individual in the subcritical GMS model.

*Abstract: We deduce explicitly the strongest individual fitness distribution on a variation for GMS model, proposed by Guiol, Machado and Schinazi. We point out to the fact that this distribution relies on the Gauss hypergeometric function and when  $p = 1/2$  on the hypergeometric function type I distribution. This is a joint work with Fábio Machado and Alejandro Roldán-Correa.*

### Cristian Coletti - UFABC.

Global survival of tree-like branching random walks.

*Abstract: The reproduction speed of a continuous-time branching random walk is proportional to a positive parameter  $\lambda$ . There is a threshold for  $\lambda$ , which is called  $\lambda_w$ , that separates almost sure global extinction from global survival. Only for some classes of branching random walks it is known that the global critical parameter  $\lambda_w$  is the inverse of a certain function of the reproduction rates, which we denote by  $K_w$ . We provide here new sufficient conditions which guarantee that the global critical parameter of tree-like branching random walks equals  $1/K_w$ . This result is part of a joint work with Bertacchi, D. and Zucca, F. (ALEA, v. 14, p. 381-402, 2017).*

### Miguel Abadi - USP.

Decay of Correlations for renewal, touching Morse and Fibonacci.

*Abstract: For a renewal process simulations a Morse code, we compute explicitly and exactly its decay of correlations. The technique is via solution of recursive linear sequence of second order. A general method is presented. The relation with the Perron Frobenius Operator appears at the end of the talk.*

### Mary Luz Rodiño - Universidad de Antioquia.

The connection between evolution algebras, random walks and graphs.

*Abstract: Evolution algebras are a new type of non-associative algebras inspired from biological phenomena. A special class of such algebras, called Markov evolution algebras, is strongly related to the theory of discrete time Markov chains. The winning of this relation is that many results coming from Probability Theory may be stated in the context of Abstract Algebra. In this work we explore the connection between evolution algebras, random walks and graphs. More precisely, we study the relationship between the evolution algebra induced by a random walk on a graph and the one determined by the same graph. This is a joint work with Paula Cadavid and Pablo Rodríguez.*

## Special Session: Survival Analysis:

**Francisco Louzada Neto - USP.**

Efficient Closed-Form MAP Estimator for Nakagami Fading Parameter.

*Abstract: The Nakagami distribution plays an important role in communication engineering problems, particularly to model fading of radio signals. Moreover, the Nakagami distribution has been used successfully in other fields such as medical imaging processing, hydrologic engineering, seismological analysis and traffic modeling of multimedia data. Important probability distributions can be obtained from the Nakagami distribution such as the Rayleigh and the half-normal distributions. Here, we present the reliability properties for this distribution and prove that its hazard rate (mean residual life) function presents increasing (decreasing) or bathtub (unimodal) shapes. A maximum a posteriori (MAP) estimator for the Nakagami fading parameter is proposed. The MAP estimator has a simple closed-form expression and can be rewritten as a bias corrected generalized moment estimator. Numerical results demonstrate that the MAP estimation scheme outperforms the existing estimation procedures and produces almost unbiased estimates for the fading parameter even for small sample size. The potentiality of our proposed methodology is illustrated in a real reliability data set. This is joint research with Pedro Luiz Ramos and Eduardo Ramos.*

**Jean-Yves Dauxois - Université de Toulouse.**

Some tests for Reliability models with different types of maintenance through a competing risks approach.

*Abstract: Let us consider an industrial system subjects to different causes of failure and different types of maintenance: a corrective maintenance is performed after a critical failure and a preventive maintenance can be performed in order to decrease the risk of critical failure. The recurrence of these types of maintenance has been often modeled in a competing risks framework. However rather few statistical inference has been carried out in these models. In this presentation we will introduce statistical tests in order to help the engineers to select the model which better fits their data. We prove the asymptotic normality of our test statistics and we carry out Monte Carlo simulations to learn how work our tests on finite size samples. Applications on a real dataset is also given.*

**Manoel dos Santos Neto - UFSCar and UFCG.**

A general class of Birnbaum-Saunders regression models for data containing zeros.

*Abstract: In this paper, we present a general class of zero-adjusted Birnbaum-Saunders regression model. This new model generalizes at least seven existing regression models in the literature. Estimation of model parameters, residual analysis and*

*influence diagnostic tools are also discussed. Finally, we apply it to two real-world case-studies to show its potential.*

**Vinicius Calsavara - A. C. Camargo Cancer Center.**

Defective Models for Cure Rate Modeling with Interval-Censored Data.

*Abstract: The regression models in survival analysis are most commonly applied for right-censored survival data. However, in some situations the time to the event is not exactly observed but it is known that the event occurred between two observed times. In practical problems, it is common to assume the moment of observation as the event occurrence time, ignoring the interval-censored mechanism. We present a cure rate defective model for interval-censored event-time data. Defective distribution is characterized by density function whose integration assumes values less than one when the domain of their parameters is different from the usual one. We consider the Gompertz and inverse Gaussian defective distributions which allow to model data containing cured elements. The parameter estimation is reached by maximum likelihood estimation procedure and Monte Carlo simulation studies are considered in order to evaluate the proposed models performance. The practice relevance of the models is illustrated through the ovarian cancer recurrence and oral lesion in children after liver transplantation datasets. Both studies were performed at A.C. Camargo Cancer Center, São Paulo, Brazil.*

## Oral Communications

**Clécio da Silva Ferreira - UFJF.**

Estimation and diagnostic analysis in skew-generalized-normal regression models.

*Abstract: The skew-generalized-normal distribution (hereafter SGN) [Arellano-Valle, RB, Gómez, HW, Quintana, FA. A new class of skew-normal distributions. Comm Statist Theory Methods 2004;33(7):1465-1480] provides a flexible family of skew-symmetric distributions containing the normal and skew-normal (hereafter SN) distributions as special cases. The SGN distribution can be represented as shape mixture of the SN distribution, in which the shape mixing distribution is normal. In this work we present some properties of the SGN distribution, as hierarchical representation, conditional and joint distributions, which will be used on the formulation of the EM algorithm. Also, we present the SGN linear regression model and the implementation of the EM algorithm to find the MLEs. After, we derive diagnostic analysis based on case-deletion measures, local influence and generalized leverage based on Zhu and Lee's approach. Different perturbation schemes are considered for local influence analysis. Finally, to illustrate the developed method, a real example is analysed. Joint work with Reinaldo B. Arellano-Valle.*

**Elizabeth Mie Hashimoto - UTFPR.**

Poisson Reparametrized Birnbaum Saunders for the Fatigue Data.

*Abstract: Mechanical components subjected to cyclic loading are susceptible to eventual failure. For this reason, researches are conducted to evaluate how much a component can be resistant. In this context, the objective of the study is to evaluate the effect of the force (kN) applied on the component and the presence of burr. Thus, the data were obtained of an experiment conducted at the Instituto Tecnológico de Aeronáutica and analysed using a Poisson reparametrized Birnbaum-Saunders time-promotion model. The results of the analysis showed that, at a significance level of 5%, there is no significant difference between the test specimens with and without burr. On the other hand, every 1kN we increase in force, we expect a reduction of approximately 9% in the lifetime. Now, considering a force greater than 12kN, less than 0.16% of the specimens are not susceptible to failure. Finally, we conclude that the Poisson reparameterized Birnbaum-Saunders time-promotion model is useful for analyzing the effect of explanatory variables in the mean lifetime without the need for logarithmic transformation in the response variable. In addition, through the model, it was possible to show that the force reduces the lifetime of the specimens and also decreases the proportion of specimens non-susceptible to the event of interest. Joint work with Janaína Fracaro de Souza Gonçalves and Milena Chanes e Souza*

**Alex Sousa - UNICAMP.**

Bayesian Wavelet Shrinkage with Beta Priors.

*Abstract: Bayesian shrinkage methods in the wavelet domain have been considerably studied in the recent decades. These methods allow the incorporation of prior information about the unknown signal such as smoothness, periodicity, sparseness, self-similarity and monotonicity. The goal of our work is to study this problem of bayesian wavelet shrinkage using the shifted beta distribution as prior distribution for the wavelet coefficients and two distributions associated with this family, the triangular and Bickel distributions. Performances of the Bayes rules associated to these priors are compared with classical shrinkage rules in the so called Donoho-Johnstone test functions. Joint work with Brani Vidakovic and Nancy L. Garcia.*

**Lorena Cáceres Tomaya - UFSCar/USP.**

Generalized fiducial inference for the Grubbs model.

*Abstract: Generalized fiducial inference for the precision of a measuring instrument without available replications on the observations is our main interest. In this work, we study two new estimation procedures for the precision parameters and product variability under the Grubbs model considering the two-instrument case. One method is based on a fiducial generalized pivotal quantity and the other one is built*

*on the method of the generalized fiducial distribution. The behavior of the point and interval estimators is assessed numerically through Monte Carlo simulation studies. Comparisons with two existing approaches are reported. Finally, the methodology is applied in the analysis of a data set from a methods comparison study. This is a joint work with Mário de Castro.*

**Amanda Morales Eudes D'Andrea - UFSCar/USP.**

Frailty general model for repairable systems under minimal repair with observed covariates.

*Abstract: In repairable systems, a key aspect to be considered is to predict the reliability of the systems under study. However, the standard methods in repairable system data analysis ignore the effect of accumulating event occurrences and the heterogeneity of the systems. Thus, this paper will be explored the frailty models, characterized by using a random effect, that is, a non-observable random variable that represents the information that could not or were not observed with a term that enables you to check the effect of accumulating event occurrences. In addition, observed covariates may affect the intensity function. The inferential method for estimation of the parameters will be displayed for models with minimal repair. Finally, an application to real data set was taken. Joint work with Vera Lucia Damasceno Tomazella.*

**Demerson André Polli - UnB and UFSCar/USP.**

Avaliação de produtos baseada na presença de  $m \geq 1$  características.

*Abstract: Um produto (ou um serviço) pode ser avaliado através da verificação da satisfação de  $m \geq 1$  características. Pelo menos dois cenários distintos de avaliação são possíveis: (1) os avaliadores atribuem, para cada item, um valor entre 0 e  $m$  correspondente à contagem de características consideradas satisfatórias, ou (2) os avaliadores atribuem, para cada item, um vetor de  $m$  variáveis dicotômicas correspondentes a cada característica. Os avaliadores pertencem a diferentes populações que consideram níveis distintos de importância para as características ou atribuem as respostas para cada característica ao acaso (sem julgar a satisfação da referida categoria). Este resumo mostra alguns modelos simples para tratar estes 2 cenários de avaliação citados. Joint work with Carlos Diniz.*

## Poster Session 1

**Alex de la Cruz Huayanay - USP/UFSCar.**

Asymmetric links and methods for imbalanced data in binary regression.

*Abstract: In the Binary Regression, the imbalanced data refers to the presence of values equal to zeros or ones in a proportion that is significantly greater than the corresponding values ones or zeros. In this work, we study the main methods developed to deal with unbalancing and we compare them to the use of asymmetric links. We specifically worked with some links proposed in Lemonte and Bazan (2017) and Bazan et al (2016). The results show that these methods do not adequately correct bias in the estimation of regression coefficients and that the models with power links and reciprocal power considered produce better results, for certain types of imbalanced data.*

*Additionally, we will present an application for imbalanced data, identifying the best model among the various proposed models.*

*The estimation of the parameters was done using a Bayesian approach, considering Hamiltonian Monte-Carlo method, utilizing No-U-Turn Sampler algorithm. The comparison of models was developed using different criteria for model comparison. Joint work with Jorge Bazán.*

**Andson Nunes da Silva - USP.**

Aplicação de novos modelos de regressão binomial para dados desbalanceados.

*Abstract: Modelos de regressão binomial fazem parte dos modelos lineares generalizados e análise de dados categorizados. Usualmente são apresentadas e usados modelos de regressão logística as quais usam funções de ligação para dados balanceados. Quando queremos ajustar um modelo regressão binomial para dados desbalanceados e usamos a regressão logística, a função de ligação pode estar mal especificada, de modo que pode levar a vícios grandes nas estimativas da resposta média como será mostrado. Assim, algumas propostas foram desenvolvidas para lidar com dados desbalanceados entre elas o uso de funções de ligações assimétricas. Neste trabalho iremos avaliar as funções de ligação propostas por Lemonte e Bazán (2017) para os modelos de regressão binária na presença de dados desbalanceados e mostraremos uma aplicação para determinar se um e-mail é um spam ou não em função de algumas covariáveis sob abordagem bayesiana.*

**Carlos Eduardo Hirth Pimentel - USP/UFSCar.**

The frog model on finite graphs.

*Abstract: Several problems are modeled by dynamics of particles that spread in a defined space. The frog model is an example, in which the particles are interpreted as frogs that jumping through sites in a given graph. In this work, we propose a frog*

model which the possible states of each frog located on a finite graph  $\mathcal{G}_n$  are empty, inactive and containing one or two frogs. Based on a probabilistic formalism, some approximations for the equations of evolution of the probabilities on  $\mathcal{G}_n$  graph are given and converge in a system of ordinary differential equations for any finite graph. Simulations are compared with ode system obtains by density dependent continuous stochastic processes methods for  $\mathcal{G}_n$  and are obtained results of the covering evolution of the possible states for complete and bipartite graphs. Joint work with Pablo Rodríguez.

### **Diego Carvalho do Nascimento - USP/UFSCar.**

Copula bivariate model: predictive analytics on estimation tinnitus psychoacoustic parameters.

*Abstract: This work aimed to analyze the conditional dependency of the Loudness perception and the magnitude of Minimal Masking Level (MML) among tinnitus frequency. The dataset has information from 626 subjects enrolled in this study were Iran-Iraq war veterans. From 2004 to 2011, 626 patients with blast-induced tinnitus and hearing loss (unilateral or bilateral) were carefully visited, and each patient was followed up for 2 years, totaling 1252 observed ears. They were referred to our clinic after their hearing loss and/or tinnitus were diagnosed by audiological tests, ENT visits, medical tests and military documents of their war participation at the time of injury. Results considered marginal normal distributions associated with studied Tinnitus psychoacoustic measurement, thereby dependence structure through normal copula were fitted in this data. In the application of the copula bivariate, the outcomes are in line with our hypothesis that positive emotional induction can reduce the loudness perception in tinnitus patients. Further investigation is needed to find out its effect on bothersome and the ability to ignore tinnitus. Joint work with Francisco Louzada, Iman G. Toostani, Zahar S. Vaziri.*

### **Eduardo Schneider Bueno de Oliveira - UFSCar/USP.**

New Gompertz based distributions to skewed bounded responses.

*Abstract: In many applications in different fields of study, the observed variable is bounded. When it comes to modeling the distribution of bounded variables, there are some models yet proposed in literature and very much known, such as the Beta distribution and the Simplex distribution.*

*Motivated by data with skewness we propose two new distributions to bounded responses, with values between 0 and 1. These distributions are based on Gompertz distribution (Lenart, 2014), with proper transformation, which we call GLim1 (Gompertz Limited 1) and GLim2 (Gompertz Limited 2).*

*Besides being able to capture skewness, each one of the proposed models is closed at one end, that is, variables with values equal to 1, by example, can be modeled using one of our proposed distribution, instead of using mixture models such as Beta inflated*

(Ospina and Ferrari, 2010), the same is true to variables with values equal to 0.

In this poster we will show the densities and the cumulative distributions of both and some important characteristics, such as the mode, skewness and kurtosis measures, as well as some important moments and the behavior of GLim1 and GLim2 with the parameters variation. We also present an useful reparametrization to the proposed distributions through quantiles, which makes possible to use them to adjust quantile regression models. Joint work with Mário de Castro Andrade Filho, Cristian Luis Bayes and Jorge Luis Bazán.

### **Fabiana Arca Cruz Tortorelli - USP/UFSCar.**

Survival analysis of the industrial engineering students dropout from UNESP.

*Abstract:* High education systems face dropouts and late graduations and then they have to deal with economics and socials issues. Our goal was to study students dropout of the Industrial Engineering course from the Câmpus Experimentais (CE) of UNESP by calculating the students proportion that will graduate using characteristics associated with the abandon students. The methodology used to find the proportion was long term survival analysis. Particularly we describe the risk function and survival function of student permanence time at the university. Therefore, we was able to identify the mean of students permanence time and the profile of the students that abandoned the course. Joint work with Juliana Cobre and Sandra Cristina de Oliveira.

### **José Fausto de Moraes - UFU.**

Preliminary rating scale of issued rapport (PRIR-10).

*Abstract: Introduction:* Rapport is understood as a high-level connection between people, and Issued Rapport is a kind of signal (voluntary or involuntary) that people send to one another in the process of interaction. In psychology, the good patient-therapist relationship requires rapport, and in this sense, instruments for measuring intensity of an established connection are very useful. *Objective:* To propose a scale (PRIR-10) to evaluate the intensity of issued rapport by university students, as well as discuss the reliability and the validity of such a scale. *Methods:* In the construction of the scale, 10 behaviors/attitudes that characterize a charismatic and popular individual were identified in the literature. The scale obtained was applied to a sample of 388 students with known AQ10 scores. The sample had 46 (11.9%) students with AQ10 above 6 (Group A) and 74 (19.1%) students with AQ10 below 3 (Group B). The KR-20 coefficient was used in evaluation of internal consistence of scale. The Mann-Whitney U test and the Spearman's Rank Correlation Coefficient were used in the study of discriminant validity. A Factor Analysis was used to evaluate the dimensional structure of the scale. *Results:* The scale was composed by ten items: a general item related to Rapport (the item D41) and nine items divided into three conceptual domains (subscales), here named by: Mutual Attentiveness, Commonality and Coordination (or Mirroring). With respect to the internal consistency of the scale,



the ten items presented KR-20 of 0.68(95%CI: 0.63 to 0.73). The factor analysis applied to the sample exhibited a KMO measure of sampling adequacy of 0.693, and the four extracted factors explained 63% of the total variance. The items D21, D22 and D23 of the scale formed the subscale of Mutual Attentiveness, which explained 19.4% of the total variance and presented KR-20 of 0.67(95%CI: 0.61 to 0.72). The items D31, D32 and D33 formed the subscale of Commonality, which explained 17.5% of the total variance and presented KR-20 of 0.68(95%CI: 0.62 to 0.73). The items D11, D12 and D13 formed the subscale of Mirroring, which explained 14.0% of the total variance and presented KR-20 of 0.44(95%CI: 0.33 to 0.53). With respect to the discrimination capacity of the scale, the average score of the PRIR-10 in the Group A was significantly lower than in the Group B ( $6.11 \pm 2.08$  vs.  $8.49 \pm 1.64$ ,  $p < 0.001$ ). A similar result were obtained with the average score of the subscale Commonality ( $1.46 \pm 1.15$  vs.  $2.51 \pm .86$ ,  $p < 0.001$ ) and the subscale Mirroring ( $1.85 \pm .96$  vs.  $2.17 \pm .67$ ,  $p < 0.001$ ). For the 388 students involved in the study, we identifies a significant negative correlation between the AQ10 and PRIR-10 scores ( $r = -0.39$ ,  $p < 0.001$ ), and between the AQ10 and D41 ( $r = -0.15$ ,  $p = 0.002$ ). The ROC curve considering the PRIR-10 scores as test variable to identify the Group B (without trouble to establish connections) produced a cut-off score of 7.5 with 79.7% of sensitivity, 78.3% of specificity and AUROC of 0.812 ( $p < 0.001$ ). Conclusion: The KR-20 calculated showed a satisfactory internal consistency ( $> 0.60$ ) for the items of scale in this study, and these items exhibited a high ability to discriminate individuals with and without trouble to establish connections. The results was also observed in some of the subscales. The factor analysis revealed that the scale has a three-dimensional structure. The findings suggest that the proposed scale has enough reliability and validity to be used to evaluate of issued rapport by University students. Joint work with Tiago Bernardes Kerr and Victor Lawrence Bernardes Santana.

### **Juliana Marambaia Maia - USP/UFSCar.**

Modeling of latent traits on nurses burnout.

*Abstract: The motivation for this study comes from a dataset related to burnout syndrome in nurses, from the RN4CAST project (Sermeus et al., 2011). This project involved 12 countries in Europe, and a large number of hospitals, nursing units, nurses and patients. Burnout is a psychological syndrome of emotional exhaustion, depersonalization and personal accomplishment. Item response models are usually applied to items with binary responses, and the probability that the  $i$ th examinee answers correctly the item  $j$ ,  $P(Y_{ij} = 1)$ , is modeled as a function of both examinee's ability and items' characteristics. A generalization of dichotomous item response model is the polytomous item response model, in which the items now assume values in more than two categories ( $m_j$  categories). A polytomous item response model that is well known in the literature is the graded response model, proposed by Samejima (1969, 1972). Now, let  $Y_{ij}$  be the response of the  $i$ -th nurse in the  $j$ -th item. Each item  $j$ ,  $j = 1, \dots, 22$ , has  $m_j = 7$  categories and  $i = 1, \dots, n$  is the number of nurses.*

Then,  $Y_{ij}|\theta_i, \xi_j \sim \text{Categorical}(P_{ij1}, \dots, P_{ij7})$ , where  $\theta_i$  is the latent trait of the  $i$ -th nurse and  $\xi_j = (a_j, b_{j1}^*, \dots, b_{j(m-1)}^*)^T$ , is the parameter vector related to the  $j$ -th item, in which  $b_{jk}^* = a_j b_{jk}$ ,  $\forall k \in 1, 2, \dots, m_j - 1$ . Now the parameter  $b_{jk}$  represents the difficulty of moving from the category  $k - 1$  to  $k$  of the item  $j$ . According to da Silva et al. (2017), in a graded response model, the probability of a nurse scoring a specific category is modeled from cumulative probabilities,  $P_{ijk}^+$ , which are the probability of the nurse  $i$  select a specific category  $k$  or other lower category in item  $j$ . So,  $P_{ijk}^+ = P(Y_{ij} \leq k|\theta_i, \xi_j) = L(\eta_{ijk})$ , where  $L(\cdot)$  represent the link function and commonly uses the logistic cumulative distribution as the link function. In this study, we use the graded response model to model the latent trait for each burnout dimension. Preliminary results will be presented. Our proposal is to use Bayesian softwares, for example, WinBugs and/or JAGS to obtain the final analysis. Joint work with Cibele M. Russo, Jorge L. Bazan, Emmanuel Lesaffre.

### **Juliana Scudilio Rodrigues - UFSCar/USP.**

Survival of patients with breast cancer submitted neoadjuvant chemotherapy treatment.

*Abstract: Breast cancer is one of the leading causes of death in the world and the second leading cause of death among women in Brazil. Identifying risk factors and analyzing the effect of treatment on patients' lives are extremely important for a better understanding of the disease. The objective of this paper are to analyze survival and estimate a proportion of cured women with breast cancer and submitted neoadjuvant chemotherapy treatment. We analyze 78 women with breast cancer and submitted neoadjuvant chemotherapy treatment in A.C Camargo Cancer Center. We used the cure rate model and classical inference for estimated the parameter of the distribution.*

*The results obtained: 25 women have died of breast cancer; the cure rate was 62%; women with tumor-infiltrating lymphocytes,  $TIL \leq 10\%$ , had a longer survival than those women with  $TIL > 10\%$ , with a cure rate 70% vs 51%; women with nearby lymph nodes do not contain cancer (N0) had a cured proportion of 83%, higher survival than those women had number of lymph nodes involved N1, N2 and N3, a cure rate of 55%; women had not any evidence of the primary tumor (T0) had a cure rate of 77% and the women had evidence of the primary tumor (T1-T4) a cure rate of the 52%.*

*Joint work with Vinicius F. Calsavara and Vera Tomazella.*

### **Lucas Pereira Lopes - USP/UFSCar.**

Pricing Rainbow Options Using Garch-Copula.

*Abstract: A financial option is a legal contract related to a certain asset. Thus, options are a kind of financial derivative that gives the right to the buyer, but not the obligation, to buy or sell an asset by a certain date at a certain price (BODIE, 2013). Many models consolidated in the literature make unrealistic assumptions to price financial options, where two of the main assumptions are: a) the variance of the*

*log-returns of the underlying assets is considered constant and; b) returns on financial assets follow a normal multivariate joint distribution. However, given the volatility of the financial market, the assumption that the variance of log-returns is constant becomes a dangerous premise in the pricing process (RACHEV; RUSCHENDORF, 1995). Moreover, assuming a normal multivariate distribution for returns requires that the correlation between the underlying assets be treated as linear, which is not true because of the stylized facts in finance (DOBRIC; SCHMID, 2005). Thus, this work has as main objective to study the process of pricing of rainbow options considering that the volatility of the assets can be modeled by a GARCH process and that the joint distribution of assets be modeled by families of copulas. To proceed with the methodology proposed in this work, we are going to use the series of two Brazilian stocks to calculate the fair price of the rainbow option. Two different sectors were chosen for the analysis, and this choice was based on the rationale of investor portfolio diversification. Therefore, we are going to analyze stocks of Itau bank (financial institution) and Ambev (beverage company). As preliminary results, it is noticed that the choice of the copula and the distributions of the errors of the marginal series lead to different results in the pricing of the option (which is evident with the graphical results), which makes the task of choosing these two elements a very important area for studies. This result evidences that, adopting a methodology to choose the distribution of residues, it is crucial to develop and perform tests to find out which copula best fits the data studied for the optimization of the financial options pricing process. Joint work with Vicente Garibay Cancho and Francisco Louzada Neto.*

## Poster Session 2

**Davi Keglevich Neiva - UFSCar.**

Repairable system modeling study under supposition of minimum repair.

*Abstract: At the last decades the efforts on studies focused on identify and predict the reliability of repairable systems increased considerably. Most of the systems are designed in a way that the failure of a single component implies that the whole system will stop to work (series association), which increases the importance of studies that indicate where the maintenance must be more effective to ensure the proper functioning of the system. Many equipments used everyday are repairable, such as vehicles, smartphones, televisions and so on. A failure in an equipment that does not enable it to perform its functions can lead to a countless damage for a company, such as the inconvenience for customers, loss of products or even accident at work. On the other hand, an intensive maintenance intervention increases a lot the expenses and requires more service by labours, which can not worth for the companies. For this reason, one of the approaches of the repairable system modeling is the minimum repair, which consider that the intervention repairs the system, but maintain it working as*

before the failure, in other words it only keeps the system in operational conditions. Therefore, the main aim is to optimize the maintenance actions, identifying and repairing only the critical parts of the system, ensuring the functioning at least expense. As usually the minimum repair is the action that presents the lowest cost there are many studies regarding this approach. There are many theoretical material that aim to describe the failure behaviour according to the equipment history considering that the repair on it is the minimum. In this study, the model obtained is applied to a case study of trucks used to transport the products of a mining company, revealing that some information from the model output can be very relevant for companies to better plan its maintenance actions, commit resources more efficiently and reduce expenses. In this work inferential method for estimation of the parameters will be displayed for models with minimal repair and a application to real data set will be taken. Joint work with Vera Tomazella.

### **Karlla D. Chagas - FCT/UNESP.**

Inferência Bayesiana para testes acelerados “step-stress” simples com dados de falha sob censura progressiva tipo II e distribuição Gama.

*Abstract: O avanço tecnológico vem crescendo progressivamente, de modo que o mercado industrial vem se mostrando cada vez mais interessado em produzir produtos de maior qualidade, visando em um segundo momento avaliar de forma adequada a confiabilidade destes produtos. Ao querer realizar inferências sobre a confiabilidade de um produto, é necessário realizar alguns tipos de testes. Intuitivamente podemos pensar que quanto maior é a qualidade de um produto produzido, mais difícil é avaliar a sua confiabilidade. Isto porque se realizarmos os testes em condições normais de funcionamento do produto, ou seja, elaborarmos um experimento colocando n produtos em testes e esperarmos até que eles falhem, muito tempo de teste pode ser exigido. Como base nisso, uma alternativa é o uso dos métodos de testes acelerados, em que este tem como objetivo acelerar a ocorrência da falha de um produto, colocando o mesmo em condições de estresse mais elevadas do que o normal. Quando estamos interessados em realizar um experimento em que será utilizada as técnicas de testes acelerados, uma das coisas mais importantes que devemos nos preocupar, é definir o tipo de carga de estresse que será aplicada nos itens. Como descrito por Nelson [1], existem diferentes cargas de estresse, sendo que neste trabalho foi adotado a carga de estresse do tipo step-stress, na qual um conjunto de itens é submetido a um nível de estresse alto por um período de tempo pré-determinado. Caso esses itens não falhem até esse período, o nível de estresse é elevado para um novo nível, e este processo se repete até o final do estudo. Após o término de um experimento de testes acelerados, os tempos de falha obtidos podem ser modelados por diferentes distribuições de probabilidades, em que neste trabalho consideramos uma distribuição Gama. Deste modo, temos como objetivo neste trabalho realizar um estudo sobre testes acelerados do tipo step-stress simples, considerando que os tempos de falha sob censura progressiva tipo II seguem uma distribuição Gama. A fim de cumprir este objetivo, será feita uma*

*análise estatística desses tempos de falha, em que utilizaremos os métodos estatísticos de inferência clássica e de inferência Bayesiana. Joint work with Fernando Antonio Moala.*

**Lucas Leite Cavalaro - UFSCar/USP.**

Intervalos de confiança bootstrap para a diferença de duas médias em variáveis do tipo proporção.

*Abstract: O estudo de taxas e proporções é comum em diversas áreas do conhecimento. Frequentemente essas variáveis podem assumir apenas valores no intervalo  $(0;1)$ . Nesses casos, elas podem ser denominadas variáveis do tipo proporção. Neste trabalho estudamos 7 métodos para a construção de intervalos de confiança para variáveis do tipo proporção: Intervalo de Confiança assumindo normalidade para diferenças de médias com variâncias diferentes; Intervalo de Confiança utilizando a distribuição assintótica do estimador de máxima verossimilhança; Intervalo de Confiança Bootstrap  $T$ ; dois métodos para obtenção de Intervalos de Confiança Bootstrap Percentis; Intervalo de Confiança Bootstrap  $BC$  (Bias-Corrected) e Intervalo de Confiança Bootstrap  $BCA$  (Bias-Corrected Accelerated). Também desenvolvemos programas para a realização de estudos dos métodos e selecionamos diferentes cenários para comparar o desempenho dos mesmos considerando as distribuições Beta e Simplex. Em cada cenário considerado, utilizando simulação de Monte Carlo obtemos, para cada método, estimativas da probabilidade de cobertura dos intervalos de confiança para a diferença de duas médias, além do comprimento dos mesmos e proporções de erros abaixo e acima do valor verdadeiro para o parâmetro estimado. Dentre os resultados obtidos, podemos destacar as boas performances dos Intervalos de Confiança assumindo normalidade para diferenças de médias com variâncias diferentes e dos Intervalos de Confiança Bootstrap  $T$ . Entretanto, em alguns dos cenários estudados estes métodos apresentam, em média, intervalos de confiança com comprimento superior aos demais intervalos de confiança bootstrap. Além disso, os resultados para as distribuições Beta e Simplex foram, em geral, semelhantes. Também realizamos uma aplicação a dados reais, com o objetivo de estudar se, na prática, os intervalos de confiança podem mudar consideravelmente dependendo do método escolhido. Através desta aplicação foi possível concluir que, dependendo do método escolhido, as conclusões acerca do problema estudado podem variar e, com isso, o estudo de simulação foi importante para identificar os métodos mais adequados para a construção de intervalos de confiança para a diferença de médias em variáveis do tipo proporção. Joint work with Brolo, Camila L.; Pereira, Gustavo H. A.*

**Marcelo Andrade da Silva - USP/UFSCar.**

Bayesian estimation of a Bifactor Generalized Partial Credit Model with different link functions.

*Abstract: Item response theory (IRT) is a class of statistical models used to es-*

imate latent traits of individuals, which are characteristics not observable, such as proficiency in an educational evaluation, satisfaction of the consumer about a product or service and aspects of the personality of the human being. It is common to find situations in which it is necessary to consider multiple characteristics for each individual submitted to an evaluation instrument, suggesting the need for a multidimensional structure of the latent traits. In this context, multidimensional IRT models (MIRT) have been widely used. There are different MIRT models with different structures for multidimensional latent traits. One of these models is called the bifactor IRT model, a particular case of the additive MIRT model, because it considers an additive structure between global and specific trait and assumes that the specific traits are orthogonal with respect to the global trait and each other. These models assume that individuals' responses to each item are directly influenced by both the global latent trait and the specific latent trait corresponding to the subtest of the item. For the estimation of parameters of bifactor IRT models, several researchers have used maximum likelihood (ML) and marginal maximum likelihood (MML) method. Bayesian estimation is little found in the literature. This work proposes a bifactor generalized partial credit model (bifac-GPC model) with flexible link functions as an alternative model to the bifactor graded response model (bifac-GR model) for the case of ordered polytomous responses. The No-U-Turn Sampler algorithm, an extension of the Hamiltonian Monte Carlo (HMC) method, is used via Stan software to estimate the parameters of the proposed model. We evaluate the recovery of bifac-GPC model parameters using the NUTS algorithm through some appropriate statistics and using the probit and logit link functions. We present an application of the proposed model, using a data set well known in the psychological area to map five supposed dimensions of personalities to the study of human characteristics. This data set contains the answers of 2800 individuals to 25 items, equally divided into 5 subtests, each subtest to evaluate one of five personality dimensions considered. Joint work with Anne Corinne Huggins-Manley and Jorge Luis Bazán.

### **Milton Miranda Neto - UFSCar/USP.**

#### Limit Theorems for the Elephant Random Walk.

*Abstract:* In this work we will study the Elephant Random Walk presented by Schutz and Trimper (2004), a discrete time, non-Markovian stochastic process with unlimited range memory that presents phase transition. Our objective will be to demonstrate the convergence almost sure for the diffusive and critical regime of the model, besides this result we also present the demonstration of the Central Limit Theorem for both regimes. For the supercritical regime we will demonstrate the convergence of the Elephant Random Walk to a non-normal random variable based on the articles (BAUR, BERTOIN, 2016) (BERCU, 2017<sup>+</sup>) and (COLETTI, Gava, SCHUTZ, 2017a). Finally, through the process correlation analysis, based on (COLETTI, Gava, SCHUTZ, 2017b) understand why the Central Limit Theorem is not satisfied for the supercritical regime.

**Rafael Soares Paixão - USP/UFSCar.**

Zero-Variance principle for Hamiltonian Monte Carlo in GJR-GARCH models.

*Abstract: In Econometrics, the family of GARCH models is widely employed to describe the variance of time series through innovations. In this context, to deal with the skewness and high kurtosis of financial returns, the GJR-GARCH can be used. Along with this model, the MCMC algorithm is able to simulate the posterior distribution in Bayesian approaches. However, this algorithm is computationally expensive to generate adequate estimators. In this paper, we improve the performance of MCMC for GJR-GARCH models. For that purpose, we employed the zero-variance principle and the Hamiltonian Monte Carlo (HMC) method, so as to generate good Bayesian estimators with few simulations and low computational cost. The results of the simulations show that our methodology efficiently achieves our purpose. Joint work with Ricardo Ehlers.*

**Ricardo Felipe Ferreira - UFSCar/USP.**

Stochastic Chains with Memory of Unbounded and Variable Length applied to Neuroscience.

*Abstract: One of the most important concerns in neuroscience is to understand the dynamics of the animal behavior, i.e., the joint action of a large number of neurons, body parts, and the environment. Mathematics is fundamental to understand this question and accompany the neuroscience progress. However, the experimenter has always only access to a small part of the neural system and, frequently, mathematical general models are considered for this problem. In this context, more effective models have been proposed as, for example, the stochastic chains with memory of unbounded and variable length on the past. In this work, we consider this class of stochastic processes to study the modelling of functional synaptic connections proposed by Zhit-ing Cai et al. (2017), which implemented a different type of context tree to estimate direct information. Joint work with DANIEL TAKAHASHI AND SANDRO GALLO.*

**Tiago Bernardes Kerr - UFU.**

High-Performance Autism in Labor (HPAL): A Study with University Students.

*Abstract: Introduction: Autism Spectrum Disorders (ASD) are a group of neurological disorders that include Autistic Disorder, Asperger's Syndrome, and unspecified behavioral disorders. In general, individual with ASD have poor social skills, poor ability to communicate verbally, obsessive attention to detail, resistance to change, sensory hyper-and/or hyposensitivity. Mathematical and artistic skills, ability to concentrate attention, strong long-term visual memory, eidetic memory, interest in methodical tasks, great ability to recognize patterns and obey rules have also been reported in ASD people. These skills have attracted the attention of high-technology companies*

who are always looking for competitive advantage. *Objective:* To identify possible determinatives of individuals with ASD characteristics, high work capacity, who can handle conflicts, communicate well, are focused, efficient and disciplined. This set of characteristics we named as High-Performance Autism in Labor (HPAL). *Methods:* In the identification of the possible determinants of HPAL people an extensive review of the literature was performed. Two evaluators defined individuals with HPAL profile independently. The definition was based on a battery of questions that include numerical skills, emotional intelligence, detail perception, focus, efficiency, and charisma. A sample with 388 university students were examined, 27 were classified as HPAL and 361 were classified as Non-HPAL by both evaluators. The *t*-test for independent samples was used in the comparison of the mean age of the groups, and the Chi-square test was used in the comparison of the percentage of men in the groups. A Logistic Regression of the HPAL outcome on the identified potential determinants was performed on the sample data. The Stepwise Backward method was used in the parsimonious selection of the variables included in the model. The autism quotient score (AQ-10) and issued rapport score (PRIR-10) were included as determinants. *Results:* We did not identify a significant difference between the mean age (21.7±2.6 versus 23.1±4.9,  $P=0.165$ ) and the proportion of men (51.6 % versus 51.8%,  $P=0.706$ ). The regression model indicated eight possible determinants of the HPAL profile that contributed significantly to the model. The Hosmer-Lemeshow test indicated that the goodness of the fit is satisfactory ( $P=0.804$ ) and the Nagelkerke *r*-square of 0.32 indicates that the model is useful for predicting HPAL profile. The predictions based on the model presented satisfactory discriminatory power (AUROC=0.87, 95%CI: 0.80 to 0.93,  $P=0.034$ ). The results indicated a positive association between HPAL and history of mother smoker (OR=3.27 with 95%CI: 1.09 to 9.81); history of sleep disturbance (OR=4.92 with 95%CI: 1.65 to 14.67); selfishness (OR=5.51 with 95%CI: 1.36 to 22.39); incompleteness (OR=7.62 and 95%CI: 2.11 to 27.60); diplomacy (OR=2.81 and 95%CI: 1.06 to 7.43); PRIR10 score (OR=2.43 and 95%CI=1.60 to 3.69) and AQ10 score (OR=1.98 and 95%CI: 1.47 to 2.66). The association between HPAL and history of phobias was negative (OR=0.28 and 95%CI: 0.09 to 0.86). *Conclusion:* The HPAL profile can be reasonably predicted by the independent variables of the logistic model adjusted. This profile is linked to the history of smoker mother, sleep disturbance, selfishness, incompleteness, absence of phobias, high PRIR-10 and AQ-10 scores. Joint work with José Fausto de Moraes.

### **Pedro Luiz Ramos - USP/UFSCar.**

Reliability-centered maintenance: analyzing failure in harvest sugarcane machine using some generalizations of the Weibull distribution.

*Abstract:* In this study, we considered five generalizations of the standard Weibull distribution to describe the lifetime of two important components of harvest sugarcane machines. The harvesters considered in the analysis does the harvest of an average of 20 tons of sugarcane per hour and their malfunction may lead to major losses,



*therefore, an effective maintenance approach is of main interesting for cost savings. For the considered distributions, mathematical background is presented. Maximum likelihood is used for parameter estimation. Further, different discrimination procedures were used to obtain the best fit for each component. At the end, we propose a maintenance scheduling for the components of the harvesters using predictive analysis. Joint work with Diego Nascimento, Camila Cocolo, Márcio J. Nicola, Carlos Alonso, Luiz G. Ribeiro, André Ennes and Francisco Louzada.*

**Victor Daniel Camarena Perez - National University of Engineering.**

Almost sure stability of discrete-time Markov jump linear systems.

*Abstract: The present work deals with the study of the almost sure stability of discrete-time Markov jump linear systems (MJLS). First, a review of the stability of discrete-time linear systems (MJLS with a single mode) is given by the Lyapunov exponent. Then the extension of this theory to the general case is addressed: use is made of Birkhoff's ergodic theorem to give a new proof of the equality for the Lyapunov exponent described in (Fang et al., 1995), from this we obtain the characterization of the stability almost certain of the MJLS.*

## Participants

Adriano Kamimura Suzuki, ICMC-USP  
 Afrânio M C Vieira, UFSCar  
 Alana Gabriela Salatim Novais, UFSCar  
 Alex de la Cruz, USP/UFSCar  
 Alex Rodrigo dos Santos Sousa, UNICAMP  
 Amanda Morales E. D'Andrea, UFSCar/USP  
 Ana P. J. do Espirito Santo, UFSCar/USP  
 André Alves Ambrósio, ICMC-USP  
 Andreza Aparecida Palma, UFSCar  
 Andson Nunes da Silva, ICMC-USP  
 Anna Caroline, UFES  
 Bruno Santos, UFBA  
 Bruno Vernaglia Zólio, UFSCar  
 Caio Lucidius N. Azevedo, UNICAMP  
 Camila Lorencetti Brolo, UFSCar  
 Carlos Hirth Pimentel, USP/UFSCar  
 Carolina Bueno Grejo, USP  
 Christian Galarza, UNICAMP  
 Cintia Isabel de Campos, EESC-USP  
 Claudia E. Escobar Montecino, USP/UFSCar  
 Clécio da Silva Ferreira, UFJF  
 Cristian Favio Coletti, UFABC  
 Cristian Villegas, ESALQ/USP  
 Cristina Nardin Zabotto, UFSCar  
 Daiane de Souza Santos, USP/UFSCar  
 Daiane Zuanetti, UFSCar  
 Daniel Perez, National Univ. of Engineering  
 Daniele Granzotto, UEM  
 Danilo Augusto Sarti, USP  
 Davi Keglevich Neiva, UFSCar  
 Démerson André Polli, UnB - UFSCar/USP  
 Diego Carvalho do Nascimento, USP/UFSCar  
 Diogo Barboza Moreira, UFSCar  
 Eduardo S. Bueno de Oliveira, UFSCar/USP  
 Edward George, University of Pennsylvania  
 Elizabeth Mie Hashimoto, UTFPR  
 Elizabeth Chipa Bedia, USP/UFSCar  
 Fabiana Arca Cruz Tortorelli, USP/UFSCar  
 Fernanda Rodrigues Vargas, UFRGS  
 Fernando Quintana, PUC de Chile  
 Francisco Antonio Loyola Lavin, USP  
 Gabriel Marcelino Alves, UFSCar  
 Gabriela Cintra Raquel, USP/UFSCar  
 George Lucas Pezzott, USP/UFSCar  
 Gesiel Rios Lopes, ICMC-USP  
 Gianpedro Robertto Mella Brigante, USP  
 Glauber Márcio Silveira Pereira, UFSCar  
 Gustavo Pereira, UFSCar  
 Hedibert Freitas Lopes, Insper  
 Hélio R. de Carvalho Nunes, Unesp-Botucatu  
 Isabela Thaís Machado de Jesus, UFSCar  
 Jean-Yves Dauxois, Université de Toulouse  
 Jeremias Leão, UFAM  
 Joao Carlos Poloniato Ferreira, UFSCar/USP  
 Jorge Bazán, USP  
 José Clelto Barros Gomes, USP-UFSCar  
 José Fausto de Moraes, UFU  
 Juliana Cobre, ICMC/USP  
 Juliana Marambaia Maia, USP/UFSCar  
 Juliana Scudilio Rodrigues, UFSCar/USP  
 Karina Alves, Universidade Luterana do Brasil  
 Karlla Delalibera Chagas, FCT-UNESP  
 Karoline Eduarda Lima Santos, USP  
 Lia Hanna Martins Morita, UFMT  
 Lorena Yanet Cáceres Tomaya, UFSCar/USP  
 Lucas Antonio Barbano, UFSCar  
 Lucas Eduardo de Moraes, UFSCar  
 Lucas Leite Cavalaro, UFSCar/USP  
 Lucas Pereira Lopes, USP/UFSCar  
 Luciana Moura Reinaldo, UFC  
 Luiz Carlos Medeiros Damasceno, UFSCar/USP  
 Luiz Gabriel Fernandes Cotrim, UFSCar/USP  
 Maira Fatoretto, ESALQ -USP  
 Manoel dos Santos Neto, UFSCar/UFCG  
 Marcello Neiva de Mello, ESALQ/USP  
 Marcelo Andrade da Silva, USP/UFSCar  
 Marcelo Bourguignon Pereira, UFRN  
 Marcos Jardel Henriques, UFSCar-USP  
 Mariana Curi, ICMC-USP  
 Marina Paez, UFRJ  
 Mário de Castro, USP  
 Mary Luz Rodiño, Universidad de Antioquia  
 Miguel Abadi, USP  
 Milene Alves Garcia, USP/UFSCar  
 Milton Miranda Neto, UFSCar/USP  
 Murilo Cantoni, UFSCar/USP  
 Nancy Garcia, UNICAMP  
 Nicholas Wagner Eugenio, IME-USP

Octávio Valentin Lourenço Agostinho, USP  
Oilson Alberto Gonzatto Junior, UFSCar/USP  
Pablo Rodriguez, USP  
Paulo Justiniano Ribeiro Junior, UFPR  
Pedro Luiz Ramos, USP/UFSCar  
Rafael Izbicki, UFSCar  
Rafael Soares Paixão, USP/UFSCar  
Rafael Stern, UFSCar  
Raphael Machado, UFSCar  
Renan de Padua, ICMC-USP  
Renata Porto Sampaio, Brazil  
Renato Gava, UFSCar  
Ricardo De Carli Novaes, USP  
Ricardo de Jesus Caldas Assis, USP/UFSCar  
Ricardo Ehlers, USP  
Ricardo Felipe Ferreira, UFSCar/USP  
Robinson N. dos Santos, Springer Nature  
Silvia Lopes de Paula Ferrari, USP  
Tainá Santana Caldas, UFRR  
Taís Roberta Ribeiro, USP/UFSCar  
Tatiane Carvalho Alvarenga, UFLA  
Teh Led Red, USP  
Thiago Fernando Ferreira Costa, USP  
Themis da Costa Abensur Leão, USP/UFSCar  
Tiago Bernardes Kerr, UFU  
Tiago Mendonça, USP  
Vanessa Rufino da Silva, USP/UFSCar  
Vera Tomazella, UFSCar  
Victor Lawrence Bernardes Santana, UFU  
Victor Vinicius Fernandes, UFSCar/USP  
Vinicius Calsavara, A.C. Camargo Cancer Cen.  
Walkira Maria de Oliveira Macerau, UFSCar  
Dr Wolfgang W Ryll, KfW Development Bank  
Yury Rojas Benites, USP/UFSCar

## Organizers



## Support

