

# *III Workshop de Verão em Matemática da UFCG*

## *Sessão de Probabilidade e Estatística*

**Duração das Palestras:** 50 minutos

**Tempo destinado para Perguntas:** 10 minutos

**Pesquisador:** Dr. Celso Rômulo Barbosa Cabral (DE-UFAM)

**Data:** 25/01/2020

**Horário:** 15h00min (Horário de Brasília)

**Título:** Misturas Finitas de Distribuições Assimétricas

**Resumo:** A modelagem estatística baseada em misturas finitas de distribuições é uma área de pesquisa com uma quantidade enorme de aplicações e que obteve um incremento significativo de popularidade nos anos recentes. Estas aplicações estão em (mas não limitam-se a) áreas como aprendizado de máquina, biologia, biometria, engenharia, computação, genética, medicina e marketing. Modelos estatísticos baseados em misturas finitas podem capturar eficientemente muitas propriedades da distribuição dos dados, tais como multimodalidade, caudas pesadas e assimetria. São modelos muito úteis para identificar heterogeneidade não observável, ou seja, quando sabemos que as observações são provenientes de grupos distintos da população, mas não sabemos como discriminá-las dentre esses grupos (que são chamados componentes da mistura). Misturas finitas também constituem uma classe de distribuições extremamente flexível. Basicamente, podemos dizer que qualquer densidade de probabilidade contínua pode ser aproximada por uma mistura finita conveniente. O principal objetivo desta apresentação é revisar uma sequência de trabalhos publicados pelo palestrante na área de misturas finitas de distribuições, em colaboração com pesquisadores brasileiros e estrangeiros, nos últimos 10 anos. O foco principal será em uma classe de misturas finitas ainda mais flexível que o usual, utilizando como componentes misturas de escala da distribuição normal assimétrica.

**Pesquisador:** Ma. Maria Ioneris Lima Oliveira (Doutoranda do PPGE - UFPE)

**Data:** 26/01/2020

**Horário:** 14h00min (Horário de Brasília)

**Título:** Modelo de Tempo de Vida Acelerado Birnbaum-Saunders Bivariado

**Resumo:** A distribuição Birnbaum-Saunders (BS) é uma distribuição de tempo de vida que foi desenvolvida com o intuito de analisar dados de fadiga de materiais. Nas últimas décadas o modelo BS univariado vem sendo bastante utilizado para este fim e também vem ganhando destaque por sua aplicabilidade em outras áreas do conhecimento. A maioria dos trabalhos relacionados aos modelos BS foram desenvolvidos assumindo que as unidades experimentais são independentes, porém a suposição de independência nem sempre é plausível. Neste sentido, apresentaremos nesta palestra o modelo de tempo de vida acelerado BS bivariado, um estudo inferencial e um estudo de simulação para verificar propriedades de um resíduo proposto. Além disso, a metodologia apresentada é aplicada a um conjunto de dados reais.

**Pesquisador:** Dr. Marcelo Bourguignon Pereira (DEST-UFRN)

**Data:** 27/01/2020

**Horário:** 14h00min (Horário de Brasília)

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

**Resumo:** 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

**Pesquisador:** Dr. Eliardo Guimarães da Costa (DEST-UFRN)

**Data:** 28/01/2020

**Horário:** 14h00min (Horário de Brasília)

**Título:** Sample size for estimating organism concentration in ballast water: A Bayesian approach

**Resumo:** Estimation of microorganism concentration in ballast water tanks is important to evaluate and possibly to prevent the introduction of invasive species in stable ecosystems. For such purpose, the number of organisms in ballast water aliquots must be counted and used to estimate their concentration with some precision requirement. Poisson and negative binomial models have been employed to describe the organism distribution in the tank, but determination of sample sizes required to generate estimates with pre-specified precision is still not well established. A Bayesian approach is a flexible alternative to accommodate adequate models that account for the heterogeneous distribution of the organisms and may provide a sequential way of enhancing the estimation procedure by updating the prior distribution along the ballast water discharging process. We adopt such an approach to compute sample sizes required to construct credible intervals obtained via two optimality criteria that have not been employed in this context. Such intervals may be used in the decision with respect to compliance with the D-2 standard of the Ballast Water Management Convention.

Joint work with **Julio Singer** (IME-USP) and **Carlos Daniel Paulino** (Universidade de Lisboa)

**Pesquisador:** Dr. Tiago Almeida de Oliveira (DE-UPEB)

**Data:** 29/01/2020

**Horário:** 14h00min (Horário de Brasília)

**Título:** Um pequeno overview sobre Machine Learning: Um estudo de caso com modelos de prognóstico para desfechos negativos em Covid-19

**Resumo:** Aprendizado de máquina, análise preditiva e outros tópicos relacionados são campos muito excitantes e poderosos da Inteligência Artificial. Neste sentido, o Aprendizado de Máquina (Machine Learning) é um campo de estudo da Inteligência Artificial cujo objeto de estudo são sistemas que podem aprender a realizar alguma tarefa por meio de experiências. Neste seminário, o objetivo é apresentar uma breve introdução aos conceitos, modelos, métodos, técnicas e aplicações do Aprendizado de Máquina a dados de COVID-19 que é a doença gerada pelo SARS-CoV-2. A nova doença coronavírus (COVID-19) é um desafio para a tomada de decisões clínicas e a efetiva alocação de recursos de saúde. Uma avaliação prognóstica precisa é necessária para melhorar a sobrevivência dos pacientes, especialmente nos países em desenvolvimento. Pretende-se apresentar um estudo que propõe prever o risco de desenvolver condições críticas em pacientes COVID-19, por meio do treinamento de algoritmos multiuso.