

---

Livro de resumos da 61<sup>a</sup> Reunião Anual da  
Região Brasileira da Sociedade Internacional de  
Biometria (RBras)

“Biometria e Bioestatística na Era de *Big Data*”

Book of abstracts of the 61st annual meeting of the Brazilian Region  
of the International Biometric Society (RBras)

“Biometry and Biostatistics in the Era of Big Data”

---

23 - 25 de maio de 2016

Salvador - BA, Brasil

***Editoração***

Paulo Jorge Canas Rodrigues  
Elisabete Batista Sampaio  
Cristiane Nazaré Pamplona de Souza  
Jalmar Manuel Farfán Carrasco  
Rodrigo de Souza Bulhões  
*Universidade Federal da Bahia, Salvador, Brasil*

Aldo William Medina Garay  
*Universidade Federal de Pernambuco, Recife, Brasil*

***Web Design***

Arthur Henrique Farias  
*Universidade Federal da Bahia, Salvador, Brasil*

***Design da Capa***

Arthur Henrique Farias  
*Universidade Federal da Bahia, Salvador, Brasil*

ISSN: 2238-4162

# Contents

---

## Part I. Introduction

---

Welcome to the 61st RBras 2016 in Salvador, Brazil .....	25
Committees .....	27

---

## Part II. Scientific Program

---

Scientific Program .....	32
--------------------------	----

---

## Part III. Social Program

---

Social Program .....	35
----------------------	----

---

## Part IV. Keynote Lectures

---

<b>KS1: Using dependence among individuals and across genomes in the genetic mapping of quantitative traits</b> .....	39
<i>Elizabeth A. Thompson</i>	
<b>KS2: Model-based geostatistics for prevalence mapping in low-resource settings</b> .....	40
<i>Peter J. Diggle</i>	
<b>KS3: Statistical thinking and methodology – A pillar for quality also in the Big Data Era</b> .....	41
<i>Pedro Luis do Nascimento Silva</i>	

---

## Part V. Short Course

---

<b>MC1: Tools for missing data and sensitivity analysis</b> .....	45
<i>Geert Molenberghs</i>	
<b>MC2: An introduction to modern statistical methods for analysis of gene-environment interactions</b> .....	45
<i>Bhramar Mukherjee</i>	
<b>MC3: Modelos de regressão para dados de contagem com o R</b> .....	46
<i>Walmes Marques Zeviani, Eduardo Elias Ribeiro Junior e Cesar Augusto Taconeli</i>	
<b>MC4: Como você compara seus modelos? Quais os critérios você conhece? Saiba quais são suas opções e conheça a medida L</b> .....	46
<i>Elayne Penha Veiga e Mario Javier Ferrua Vivanco</i>	
<b>MC5: Pesquisa reproduzível com o R: de documentos dinâmicos a pacotes</b> .....	47
<i>Fernando Mayer e Walmes Zeviani</i>	

<b>MC6: Geoestatística para análise e modelagem de dados biométricos</b> .....	48
<i>Mário Sebastiao Tuzine, Anderson Pedro Batista e José Márcio de Mello</i>	
<b>MC7: Diagrama de Hasse no planejamento de experimentos e na análise de dados</b> ....	48
<i>Clarice G.B. Demétrio e Renata Alcarde Serمارini</i>	
<b>MC8: Análise de dados genéticos: Um problema de Big Data a cada novo paciente</b> ...	49
<i>Marcus Nunes</i>	
<b>MC9: Redução de dimensionalidade – Aplicações à genômica</b> .....	49
<i>Júlia M. Pavan Soler e Adèle H. Ribeiro</i>	

---

## Part VI. Invited Paper Sessions

---

<b>IPS2 – Twenty years of P-splines, and the future is bright</b> .....	53
<b>IPS2.1: The analysis of Big Data using P-splines and sparse matrix algebra</b> .....	53
<i>Martin Boer</i>	
<b>IPS2.2: Spatial and spatio-temporal modelling of aggregated Counts with applications in epidemiology: Methods and tools</b> .....	53
<i>Dae-Jin Lee</i>	
<b>IPS2.3: Spatial models for field trials using P-splines</b> .....	54
<i>María Xosé Rodríguez-Álvarez</i>	
<b>IPS3 – Modelos Mistos e Aplicações</b> .....	55
<b>IPS3.1: Modelos mistos potencialidades e desafios nas ciências biológicas</b> .....	55
<i>Viviana Giampaoli</i>	
<b>IPS3.2: A mixed model for positive responses augmented by zeros</b> .....	55
<i>Mariana Motta Rodrigues</i>	
<b>IPS4 – Robust statistics in the 21st century: New developments</b> .....	57
<b>IPS4.1: Functional outlier detection with applications to MRI image data</b> .....	57
<i>Mia Hubert, Peter J. Rousseeuw, Jakob Raymaekers and Pieter Segaert</i>	
<b>IPS4.2: Detecting anomalous data cells</b> .....	57
<i>Peter J. Rousseeuw and Wannes van den Bossche</i>	
<b>IPS4.3: An R package for robust orthogonal regression for compositional data</b> .....	58
<i>Valentin Todorov, Klára Hrušová, Karel Hron and Peter Filzmoser</i>	
<b>IPS5 – Innovative statistical methods with applications in medicine and biology</b> .....	59
<b>IPS5.1: Bivariate Clayton copula-based SUR Tobit model: A modified inference function for margins, interval estimation and an application to food consumption data</b> ....	59
<i>Paulo Henrique Ferreira and Francisco Louzada</i>	
<b>IPS5.2: SISGERES: A statistical system for urban solid waste management</b> .....	60
<i>Francisco Louzada</i>	

<b>IPS5.3: Survival modelling for prediction of intrapartum interventions</b> .....	60
<i>Gleici da Silva Castro Perdoná</i>	
<b>IPS6 – New approaches for big and complex data in biostatistics</b> .....	61
<b>IPS6.1: On the robust estimation of trait heritability</b> .....	61
<i>Vanda M. Lourenço, Miguel Fonseca, Ana M. Pires and Paulo C. Rodrigues</i>	
<b>IPS6.2: Set-based interaction analysis: Some new observations</b> .....	61
<i>Bhramar Mukherjee</i>	
<b>IPS6.3: Pseudo-likelihood and split-sample methods in small and very large trials</b> ....	62
<i>Geert Molenberghs</i>	
<b>IPS7 – Estatística Experimental</b> .....	63
<b>IPS7.1: On compound design criteria</b> .....	63
<i>Luzia Aparecida Trinca</i>	
<b>IPS7.2: Variance component estimation on discrete time competing risks analysis with masked causes</b> .....	63
<i>Rafael Pimentel Maia</i>	
<b>IPS7.3: Case study of sugarcane trials with commercial varieties and test lines</b> .....	64
<i>Alessandra dos Santos, Chris J. Brien, Clarice Garcia Borges Demétrio and Renata Alcarde Sermarini</i>	
<b>IPS8 – Sessão ABE – Projetos de divulgação e educação estatística</b> .....	65
<b>IPS8.1: Literacia, capacitação e divulgação estatística no ISI</b> .....	65
<i>Pedro Luis do Nascimento Silva</i>	
<b>IPS8.2: Reflexões sobre o potencial do ensino de estatística na educação básica</b> .....	65
<i>Irene Mauricio Cazorla</i>	
<b>IPS8.3: Projeto Bussab de divulgação e educação estatística</b> .....	66
<i>Francisco Louzada</i>	
<b>IPS9 – Sessão EMBRAPA</b> .....	67
<b>IPS9.1: Embrapa: “Quem somos nós?”</b> .....	67
<i>Waldomiro Barioni Jr. e Carlos Alberto da Silva Ledo</i>	
<b>IPS9.2: Métodos quantitativos e computação científica na Embrapa</b> .....	67
<i>Geraldo da Silva e Souza</i>	
<b>IPS9.3: Armazenamento e processamento de grandes volumes de dados na Embrapa</b> ..	68
<i>Roberto Hiroshi Higa</i>	

---

**Part VII. Round Tables**


---

<b>RT1: Meet the Presidents</b> .....	71
<i>Paulo Canas Rodrigues and John Hinde</i>	

<b>RT2: Tecnologia e estatística aplicadas a políticas de saúde</b> .....	72
<i>Doris Fontes</i>	
<hr/>	
<b>Part VIII. Contributed Papers</b>	
<hr/>	
<b>CPS1 - Recent advances in survival analysis</b> .....	75
<b>CPS1.1: The exponentiated Kumaraswamy inverse Weibull distribution</b> .....	75
<i>Jailson A. Rodrigues, Gholamhossein Hamedani, Ana Paula Silva and Angela Silva</i>	
<b>CPS1.2: Regression models for the restricted mean residual time</b> .....	75
<i>Giuliana Cortese, Stine Holmboe and Thomas H. Scheike</i>	
<b>CPS1.3: A flexible cure rate model with frailty term in latent risks: a Bayesian approach</b>	76
<i>Agatha Rodrigues and Vinicius Calsavara</i>	
<b>CPS1.4: Subconjuntos de validação em análise de sobrevivência com censura intervalar e erros de classificação</b> .....	76
<i>Magda Pires, Enrico Colosimo, Guilherme Veloso e Raquel Ferreira</i>	
<b>CPS2 – Recent advances in mixed models and applications</b> .....	77
<b>CPS2.1: Modelos mistos de regressão segmentada: Um exemplo</b> .....	77
<i>Julio Singer, Francisco Marcelo Monteiro-Da-Rocha, Antonio Carlos Pedroso-De-Lima, Giovani Silva, Giuliana Coatti e Mayana Zatz</i>	
<b>CPS2.2: Análise de dados longitudinais da evapotranspiração de alface em função de níveis de água</b> .....	77
<i>Omar Pereira, Altair Bertonha e Isolde Previdelli</i>	
<b>CPS2.3: The longitudinal approach to evaluating changes behavioural in rats with cerebral ischemia induced by chronic cerebral hypoperfusion</b> .....	78
<i>Matheus Henrique Dal Molin Ribeiro, Omar Pereira, Humberto Milani and Isolde Previdelli</i>	
<b>CPS2.4: Modelo linear misto: Aplicações no mapeamento de determinantes genéticos associados com depressão e ansiedade</b> .....	78
<i>Tâmara P. Taporoski, André B. Negrão, Andréa R. V. R. Horimoto, Nubia E. Duarte, Mariza de Andrade, José E. Krieger, Malcolm von Schantz, Homero Vallada e Alexandre C. Pereira</i>	
<b>CPS3 – Applications in biometry and biostatistics</b> .....	80
<b>CPS3.1: Accuracy in the estimates of zucchini production</b> .....	80
<i>Alessandro D. Lúcio and Vilson Benz</i>	
<b>CPS3.2: Sample size for estimating the concentration of organisms in ballast water</b> ...	80
<i>Eliardo Costa, Rubens Lopes and Julio Singer</i>	
<b>CPS3.3: Distribuições zero-modificadas biparamétricas: Uma aplicação aos dados de variações de temperatura média global</b> .....	81
<i>Luisa Hebling e Katiane Conceição</i>	

<b>CPS3.4: Estudos, em termos estatísticos, das condições de moradia para pessoas com deficiências</b> .....	81
<i>Paulo Oliveira</i>	
<b>CPS4 – Regression and spatio-temporal modelling</b> .....	82
<b>CPS4.1: On quantile residuals in beta regression</b> .....	82
<i>Gustavo Pereira</i>	
<b>CPS4.2: Previsão de preços para soja e café com base em variáveis climáticas: Aplicação da análise espectral multivariada</b> .....	82
<i>Carlos Pinheiro e Valter Senna</i>	
<b>CPS4.3: Skew scale mixtures of normal distributions and applications</b> .....	83
<i>Clécio Ferreira and Victor Hugo Lachos</i>	
<b>CPS4.4: Análise de fronteiras geográficas e limites no processo temporal para avaliação de conglomerados espaço temporais</b> .....	83
<i>Spencer Barbosa da Silva, Anderson Ribeiro Duarte, Fernando Luiz Pereira de Oliveira1, Alexandre Celestino Leite Almeida e Luiz Duczmal</i>	
<b>CPS5 – Computer intensive methods and software development</b> .....	84
<b>CPS5.1: A new likelihood method for three-parameter gamma distribution fitting</b> ....	84
<i>Simplíce Dossou Gbete and Quédraogo Etienne</i>	
<b>CPS5.2: A new process for modeling count time series with overdispersion</b> .....	84
<i>Marcelo Bourguignon Pereira</i>	
<b>CPS5.3: O modelo de regressão Weibull Modificado Exponenciado de longa duração para dados oncológicos: Estudo de simulação e desenvolvimento de uma ferramenta Web</b> .....	85
<i>Hayala Souza, Gleici Perdoná, Francisco Louzada, Ana Paula de Andrade, Francisco Barbosa Junior e Fernanda Peria</i>	
<b>CPS5.4: Ferramenta para o cálculo dos limites de controle para variáveis, baseado em R utilizando interface gráfica TclTk</b> .....	85
<i>Jonathan Silva, Flaviane Peccin Brevi e Fernando Henrique Rosa</i>	

---

## Part IX. Contributed Posters 1

---

<b>Análise descritiva e exploratória e o uso de outras aplicações metodológicas [<i>Descriptive and exploratory statistics and the use of other methodological applications</i>]</b> .....	89
<b>CP1 – SAS codes for quantifying influence of the El Niño/Southern oscillation phenomenon on climate variable distributions</b> .....	89
<i>Aline Maia</i>	
<b>CP2 – Análise estatística em saúde reprodutiva pré e pós-gestacional de adolescentes na cidade de Campina Grande, Paraíba</b> .....	89
<i>Alisson Brito, Rodolfo Silva, Marcio Albuquerque e Walber Colaço</i>	

<b>CP3 – Detecção sensorial de grãos precoces no café torrado e moído</b> .....	90
<i>Allan Fernandes, Paulo Lima, Mariane Rabelo, Rosemary Pereira e Renato R. de Lima</i>	
<b>CP4 – Impact evaluation using propensity score matching for analysis of school enrollment rate data from rural Burkina Faso</b> .....	90
<i>Ana Clara Campos and Leila D.A.F. Amorim</i>	
<b>CP5 – Explorando tópicos de estatística básica com o GeoGebra</b> .....	91
<i>Ana Maria Lemes e Adriana Strieder Philippsen</i>	
<b>CP6 – Fatores que impactam o desempenho dos alunos do 9º ano do estado de São Paulo nas provas do SAEB 2013</b> .....	91
<i>Angelica Cristina Custodio Cruz, Pedro Ferreira Filho e Estela Maris Pereira Bereta</i>	
<b>CP7 – Comparação de métodos de agrupamento em acessos de tabaco</b> .....	92
<i>Antonio Leandro da Silva Conceição, Carlos Alberto da Silva Ledo, Mauricio dos Santos da Silva, Sandra Domingos João Afonso e Ricardo Franco Cunha Moreira</i>	
<b>CP8 – Percepção ambiental dos moradores da cidade de Coremas-PB em relação ao esgotamento sanitário</b> .....	93
<i>Carlos Sérgio Araújo dos Santos, Daniel Jackson Andrade de Sousa, Gabriel Carlos Moura Pessôa, Alan Dél Carlos Gomes Chaves e Ricardo Ricelli P. de Almeida</i>	
<b>CP9 – Comparação de métodos para avaliação de impacto do programa KALAHÍ na redução da pobreza nas Filipinas</b> .....	93
<i>Cristiane N.P. Souza e Leila D.A.F. Amorim</i>	
<b>CP10 – Programas governamentais e seus efeitos na mortalidade infantil do semiárido Brasileiro</b> .....	94
<i>Everlane Araújo e Neir A. Paes</i>	
<b>CP11 – Testes de não-inferioridade: Uma proposta para validação de equipamentos</b> ...	94
<i>Flaviane Peccin, Fernando Henrique Rosa e Jonathan Silva</i>	
<b>CP12 – Precipitação pluviométrica mensal em Uberlândia-MG: Análise descritiva e teste de normalidade</b> .....	95
<i>Gabriela Maria Rocha Bolaina, Matheus Ferreira do Amaral e Ednaldo Carvalho Guimarães</i>	
<b>CP13 – IPESE: Uma análise da performance socioeconômica dos municípios Baianos</b> .	95
<i>Jonatas Espírito Santo, Alex Santos, Jadson Silva e Urandi Freitas</i>	
<b>CP14 – Doença de chagas aguda: Aspectos associados a confirmação do caso</b> .....	96
<i>José Luiz De Carvalho Lisboa, Cristiane Nazaré Pamplona de Souza, Edson Marcos Leal Soares Ramos e Sílvia dos Santos de Almeida</i>	
<b>CP15 – Relação entre transtorno de personalidade antissocial, usuários de drogas e criminalidade</b> .....	96
<i>Joyce Andrade, Thomas Barbosa e Elizeu Alves</i>	
<b>CP16 – Estudos probabilísticos do índice de desenvolvimento humano municipal (IDH-M) do norte de Minas Gerais (1991 a 2010) - análise via Krigagem ordinária e indicativa</b>	97
<i>Maria de Fátima Ferreira Almeida, Jacqueline Domingues, Vivian Aparecida Brancaglioni e José Sílvio Govone</i>	

<b>CP17 – Análise univariada dos índices Bovespa e Petrobras entre 2005-2015</b> .....	97
<i>Maria Eugênia Oliveira Batista, Luciene Resende Gonçalves e Gabriel Rodrigo Gomes Pessanha</i>	
<b>CP18 – Modelos lineares generalizados na análise da estrutura de comprimento da manjuba (<i>Opisthonema oglinum</i>) capturada no canal de Santa Cruz-PE</b> .....	98
<i>Nathielly Rego, Renata Silva e Humber Andrade</i>	
<b>CP19 – Descrição do consumo diário de ração para gatos usando modelos lineares e não lineares</b> .....	98
<i>Paula Ribeiro Santos, Izabela Regina Cardoso de Oliveira, Jéssica Santana dos Reis e Flávia M. de Oliveira B. Saad</i>	
<b>CP20 – Cadeia de Markov aplicada a bioacústica</b> .....	99
<i>Rayssa Ribeiro, João Marcelo Protazio and Maria Luisa Silva</i>	
<b>CP21 – Acidente de trânsito de motocicleta e carro em João Pessoa</b> .....	99
<i>Rodolfo Silva, Alisson Brito e Mácio Albuquerque</i>	
<b>CP22 – Panorama dos três corredores com maior número de acidentes na cidade de Campina Grande-PB</b> .....	99
<i>Sônia Santos, Wanessa Wanderley, Tanise Souza, Eduarda Bizerra e Ricardo Olinda</i>	
<b>CP23 – Redução da mortalidade infantil no nordeste Brasileiro para os anos de 2010 e 2010 e perspectivas até 2015</b> .....	100
<i>Vanessa Santos, Suy Santos e Mateus Reis</i>	
<b>Modelos de Regressão [<i>Regression Models</i>]</b> .....	101
<b>CP24 – Testes de Paralelismo: Aplicação para comparação de curvas analíticas de água</b>	101
<i>Adson Costanzi Filho, Flaviane Peccin, Thais Aline Pohren Ferreira e Jonathan Silva</i>	
<b>CP25 – Aplicações da decomposição espectral da matriz de parentesco no modelo misto poligênico</b> .....	101
<i>Alex Ribeiro, Daniel F. Ferreira e Júlia Pavan</i>	
<b>CP26 – Aplicação e comparação de MLG para precificação do contrato do seguro agrícola no estado de Sergipe</b> .....	102
<i>Amanda da Silva Lira, Luelton Nunes Santos, Giovani Oliveira Santos, Elielma Santana de Jesus e Kleber Regis Santoro</i>	
<b>CP27 – Diagnostic accuracy and inter-observer agreement measures to dependent data</b>	102
<i>Ana Carolina Cintra Nunes Mafra, Elivane da Silva Victor and Thaís Caldara Mussi Andrade</i>	
<b>CP28 – Uso da transformação isotônica no ajuste do acúmulo de massa seca por modelo de regressão não linear e análise da variância</b> .....	103
<i>Augusto Moraes, Laís Silva, Messias J.B. Andrade, Dâmiany Oliveira, Elisa Moraes e Thaís Ribeiro</i>	
<b>CP29 – Uma comparação entre modelo fuzzy e regressão linear na relação entre indicadores de capital de giro e o risco da empresa</b> .....	103
<i>Beatriz Rezzieri Marchezini, Lucas Pereira Lopes e Gabriel Rodrigo Gomes Pessanha</i>	

<b>CP30 – Regressão logística com dados repetidos aplicada à modelagem de incontinência urinária feminina</b> .....	104
<i>Cecília Candolo, Ana Carolinne Portela Rocha e Fernanda Vieira Rodovalho Callegari</i>	
<b>CP31 – Estudo da degradabilidade in situ da fibra em detergente neutro utilizando modelos não lineares mistos</b> .....	104
<i>César Gonçalves de Lima, Euclides Braga Malheiros, Simone Daniela Sartorio e José Valmir Feitosa</i>	
<b>CP32 – Uso de modelos não lineares para modelagem da área basal e volume de <i>Eucalyptus</i> sp.</b> .....	105
<i>Elliezer de Almeida Melo, Joyce de Almeida Alves, Ernani Lopes Possato, Marcos Antônio Isaac Jr, Ernandes Guedes Moura e Natalino Calegario</i>	
<b>CP33 – Uso de modelos não lineares para modelagem do crescimento das árvores dominantes e codominantes de <i>Eucalyptus</i> sp.</b> .....	105
<i>Elliezer de Almeida Melo, Joyce de Almeida Alves, Ernani Lopes Possato, Marcos Antônio Isaac Jr, Gabriela Cristina Costa Silva e Natalino Calegario</i>	
<b>CP34 – Análise do desempenho da probabilidade de cobertura para o modelo Garma com distribuição condicional de Poisson</b> .....	106
<i>Elza Blank, Adriana Strieder Philippsen e Marinho G. Andrade</i>	
<b>CP35 – Análise do ciclo reprodutivo das lagostas vermelha (<i>Panulirus Argus</i>) e verde (<i>Panulirus Laevicauda</i>) com modelos lineares generalizados</b> .....	106
<i>Evelyn Oliveira e Humber Andrade</i>	
<b>CP36 – Modelos sigmoidais para descrever o afilamento do fuste de árvores de eucalipto</b>	107
<i>Joyce de Almeida Alves, Ernani Lopes Possato, Geisi Azevedo Silva, Elliezer de Almeida Melo e Natalino Calegario</i>	
<b>CP37 – Testes de hipótese para os parâmetros do modelo de regressão logística</b> .....	107
<i>Juliana Maia e Verônica Lima</i>	
<b>CP38 – Utilização de modelos não lineares na descrição do acúmulo de boro no feijoeiro</b>	108
<i>Kelly Pereira de Lima, Augusto Ramalho Moraes, Messias José Bastos de Andrade, Sidcleide Barbosa de Sousa e Dâmiany Pádua Oliveira</i>	
<b>CP39 – Modelos lineares hierárquicos normais</b> .....	108
<i>Kelly Pereira de Lima, Sidcleide Barbosa de Sousa Ronald Targino Nojosa e Augusto Ramalho Moraes</i>	
<b>CP40 – Comparação dos modelos de Gompertz e Verhulst no ajuste do acúmulo de nitrogênio no feijoeiro cv. Jalo</b> .....	109
<i>Laís Silva, Allan Fernandes, Thaís Ribeiro, Augusto Moraes, Joel Muniz, Messias J.B. Andrade e Neiva Vieira</i>	
<b>CP41 – Identificação dos fatores de riscos associados à ocorrência de anomalias congênitas em recém nascidos utilizando a regressão logística</b> .....	109
<i>Lidia Dayse Araujo de Souza, Isis Milane Batista de Lima, Ianne Rafaella Santos Melo e Jessica da Silva Braga</i>	

<b>CP42 – Modelo linear misto na análise de perfis de pesos de filhotes de ratos com mudança acentuada de tendência</b> .....	110
<i>Luzia Pedroso de Oliveira, César Gonçalves de Lima, Vera Lúcia Scherholz Salgado de Castro e Aline de Holanda Nunes Maia</i>	
<b>CP43 – Modelo logístico para ajuste de curvas de crescimento de <i>Eucalyptus</i> sp.</b> .....	110
<i>Magali Teresopolis Reis Amaral e Carlo Roberto Padovani</i>	
<b>CP44 – Comparação dos modelos individuais e híbridos de Credit Scoring utilizando regressão logística e rede neural</b> .....	111
<i>Mariana Nolde Pacheco e Lisiane Priscila Roldão Selau</i>	
<b>CP45 – Evolution of electricity consumption residential in geographical regions of Brazil: Comparison of linear regression coefficients</b> .....	111
<i>Matheus Ferreira do Amaral, Raphael Silva, Gabriela Maria Rocha Bolaina and Ednaldo C. Guimarães</i>	
<b>CP46 – Adjustment of nonlinear models applied to data from the artificial population dynamics of blowflies</b> .....	112
<i>Paulo Henrique Sales Guimarães</i>	
<b>CP47 – Data analysis with repeated measures for an experiment sugar cane</b> .....	112
<i>Ricardo Olinda, Pedro Chinelato, João Gil Luna and Ednário Mendonça</i>	
<b>CP48 – Utilização da linguagem Stan na análise de dados de dose-resposta</b> .....	112
<i>Roseli Aparecida Leandro, Erasnilson Vieira Camilo e José Bruno Malaquias</i>	
<b>CP49 – Additive regression model: An example without and with dynamic covariates</b>	113
<i>Rosemeire Fiaccone</i>	
<b>CP50 – Análise de trilha na influência de componentes morfométricos do bagre americano (<i>Ictalurus punctatus</i>) no seu rendimento de filetagem</b> .....	113
<i>Tadeu Vilela de Souza, Adriano Carvalho Costa, João Domingos Scalon, Rilke Tadeu Fonseca de Freitas e Thiago Archangelo Freato</i>	
<b>CP51 – Regressão ridge: Uma abordagem para solucionar o problema da multicolinearidade</b> .....	114
<i>Tatiana Kazue Yamanouchi e Pollyane Vieira da Silva</i>	
<b>CP52 – Ajuste do modelo de Kuhn na descrição da isoterma de adsorção aplicada a dados de guaraná em pó</b> .....	114
<i>Thais Ribeiro, Laís Silva, Joel Muniz, Augusto R. Moraes e Gelson Ribeiro Júnior</i>	
<b>CP53 – Uso dos modelos exponencial simples com três parâmetros e Henderson para descrever a secagem de polpa de jabuticaba</b> .....	115
<i>Thais Ribeiro, Michele Barbosa, Joel Muniz e Marcio Balestre</i>	
<b>CP54 – Qualidade do ajuste do modelo Poisson-Weibull de longa duração na presença de covariáveis</b> .....	115
<i>Valdemiro Vigas, Giovana Oliveira e Francisco Louzada</i>	
<b>Bioestatística e Modelos Biométricos 1 [<i>Biostatistics and Biometrical Models 1</i>]</b> .....	117

<b>CP55 – Parameterization of Davis dynamic mathematical model to individually estimate the performance of cross beef cattle</b> .....	117
<i>Adrielle Giaretta Biase, Carlos Tadeu dos Santos Dias, Luís Gustavo Barioni, Tiago Zanett Albertini, Lucieta Guerreiro Martorano, James W. Oltjen, Dante Pazzanese Lanna, Patrícia Perondi Anchão Oliveira, Sérgio Raposo de Medeiros and Roberto Augusto de Almeida Torres Júnior</i>	
<b>CP56 – Dimensões de unidades experimentais básicas e de parcelas em tremoço branco</b>	118
<i>Alberto Cargnelutti Filho, Bruna Mendonça Alves, Marcos Toebe e Giovani Facco</i>	
<b>CP57 – Practical application of the sign test on data from trial and meta-analysis</b> ....	118
<i>Alfredo José Barreto Luiz, Adriane da Rocha Costa, Eunice Reis Batista and Rosa Toyoko Shiraishi Frighetto</i>	
<b>CP58 – Martingales: Teoria, simulações e aplicações</b> .....	119
<i>Allan Robert Silva e Cláudio T. Cristino</i>	
<b>CP59 – Avaliação da associação entre o desenvolvimento humano da população de Sergipe em relação à pecuária leiteira</b> .....	119
<i>Amanda da Silva Lira, Gabrielle Monteiro Souza, Vanessa Kelly dos Santos, Luydson Jamyson do Nascimento Souza e Kleber Régis Santoro</i>	
<b>CP60 – Use of factor analysis to evaluate everyday themes students of the 6th grade of elementary school</b> .....	120
<i>Amanda Liz Pacífico Manfrim Peticarrari, Clara Augusta Marques Natalin, Andreia da Silva Meyer and Aline Maria Pacífico Manfrim</i>	
<b>CP61 – Longitudinal analysis of inter-rater agreement and associated factors in medical records audit</b> .....	120
<i>Ana Carolina Cintra Nunes Mafra, Renata Rafaella Santos Tadeucci Medina, Ederson Haroldo Pereira Almeida and Mario Maia Bracco</i>	
<b>CP62 – Latent class analysis with covariates for evaluation of asthma symptoms</b> .....	121
<i>Beatriz Santos, Leila D.A.F. Amorim and Ana Clara Campos</i>	
<b>CP63 – Relações lineares entre caracteres em genótipos transgênicos de milho</b> .....	121
<i>Bruna Mendonça Alves e Alberto Cargnelutti Filho</i>	
<b>CP64 – Inclusão de variável multicategórica na análise de trilha da produtividade de alho</b> .....	122
<i>Brunna Rithielly Rezende, Divino de Sousa Costa, Hygor Amaral Santana, Wilhan Valasco dos Santos e Anderson Rodrigo da Silva</i>	
<b>CP65 – Tamanho de amostra para estimar o coeficiente de correlação de Pearson em experimentos com tomate cereja</b> .....	122
<i>Bruno Giacomini Sari, Alessandro Dal'col Lúcio, Cintya Souza Santana, André Luis Tischer, Andrielle Magrini Rodrigues, Letícia Gall Dreifke e Lucas Drebes</i>	
<b>CP66 – Relações entre caracteres produtivos de tomate tipo cereja</b> .....	123
<i>Cintya Souza Santana, Alessandro Dal'col Lúcio, Bruno Giacomini Sari, Lucas Drebes, Andrielle Magrini Rodrigues, André Luis Tischler e Letícia Gall Dreifke</i>	

<b>CP67 – Simulação morfológica na caracterização espacial da perda de solo por erosão em cultivo de cana-de-açúcar</b> .....	123
<i>Daniel de Bortoli Teixeira, João Fernandes da Silva Júnior, Marcílio Vieira Martins Filho e Gener Tadeu Pereira</i>	
<b>CP68 – Identificação do período de prevenção contra o vírus sincicial respiratório e ocorrência de bronquiolite viral aguda</b> .....	124
<i>Edilenia Queiroz Pereira, Márcia Lorena Alves dos Santos e Eniuce Menezes de Souza</i>	
<b>CP69 – Application of marginal structural models to irregular longitudinal data</b> .....	124
<i>Edmore Chamapiwa, David Reeves, Evangelos Kontopantelis and Darren Ashcroft</i>	
<b>CP70 – Modelling the rate of carbon dioxide absorption of the grassy species <i>Echinochloa crus-galli</i> by mixing finite distributions</b> .....	125
<i>Ednário Mendonça, Ricardo Olinda, Tiago Oliveira and Thuenne Barros de Oliveira</i>	
<b>CP71 – Calculation of probabilities for daily maximum rainfall in the city of Jatai, Goiás</b> .....	125
<i>Elias Silva de Medeiros, Marina Rodrigues Maestre and Pablo Lourenço Ribeiro de Almeida</i>	
<b>CP72 – Análise quantitativa dos traçados eletrocardiográficos de pacientes do projeto Teleminas Saúde</b> .....	126
<i>Franciele Brito, Aurélia A. de A. Rodrigues, Rodolfo Ribeiro, Elmiro Resende e João Batista Destro Filho</i>	
<b>CP73 – Modelos assimétricos para dados de consumo de vitaminas em idosos</b> .....	127
<i>Giovana Fumes and José Eduardo Corrente</i>	
<b>CP74 – Rede de correlação fenotípica de variáveis de alho</b> .....	127
<i>Hygor Amaral Santana, Divino de Sousa Costa, Wilhan Valasco dos Santos, Brunna Rithielly Rezende e Anderson Rodrigo da Silva</i>	
<b>CP75 – Avaliação da qualidade dos registros neonatais através de linkage e imputação de bancos de dados</b> .....	128
<i>Ianne Rafaella Santos Melo, Jessica da Silva Braga, João Batista Carvalho e Neir Antunes Paes</i>	
<b>CP76 – Equações de volume individual para <i>Toona ciliata</i></b> .....	128
<i>Joyce de Almeida Alves, Geisi Azevedo Silva, Ernani Lopes Possato, Marcos Antônio Isaac Jr e Natalino Calegario</i>	
<b>CP77 – O método simplex aplicado a dados de mistura isotópica em suco de frutas</b> ...	129
<i>Maria Márcia Pereira Sartori, Luciana Balter Francisco, Beatriz Vedovato e Waldemar Gastoni Venturini Filho</i>	
<b>CP78 – Modelos com platô para estimativa do nível de zinco inorgânico na dieta de frangos de corte de 1 a 21 dias de idade</b> .....	129
<i>Maria Márcia Pereira Sartori, Juliana Cristina Ramos Rezende, Antônio Celso Pezzato e José Roberto Sartori</i>	
<b>CP79 – Modelagem linear mista aplicada a dados de expectativa de vida da população</b>	130
<i>Marina Rodrigues Maestre e Elias Silva de Medeiros</i>	

<b>CP80 – Evaluation of allelopathic effect of <i>Senecio</i> sp. in the germination of pearl millet seeds (<i>Pennisetum glaucum</i>) through a log-gamma model</b> . . . . .	130
<i>Matheus Henrique Dal Molin Ribeiro, Vanderlei Aparecido de Lima, Patrícia Carla Giloni de Lima and Flávia Boruch</i>	
<b>CP81 – Seleção de modelos lineares misto na estimativa de quantidade de beta-caseína do leite UAT e comparação entre os softwares SAS e R</b> . . . . .	131
<i>Michele Barbosa, Márcio Balestre e César Gonçalves de Lima</i>	
<b>CP82 – Time-to-event analysis of subcellular exocytotic events of insulin release</b> . . . . .	131
<i>Morten Gram Pedersen, Giuliana Cortese, Nikhil Gandasi and Sebastian Barg</i>	
<b>CP83 – Efeito da interação entre fontes e doses de biomassa carbonizada na produção de mudas de uma espécie florestal na região Amazônica</b> . . . . .	132
<i>Paulo Souza, Gustavo Silva, Carlos Felipe Link, Roni Balduino, Ernandes Moura, Jodean Silva e Fabrício Andrade</i>	
<b>CP84 – Análise de modelos probabilísticos para os dados mensais de chuva no município de Toledo - PR</b> . . . . .	132
<i>Rúbia Wilhelms França e Rosangela Aparecida Botinha Assumpção</i>	

---

## Part X. Contributed Posters 2

---

<b>Bioestatística e Modelos Biométricos 2 [<i>Biostatistics and Biometrical Models 2</i>]</b> . . . . .	135
<b>CP85 – Previsão do número de pousos no aeroporto Santa Maria em Aracaju-SE</b> . . . . .	135
<i>Carlos Silva, Juliana Silva e Suzana Russo</i>	
<b>CP86 – Um método de amostragem georreferenciado para fiscalização, controle e erradicação da doença grenning (<i>Huanglongbing/HLB</i>) da cultura do citros com base na distribuição Beta-Binomial</b> . . . . .	135
<i>Edilenia Queiroz Pereira, Marcos Jardel Henriques, Oilson Alberto Gonzatto Junior, Terezinha Aparecida Guedes, Vanderly Janeiro e Ana Beatriz Tozzo Martins</i>	
<b>CP87 – Otimização na seleção entre famílias de cana-de-açúcar via classificação por K-nearest neighbor</b> . . . . .	136
<i>Édimo Moreira, Luiz Peternelli e Kelly Rosa</i>	
<b>CP88 – Modelagem da variabilidade em experimentos de mistura com variáveis de ruído</b>	136
<i>Edmilson Rodrigues Pinto, Leandro Alves Pereira e Aurélia A. de A. Rodrigues</i>	
<b>CP89 – Análise de trilha em caracteres de genótipos de milho de ciclo precoce</b> . . . . .	137
<i>Jéssica Andiará Kleinpaul, Alberto Cargnelutti Filho, Bruna Mendonça Alves, Giovani Facco, Rosana Marzari Thomasi, Daniela Barbieri Uliana, Cleiton Antonio Wartha, Rafael Vieira Pezzini e Daniela Lixinski Silveira</i>	
<b>CP90 – Análise de agrupamento em genótipos de milho de ciclos precoce e superprecoce</b>	137
<i>Lucas Silva Stefanelo, Alberto Cargnelutti Filho, André Lavezo, Bruna Mendonça Alves, Gabriela Görgen Chaves, Jéssica Andiará Kleinpaul, Cláudia Marques de Bem, Cleiton Antonio Wartha e Daniela Lixinski Silveira</i>	

<b>CP91 – Análise de trilha em caracteres de genótipos de milho de ciclo superprecoce . . .</b>	<b>138</b>
<i>Rosana Marzari Thomasi, Alberto Cargnelutti Filho, Bruna Mendonça Alves, Denison Esequiel Schabarum, Jéssica Andiará Kleinpaul, Cleiton Antonio Wartha, Daniela Barbieri Uliana, Fernanda Martins Simões e Rafael Vieira Pezzini</i>	
<b>CP92 – Improvement of impact assessment through combination of alternative strategies: Empirical results of a training program for rice production in Indonesia . . . . .</b>	<b>139</b>
<i>Samila Sena and Leila D.A.F. Amorim</i>	
<b>CP93 – Aceitabilidade de pão tipo bisnaguinha adquirido na cidade de Araras/SP por meio de mapas de preferência . . . . .</b>	<b>139</b>
<i>Simone Daniela Sartorio, Marta Regina Verruma-Bernardi, Carolina Ibelli Bianco, Jéssica Adriana de Jesus e Ionele do Socorro Ferreira Almeida</i>	
<b>Análise de Sobrevivência [<i>Survival Analysis</i>] . . . . .</b>	<b>141</b>
<b>CP94 – The Kumaraswamy transmuted Fréchet distribution . . . . .</b>	<b>141</b>
<i>Angela Lima da Silva, Giovana Oliveira Silva, Jailson de Araujo Rodrigues and Marcio Claudio Mercês Brito</i>	
<b>CP95 – Modelo de regressão para eventos recorrentes utilizando a distribuição Weibull modificada em análise de sobrevivência . . . . .</b>	<b>141</b>
<i>Jackson Santos da Conceição e Giovana Oliveira Silva</i>	
<b>CP96 – The exponentiated generalized Lindley distribution: Properties and application</b>	<b>142</b>
<i>Jailson A. Rodrigues, Ana Paula Silva and Angela Silva</i>	
<b>CP97 – Análise de sobrevivência aplicada a durabilidade de pneus 11.00R22 . . . . .</b>	<b>142</b>
<i>Pablo Lourenço, Elias Silva, Tiago Almeida e Isaac Pereira</i>	
<b>Estatística Multivariada [<i>Multivariate Statistics</i>] . . . . .</b>	<b>143</b>
<b>CP98 – Analysis with distal outcomes: latent effects on ENADE performance . . . . .</b>	<b>143</b>
<i>Ana Clara Campos, Sheila Regina Pereira, Samila Sena, Lilia Carolina Costa and Leila D.A.F. Amorim</i>	
<b>CP99 – A new approach for testing the genetic covariance . . . . .</b>	<b>143</b>
<i>Anderson Silva, Carlos Tadeu dos Santos Dias, Claudio Souza Junior, Trevor Bailey and Wojtek Krzanowski</i>	
<b>CP100 – Correlação canônica entre caracteres fenológicos, morfológicos e produtivos de cultivares de milho . . . . .</b>	<b>144</b>
<i>André Lavezo, Alberto Cargnelutti Filho, Lucas da Silva Stefanelo, Bruna Mendonça Alves, Jéssica Andiará Kleinpaul, Diego Nicolau Follmann, Gabriela Górgen Chaves, Cleiton Antonio Wartha e Fernanda Martins Simões</i>	
<b>CP101 – LISA cluster maps of incidence rates of occupational accidents in Brazil . . . . .</b>	<b>144</b>
<i>Andre Luis Santiago Maia, Isabelle Galdino de Almeida and Gecynalda Soares da Silva Gomes</i>	
<b>CP102 – Seleção de descritores agronômicos em acessos de tabaco por meio de Técnicas Multivariadas . . . . .</b>	<b>145</b>
<i>Antonio Leandro da Silva Conceição, Carlos Alberto da Silva Ledo, Mauricio dos Santos da Silva, Clailto Carvalho dos Santos a Sandra Domingos João Afonso</i>	

<b>CP103 – Técnicas multivariadas aplicadas a biomarcadores metabólicos e o método Z-scan</b> .....	145
<i>Aubin Elisete, Viviana Giampaoli, Nágila Damaceno, Antonio Figueiredo Neto, Maria Camila Freitas e Milena Barbosa</i>	
<b>CP104 – Análise de componentes principais na identificação dos principais fatores relacionados à satisfação dos alunos: uma análise em uma IES Pública</b> .....	146
<i>Beatriz Rezzieri Marchezini, Lucas Pereira Lopes e Gabriel Rodrigo Gomes Pessanha</i>	
<b>CP105 – A dependência espacial e análise discriminante canônica para a bovinocultura no estado de Mato Grosso</b> .....	147
<i>Camila Ozelame, Kuang Hongyu e Anderson Oliveira</i>	
<b>CP106 – Análise do aquecimento de motores elétricos utilizando a análise de componentes principais Biplot</b> .....	147
<i>Cláudio Cruz da Silva, Aline Xavier Porto Alves, Maria Joseane Cruz da Silva e Sônia Eliane Gonçalves dos Santos</i>	
<b>CP107 – Redução de dimensionalidade e visualização: comparação entre os modelos AMMI Ponderado e AMMI Robusto</b> .....	148
<i>Cristiane N. Pamplona Souza e Paulo Canas Rodrigues</i>	
<b>CP108 – Análise de agrupamento dos municípios do estado de Sergipe em relação à precipitação e região</b> .....	148
<i>Eucymara Nunes e Inajá Francisco Sousa</i>	
<b>CP109 – Análise de fatores no agrupamento de caracteres de acessos de pimentas (Capsicum chinense)</b> .....	149
<i>Giselda M. Pereira, Maicon Nardino, Elisia R. Corrêa, Willian Barros, Ivo R. Silva, Adilson R. Schuelter and Fernando Finger</i>	
<b>CP110 – Classificação no desfecho da doença meningocócica invasiva usando análise discriminante</b> .....	149
<i>Izabela R. C. Oliveira, Daniel F. Ferreira, Eliana T. Masuda, Eliseu A. Waldman e Francisco Chiaravalloti Neto</i>	
<b>CP111 – Componentes independentes, análise de agrupamento e medidas de similaridades no monitoramento da potência elétrica</b> .....	150
<i>Jair Rocha do Prado, Thelma Sáfadi and Joaquim Paulo da Silva</i>	
<b>CP112 – Um estudo de diferentes técnicas de seleção do modelo AMMI</b> .....	150
<i>João Vitor Ribeiro Silva, Lúcio B. de Araújo, Mírian F. C. Araújo and Carlos Tadeu dos Santos Dias</i>	
<b>CP113 – Utilização dos modelos AMMI e GGE Biplot para os dados de interação genótipo por ambiente</b> .....	151
<i>Kuang Hongyu, Fabiane Silva, Lúcio Araújo and Carlos Tadeu dos Santos Dias</i>	
<b>CP114 – Avaliação dos registros de nascimentos das microrregiões do Semiárido Brasileiro utilizando análise fatorial</b> .....	151
<i>Lúgia Maia, Clarissa Cavalcanti, Diego Silva e Neir Paes</i>	

<b>CP115 – Aplicação do modelo multivariado de resposta ao item em um exame vestibular da UFPA</b> .....	152
<i>Manoel Vitor de Souza Veloso, Caio Azevedo e Júlio Bueno</i>	
<b>CP116 – Relações de interdependência da emissão de CO<sub>2</sub> e atributos de solo em áreas de cana-de-açúcar, analisadas por meio de estatística multivariada</b> .....	152
<i>Mara Regina Moitinho, Daniel de Bortoli Teixeira, Alan Rodrigo Panosso e Newton La Scala Jr</i>	
<b>CP117 – Determinantes da mortalidade na infância no Brasil: Uma análise por componentes principais</b> .....	153
<i>Marcela de Marillac Carvalho and Manoel Vitor de Souza Veloso</i>	
<b>CP118 – Identificação da similaridade entre tratamentos em locais distintos com a utilização da análise de múltiplos fatores</b> .....	153
<i>Paulo Cesar Ossani, Tatiane Carvalho Alvarenga e Renato Ribeiro de Lima</i>	
<b>CP119 – Obtenção de limite de acurácia seletiva e coeficiente de variação para classificação de ensaios de cana de açúcar</b> .....	153
<i>Sandra Elisa Fonrozo, Rubens Leite do Canto Braga Júnior e Guilherme Augusto Peres Silva</i>	
<b>CP120 – Análise de correlação canônica da emissão de CO<sub>2</sub> e atributos do solo em área de soja na região de cerrado</b> .....	154
<i>Vivian A. Brancaglioni, Jacqueline Domingues, Maria F. F. Almeida, Alan R. Panosso, Newton La Scala Junior e José S. Govone</i>	
<b>Big Data, Redes Neurais e Estatística Espacial [<i>Big Data, Neural Networks and Spatial Statistics</i>]</b> .....	155
<b>CP121 – Aplicação de máquinas de suporte vetorial no reconhecimento de neuropatias periféricas</b> .....	155
<i>Albaro Paiva, Carenne Ludeña e Tiago Espínola</i>	
<b>CP122 – Accuracy Assessment in Probabilistic Linkage Procedures using Big Data</b> ...	155
<i>André Alves, Samila Sena, Rosemeire Fiaccone, Leila D. A. F. Amorim, George Barbosa, Marcos Barreto and Maurício Lima Barreto</i>	
<b>CP123 – Estudo exploratório da configuração espacial pontual dos alvéolos do pão francês</b> .....	156
<i>Carolina Cristina Bicalho, Raul Antônio Viana Madeira e João Domingos Scalon</i>	
<b>CP124 – Análise Espacial dos Resultados do ENEM 2014 das Escolas do Município de Fortaleza/CE</b> .....	156
<i>Débora Ferreira de Assis e Júlio Francisco Barros Neto</i>	
<b>CP125 – Delineamento de unidades de mapeamento com base na incerteza espacial da suscetibilidade magnética e teor de argila</b> .....	157
<i>Gener Tadeu Pereira, Daniel de Bortoli Teixeira, Diego Silva Siqueira, Vinícius Vasconcellos e José Marques Júnior</i>	
<b>CP126 – Redes neurais aplicadas à classificação de imagens radiográficas de sementes</b>	157
<i>Isabel Leite, Felipe Gonzalez, Thelma Sáfiadi e Maria Laene Carvalho</i>	

<b>CP127 – Análise da autocorrelação espacial dos setores censitários de Rio Claro - SP para os casos de dengue, através do índice de Moran</b> .....	158
<i>Jaqueline Domingues, Maria F. F. Almeida, Vivian Ap. Brancaglioni e José S. Govone</i>	
<b>CP128 – The process of search networks applied to the MDM</b> .....	158
<i>Lilia Carolina Costa</i>	
<b>CP129 – Distribuição espacial dos focos de queimadas no estado de São Paulo no período de 2008 a 2015</b> .....	159
<i>Mayara P.L.S. Sicchieri e Estela C. Carneseca</i>	
<b>CP130 – Funções L homogênea e não homogênea em testes de hipótese de completa aleatoriedade espacial</b> .....	159
<i>Sicleide Barbosa de Sousa, João Domingos Scalon e Kelly Pereira de Lima</i>	
<b>CP131 – Análise da distribuição espacial da resistência do solo à penetração em área experimental</b> .....	160
<i>Wilhan Valasco dos Santos, Divino de Sousa Costa, Brunna Rithielly Rezende, Hygor Amaral Santana e Anderson Rodrigo da Silva</i>	
<b>Geoestatística e estatística não paramétrica [<i>Geostatistics and non-parametric statistics</i>]</b>	161
<b>CP132 – Applied Nonparametric Analysis of covariance: One study of the effect of the bandwidth choosing</b> .....	161
<i>Fabio da Silva</i>	
<b>CP133 – Influence of symmetry of the data in a geostatistical analysis</b> .....	161
<i>Lidiane Rosa, Gerson Santos, Adriana Santos and Italo Ferreira</i>	
<b>CP134 – Estudo da variabilidade da emissão de CO<sub>2</sub> do solo em área de reforma do canal, utilizando-se análises multivariada e geoestatística</b> .....	162
<i>Mara Regina Moitinho, Antonio Sergio Ferraudo, Daniel de Bortoli Teixeira, Fabiana De Souza Cannavan, Siu Mui Tsai, Elton da Silva Bicalho e Newton La Scala Junior</i>	
<b>CP135 – Comparação entre modelo paramétrico e não paramétrico para dados de contagem</b> .....	162
<i>Paulo Henrique Sales Guimarães e Sue Éllen Ester Queiroz</i>	
<b>Séries Temporais, Cópulas e Econometria [<i>Time series, Copulas and Econometrics</i>]</b> ..	164
<b>CP136 – Comparação de métodos de estimação do risco relativo</b> .....	164
<i>Agda Galletti, Mateus Ananias, Rodrigo Lassance, Bernardo Andrade e Joanlise Marco de Leon Andrade</i>	
<b>CP137 – Análise de previsibilidade da velocidade do vento no Nordeste Brasileiro</b> ....	164
<i>Antonio Silva, Leda Santana, Luiz Dore e Tatijana Stosic</i>	
<b>CP138 – Análise da produção de biodiesel no Brasil usando a metodologia de Box e Jenkins</b> .....	165
<i>Bárbara Alves Lopes e Leticia Lima Milani</i>	

<b>CP139 – Análise econométrica da estabilidade do mercado acionário do setor bancário Brasileiro</b> .....	165
<i>Beatriz de Souza Bernadinho, Gabriel Rodrigo Gomes Pessanha, Caio Tonolli Dias Gentile e Lucas Pereira Lopes</i>	
<b>CP140 – Moving block quantile residual Bootstrap in GARMA models: An application for the time series of dengue case count</b> .....	166
<i>Breno Silveira de Andrade, Jacek Léskow and Marinho Gomes de Andrade Filho</i>	
<b>CP141 – Determinantes da concentração de mercado: Uma análise econométrica do setor bancário Brasileiro no período de 1995 a 2015</b> .....	166
<i>Caio Tonolli Dias Gentile, Gabriel Rodrigo Gomes Pessanha, Lucas Pereira Lopes e Beatriz de Souza Bernadinho</i>	
<b>CP142 – Ajuste de modelo de séries temporais para o número de óbitos por causas externas no Brasil</b> .....	167
<i>Camila Moreira Ribeiro e Leticia Lima Milani</i>	
<b>CP143 – Uso de séries temporais na análise do consumo de energia elétrica da Universidade Federal de São João Del-Rei</b> .....	167
<i>Caroline Oliveira Santos, Letícia Lima Milani, Andréa Cristiane Delfino e Ana Lúcia Mateus</i>	
<b>CP144 – Análise do volume do reservatório da usina hidrelétrica de Três Marias usando modelos de séries temporais</b> .....	168
<i>Charles Shalimar Felipe da Silva, Paulo César Moraes Ribeiro, Sidcleide Barbosa de Sousa, Ricardo Vitor Ribeiro dos Santos e Marcelo Silva de Oliveira</i>	
<b>CP145 – Um teste para qualidade de ajuste de Cópulas bivariadas via marginal obtida na diagonal</b> .....	168
<i>Danielle Prado, Lucas Chaves e Devanil Souza</i>	
<b>CP146 – Determinantes do retorno das ações: Um estudo econométrico na Eletrobrás no período de 1994 a 2014</b> .....	169
<i>Ericksen Souza, Gabriel Rodrigo Gomes Pessanha e Luciene Resende Gonçalves</i>	
<b>CP147 – Estudo do Índice de Criminalidade contra a Vida (ICV) em Maputo/Moçambique usando séries temporais</b> .....	169
<i>Joel Jorge Nuvunga, Luciano Antonio de Oliveira, Alessandra Querino da Silva, Carlos Pereira da Silva, Andrezza Kellen Alves Pamplona e Marcio Balestre</i>	
<b>CP148 – Análise estatística no setor financeiro - estudo de um caso de uma instituição paulista</b> .....	170
<i>Kaue Gustavo Reis e Vilma M. Tachibana</i>	
<b>CP149 – Análise da precipitação anual em São João do Tigre no Cariri paraibano utilizando modelos ARMA</b> .....	170
<i>Kerolly Kedma Felix do Nascimento, Fábio Sandro dos Santos, Hérica Santos da Silva, Ikaro Daniel de Carvalho Barreto e Tiago Alessandro Espínola Ferreira</i>	
<b>CP150 – Análise da complexidade climática do município de Areia – PB através do SampEn</b> .....	171

*Leandro dos Santos Silva, Ewerton Matias de Lima, Lázaro de Souto Araújo, Péricles Borges de Farias, José Rodrigo Santos Silva e Eucymara França Nunes Santos*

- CP151 – Análise de séries temporais do número de óbitos infantis de Minas Gerais ...** 171  
*Mariana de Souza Gonçalves e Letícia Lima Milani*
- CP152 – Análise espectral singular aleatorizada na análise de grandes series temporais** 172  
*Petala Gardenia da Silva Estrela Tuy, Paulo Canas Rodrigues and Rahim Mahmoudvand*
- CP153 – Gap-filling in time series by multivariate singular spectrum analysis . . . . .** 172  
*Rahim Mahmoudvand*
- CP154 – Analysis of consumer price index metropolitan region of Belo Horizonte/MG using time series . . . . .** 173  
*Ricardo Vitor Ribeiro dos Santos, Érica Fernanda da Cruz, Thelma Sáfadi, Tamara Aperecida Nogueira dos Anjos, Charles Shalimar Felipe da Silva and Cristian Tiago Mendes*
- CP155 – Projeção da abate, rebanho e peso total de carcaças de aves em Mato Grosso pelo modelo VAR . . . . .** 173  
*Rondiny Carneiro, Edilson Júnior, Anderson Oliveira, Marcelino Pascoa e Juliano Bortolini*
- CP156 – Dynamic analysis for event history data: An example of recurrent event . . . .** 173  
*Rosemeire Fiaccone, Robin Henderson and Fabio da Silva*
- CP157 – Intervenções na evolução temporal do preço da cesta básica de Salvador . . . .** 174  
*Tamara Aperecida Nogueira dos Anjos, Ricardo Vitor Ribeiro dos Santos, Érica Fernanda da Cruz e Thelma Sáfadi*
- Estatística Computacional [*Computational Statistics*] . . . . .** 175
- CP158 – Redes neurais artificiais na identificação de padrões de ocorrência de Fusões e Aquisições (F&A) nas instituições financeiras Brasileiras . . . . .** 175  
*Lucas Pereira Lopes, Gabriel Rodrigo Gomes Pessanha, Beatriz de Souza Bernadinho e Caio Tonolli Dias Gentile*
- CP159 – Criação de redes neurais no Statistica: apresentação dos principais recursos do software . . . . .** 175  
*Mariana Nolde Pacheco e Lisiane Priscila Roldão Selau*
- CP160 – Predição da rentabilidade: Uma aplicação da metodologia de redes neurais artificiais em empresas brasileiras de capitalização . . . . .** 176  
*Richard Hernani Pereira, Leandro Ferreira, Gabriel Rodrigo Gomes Pessanha e Lucas Pereira Lopes*
- Estatística Bayesiana [*Bayesian Statistics*] . . . . .** 177
- CP161 – Bayesian binary overdispersed longitudinal model via marginalized likelihood** 177  
*Afranio M.C. Vieira and Everton B. Rocha*
- CP162 – Avaliação de híbridos de milho utilizando o modelo AMMI com efeitos fixos e o AMMI-Bayesiano . . . . .** 177  
*Alessandra Querino da Silva, Luciano Antonio de Oliveira, Carlos Pereira da Silva, Joel Jorge Nwunga, Andrezza Kellen Alves Pamplona e Marcio Balestre*

<b>CP163 – Análise bayesiana de delineamentos triangulares para experimentos sensoriais</b>	178
<i>Camilla Barroso e Júlio Bueno Filho</i>	
<b>CP164 – Uma aplicação do modelo AMMI-bayesiano-shirincage na análise de estabilidade e adaptabilidade genotípica</b>	178
<i>Carlos Pereira da Silva, Luciano Antonio de Oliveira, Andrezza Kellen Alves Pamplona, Joel Jorge Nuvunga, Alessandra Querino da Silva e Marcio Balestre</i>	
<b>CP165 – Cálculo das condicionais a posteriori para a análise bayesiana do modelo fatorial dinâmico para um vetor de séries temporais utilizando distribuição t multivariada</b>	179
<i>Larissa Ribeiro de Andrade, Daniel Furtado Ferreira e Thelma Sáfiadi</i>	
<b>CP166 – Ajuste de modelos não lineares com enfoque bayesiano aos dados de comprimento do fruto da pereira Asiática</b>	179
<i>Rick Mangueira, Taciana Savian e Erasnilson Vieira Camilo</i>	
<b>CP167 – Método Bayesiano para estimação de efeitos latentes com desfechos distais</b>	180
<i>Silvia Regina Moraes, Lilia Carolina Costa e Leila D.A.F. Amorim</i>	
<b>CP168 – Construção de prioris utilizando o conceito de máxima entropia: aplicação a dados genéticos</b>	180
<i>Tais Alvarenga, Ernandes G. Moura e Marcio Balestre</i>	
<b>Index</b>	181



Part I

**Introduction**



## Welcome to the 61st RBras 2016

On the behalf of the Local Organizing Committee we are pleased to welcome you to Salvador. This is the 61st edition of our regional meeting and the first time we have the pleasure to organize it in the beautiful city of Salvador.

The 61st RBras 2016 is sponsored by the Brazilian Region of the International Biometric Society (RBras) and the International Biometric Society (IBS) itself. Its purpose is to bring together researchers that use statistical methods for decision making, to share and discuss ways to improve the access to knowledge, and promote interdisciplinary collaborations.

The scientific program includes three keynote speakers who are the presidents of three of the four biggest worldwide statistical societies, nine short courses, nine invited paper sessions, two round tables, five contributed paper sessions and two poster sessions, with a total of almost 250 contributions. The theme of the 61st RBras 2016 is “Biometry and Biostatistics in the Era of Big Data” which is one of the current hot topics in statistical sciences and data analytics.

The 61st RBras 2016 will offer opportunities to meet each other, to share scientific and professional experiences, and to promote new collaborations. This meeting will cover many of the research areas of Biometry, Biostatistics and Applied Statistics.

Students and young statisticians (up to 5 years after their last academic degree) attending this conference have the possibility to compete for several awards: “Best PhD thesis”, “Best RBras 2016 Paper Award”, “Best RBras 2016 Poster by a Graduation Student”, “Best RBras 2016 Poster by a MSc Student” and “Best RBras 2016 Poster by a PhD Student”.

The program also includes social events which will allow the participants to get to know each other and to experience the culture and history of Bahia, in addition to the taste of the well known Bahia’s cuisine and hospitality.

The organizers would like to thank to all institutions that have provided financial and other support to make this organization possible. Many thanks to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Timberlake Brasil for the financial support which made possible the organization of the 61st RBras 2016. We also thank Salvador Destination for intermediating the negotiations with the conference venue, the Instituto Anísio Teixeira for providing accommodation for students, SALTUR for providing the cultural performances, and our colleagues from the RBras Board of Directors and to the International Biometric Society for their support. Last, but not the least, we thank the keynote speakers, the lecturers of the short courses, the speakers and discussants in the invited paper sessions and round tables, the speakers in the contributed sessions and the poster presenters, for their contribution to make a great scientific program. Thank you all for being here!

We wish you an enjoyable stay and a good time in Salvador!

On the behalf of the Local Organizing Committee,

Paulo Canas Rodrigues  
Chair of the LOC of the 61st RBras 2016



## Comissão Organizadora Local

[*Local Organizing Committee*]

- Paulo Canas Rodrigues (Coordenador), Universidade Federal da Bahia, Brasil
- Rodrigo Bulhões, Universidade Federal da Bahia, Brasil
- Jalmar Carrasco, Universidade Federal da Bahia, Brasil
- Aldo Medina, Universidade Federal de Pernambuco, Brasil
- Arthur Farias, Universidade Federal da Bahia, Brasil
- Elisabete Sampaio, Universidade Federal da Bahia, Brasil
- Cristiane Souza, Universidade Federal da Bahia, Brasil

## Comissão Organizadora Local – Colaboradores

[*Local Organizing Committee – Collaborators*]

- Jairan Dionizio, Universidade Federal da Bahia, Brasil
- Alexandro Teles de Oliveira, Universidade Federal da Bahia, Brasil
- Daisy Santana, Universidade Federal da Bahia, Brasil

## Comissão Científica

[*Scientific Program Committee*]

- John Hinde (Chair), National University of Ireland Galway, Ireland
- Paulo Canas Rodrigues (Chair), Federal University of Bahia, Brazil
- Aparecida Souza, São Paulo State University, Brazil
- Bhramar Mukherjee, University of Michigan, USA
- Clarice Garcia Borges Demétrio, University of São Paulo, Brazil
- Emmanuel Lesaffre, Katholieke Universiteit Leuven, Belgium
- Fernando Oliveira, Federal University of Ouro Preto, Brazil
- Geert Molenberghs, Universiteit Hasselt and Katholieke Universiteit Leuven, Belgium
- Hans-Peter Piepho, Hohenheim University, Germany
- Renato Ribeiro de Lima, Federal University of Lavras, Brazil
- Ricardo Olinda, State University of Paraíba, Brazil
- Vicente Núñez-Antón, University of the Basque Country, Spain

## **Diretoria da RBras (2014–2016)**

*[Board of Directors of RBras (2014–2016)]*

- Presidente: Paulo Justiniano Ribeiro Junior (UFPR, PR)
- Vice-Presidente: Joel Augusto Muniz (UFLA, MG)
- 1º Secretário: Afrânio Márcio Corrêa Vieira (UFSCar, SP)
- 2º Secretário: Roseli Aparecida Leandro (ESALQ/USP, SP)
- 1º Tesoureiro: Luzia Aparecida Trinca (IBB/UNESP, SP)
- 2º Tesoureiro: Euclides Braga Malheiros (FCAV/UNESP, SP)

## **Conselho Diretor da RBras**

*[Council of RBras]*

- Cecília Candolo (UFSCar, SP; 2013–2016)
- Ricardo Alves de Olinda (UEPB, PB; 2013–2016)
- Vera Lúcia Damasceno Tomazella (UFSCar, SP; 2014–2017)
- Luiz Alexandre Peternelli (UFV, MG; 2014–2017)
- Alessandro Dal'Col Lucio (UFSM, RS; 2015–2018)
- Aparecida D.P. Souza (UNESP-PP, SP; 2015–2018)

## **Membros da RBras no Representative Council da IBS**

*[RBras Representative Council in IBS]*

- Julia Maria Pavan Soler (IME/USP, SP; 2013–2017)
- Renato Martins Assunção (UFMG, MG; 2011–2017)
- Luzia Aparecida Trinca (UNESP, Botucatu, SP, 2015–2019)

Part II

**Scientific Program**



## Programação Científica (Scientific Program)

Time	Monday, 23/05/2016	Tuesday, 24/05/2016	Wednesday, 25/05/2016
08:00–08:30	Registration	<b>MC2:</b> {Conde}	<b>MC4:</b> {Morro de São Paulo} <b>MC5:</b> {Praia do Forte} <b>MC8:</b> {Arraial D'Ajuda}
08:30–09:00		<b>MC4:</b> {Morro de São Paulo}	
09:00–09:30	<b>Opening Ceremony</b> {Salvador II}	<b>MC5:</b> {Praia do Forte}	
09:30–10:00	<b>KS1:</b> Elizabeth Thompson	<b>MC8:</b> {Salvador II}	
10:00–10:30	{Salvador II}	<b>KS2:</b> Peter Diggle	<b>Coffee Break</b>
10:30–11:00	<b>Coffee Break</b>	{Salvador II}	
11:00–11:30	<b>IPS3:</b> {Salvador II}	<b>Coffee Break</b>	<b>IPS8:</b> {Praia do Forte}
11:30–12:00	<b>IPS7:</b> {Arraial D'Ajuda}	<b>Poster Session 1</b> {Porto Seguro}	<b>MC1:</b> {Arraial D'Ajuda}
11:30–12:00	<b>IPS9:</b> {Conde}		<b>MC6:</b> {Morro de São Paulo}
12:00–12:30	<b>MC1:</b> {Praia do Forte}		
12:30–13:00	<b>MC6:</b> {Morro de São Paulo}		<b>KS3:</b> Pedro Silva {Morro de São Paulo}
13:00–13:30	<b>Lunch</b>	<b>Lunch</b>	<b>Closing Ceremony</b> {Morro de São Paulo}
13:30–14:00			
14:00–14:30	<b>IPS2:</b> {Salvador II}	<b>IPS4:</b> {Salvador II}	
14:30–15:00	<b>IPS5:</b> {Arraial D'Ajuda}	<b>MC3:</b> {Morro de São Paulo}	
15:00–15:30	<b>MC3:</b> {Morro de São Paulo}	<b>MC7:</b> {Conde}	
15:30–16:00	<b>MC7:</b> {Conde}	<b>MC9:</b> {Praia do Forte}	
15:30–16:00	<b>MC9:</b> {Praia do Forte}		
16:00–16:30	<b>Coffee Break</b>	<b>Coffee Break</b>	
16:30–17:00	<b>IPS6:</b> {Salvador II}	<b>General Assembly of RBras</b> {Salvador II}	
17:00–17:30	<b>CPS1:</b> {Arraial D'Ajuda}		
17:00–17:30	<b>CPS2:</b> {Praia do Forte}		
17:30–18:00	<b>CPS5:</b> {Morro de São Paulo}		
17:30–18:00	<b>RT2:</b> {Conde}		
18:00–18:30	<b>RT1:</b> {Salvador II}	<b>Poster Session 2</b> {Porto Seguro}	
18:30–19:00	<b>CPS3:</b> {Morro de São Paulo}		
18:30–19:00	<b>CPS4:</b> {Arraial D'Ajuda}		
19:00–19:30	<b>CTD:</b> {Praia do Forte}		
19:45–21:15	<b>Cocktail</b>		

Invited Paper Sessions	Short Courses
<b>IPS2:</b> Twenty years of P-splines, and the future is bright <b>IPS3:</b> Modelos Mistos e Aplicações <b>IPS4:</b> Robust Statistics in the 21st Century: New developments and applications <b>IPS5:</b> Innovative statistical methods with applications in medicine and biology <b>IPS6:</b> New approaches for big and complex data in biostatistics <b>IPS7:</b> Estatística Experimental <b>IPS8:</b> Projetos de divulgação e educação estatística <b>IPS9:</b> Sessão EMBRAPA	<b>MC1:</b> Tools for missing data and sensitivity analysis <b>MC2:</b> An introduction to modern statistical methods for analysis of gene-environment interactions <b>MC3:</b> Modelos de Regressão para Dados de Contagem com o R <b>MC4:</b> Como você compara seus modelos? Quais os critérios você conhece? Saiba quais são suas opções e conheça a medida L <b>MC5:</b> Pesquisa reproduzível com o R: de documentos dinâmicos à pacotes <b>MC6:</b> Geoestatística para análise e modelagem de dados biométricos <b>MC7:</b> Diagrama de Hasse no planejamento de experimentos e na análise de dados <b>MC8:</b> Análise de dados genéticos: Um problema de Big Data a cada novo paciente <b>MC9:</b> Redução de dimensionalidade – Aplicações à genômica
Round Tables	Contributed Paper Sessions
<b>RT1:</b> Meet the Presidents <b>RT2:</b> Tecnologia e estatística aplicadas a políticas de saúde  <b>CTD:</b> Concurso de Tese de Doutorado	<b>CPS1:</b> Recent advances in survival analysis <b>CPS2:</b> Recent advances in mixed models and applications <b>CPS3:</b> Applications in biometry and biostatistics <b>CPS4:</b> Regression and spatio-temporal modelling <b>CPS5:</b> Computer intensive methods and software development



Part III

**Social Program**



# Cerimônia de Abertura (*Opening Ceremony*)

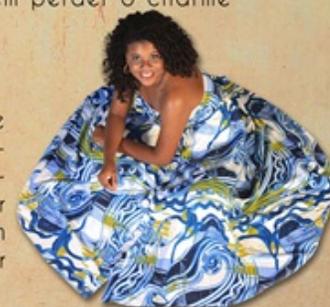
## CLAUDYA COSTTA

Reunindo talento, experiência e carisma a cantora e compositora, Claudya Costa já é destaque no cenário musical baiano e surpreende o público com a exuberância cênica dos seus shows que traçam um panorama de diversidade rítmica, absorvendo vários estilos, fruto do hibridismo de ritmos característicos do Brasil, inserindo-se na revitalizada onda do samba e outras sonoridades que fazem parte da formação musical brasileira mestiça e rica.

Energia, Ousadia e Atitude são próprios da cantora. Seus shows são surpreendentes e sedutores, marcados pela elegância, leveza, refinamento e interatividade. Um samba que engloba as mais diversas criações artísticas da contemporaneidade, sem perder o charme da tradição e o ar do clássico numa inovação pra lá de luxuosa.

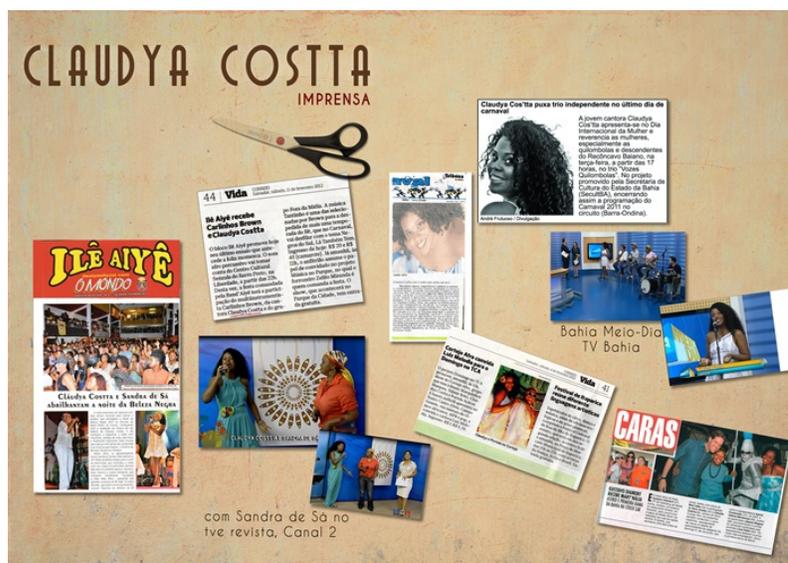
### SHOW PLURALIDADE SINGULAR

Neste show, Claudya traz seu universo pessoal, pleno de referências que dialogam com o ouvinte, variedade rítmica e arranjos simples e sofisticados. Uma conversa ao pé de ouvido com o inconsciente coletivo, filtrando influências e temperando estilos e autores. Pluralidade Singular transita entre o samba e todo universo da Música Popular Brasileira com influências do jazz, do blues, do funk, do Soul, além de interpretar canções autorais e recriar clássicos do nosso gênero.



O Show é mesclado, onde Claudya dá voz aos seus contemporâneos e provoca emoção, balançando e animando o público, mas também faz pensar, aguça o senso crítico e vislumbra a cultura como conquista da cidadania, valorizando a diversidade.

Claudya Costa estará acompanhada de uma banda que é tão criativa e original quanto ela, que hipnotiza os presentes, transferindo-os para um universo particular.



## Coquetel (Cocktail)

# Orquestra de Pandeiros

De Lauro de Freitas / BA



### Release

A Orquestra de Pandeiros, traz para as crianças e adolescentes da cidade de Lauro de Freitas a arte da música através do instrumento “Pandeiro”.

Com o conceito de “Bateria ambulante”, ensina aos alunos e participantes a tocar o instrumento, interagindo em grupo, preparando os musicais da Orquestra, trabalhando a Percepção musical, para fazer apresentações profissionais, passando a existir no cenário cultural.

O Projeto foi criado por “Luciano Pinto”, fundador da “A TAPA Orquestra de Pandeiros e Tonéis”, músico formado em Licenciatura em música pela Universidade Católica de Salvador, IMUCSAL, Diretor Musical e Tecladista da Artista Claudia Leite. Desde o ano de 2001 a “A TAPA”, vem fazendo oficinas de pandeiro em parceria com a Contemporânea.

#### Profissionais envolvidos:

Luciano Pinto - Professor de Música e Diretor Musical da Orquestra.  
 Roberto Aranha - Instrutor e Músico  
 André Buzinga - Instrutor e Músico  
 Lucas de Gal - Instrutor e Músico



Part IV

**Keynote Lectures**



## KS1: Using dependence among individuals and across genomes in the genetic mapping of quantitative traits

Elizabeth A. Thompson<sup>1,2</sup>

<sup>1</sup> University of Washington, USA

<sup>2</sup> President of the International Biometrical Society

**Email:** eathomp@u.washington.edu

### Abstract

The relatedness of individuals is reflected in the close similarity of segments of their genomes that are descended from shared common ancestors. The descent of genome segments also leads to population-level dependence in DNA observed at contiguous genome locations. Modelling this dependence in the DNA both among individuals and across genome locations is key to using modern genomic data in the mapping the locations of DNA that contribute to a quantitative trait. Traditionally, genetic marker data were sparse, and defined pedigree relationships were used to infer the shared descent (IBD) leading to correlations among relatives. With modern genetic data, this pedigree prior is unnecessary, and both local and genome-wide IBD may be estimated from genetic data alone. However, even pairwise local IBD is not easily estimated, and uncertainties and biases in estimation can impact genetic mapping using variance component models. We present current methods for model-based pairwise IBD estimation across the genome, and show how improved IBD estimation can lead to more accurate and powerful approaches to the estimation of genetic mapping likelihoods.

## KS2: Model-based geostatistics for prevalence mapping in low-resource settings

Peter J. Diggle<sup>1,2</sup>

<sup>1</sup> CHICAS, Lancaster Medical School, Lancaster University

<sup>2</sup> President of the Royal Statistical Society

**Email:** p.diggle@lancaster.ac.uk

### Abstract

In low-resource settings, prevalence mapping relies on empirical prevalence data from a finite, often spatially sparse, set of surveys of communities within the region of interest, possibly supplemented by remotely sensed images that can act as proxies for environmental risk factors. A standard geostatistical model for data of this kind is a generalized linear mixed model with logistic link, binomial error distribution and a Gaussian spatial process as a stochastic component of the linear predictor.

In this talk, I will first review statistical methods and software associated with this standard model, then consider several methodological extensions whose development has been motivated by the requirements of specific applications including river-blindness mapping Africa-wide.

## **KS3: Statistical thinking and methodology – A pillar for quality also in the Big Data Era**

**Pedro Luis do Nascimento Silva<sup>1,2</sup>**

<sup>1</sup> ENCE/IBGE, Brazil

<sup>2</sup> President of the International Statistical Institute

**Email:** pedronsilva@gmail.com

### **Abstract**

We live in an era of unprecedented data availability and accessibility, and where 'Big Data' emerged as a 'new paradigm' for those seeking to 'learn from data'. Yet there are still substantial data gaps for many topics or areas of interest. In addition, much of the available data lack the quality needed for its safe use in various applications.

I will review the many ways in which Statistical thinking and methods are used to obtain, process, combine, analyse and disseminate data in the field of Official Statistics. Then I will discuss how statistical methodology is key to achieve specified levels of inferential accuracy, and will link this to a broader view of quality, where the goal is to obtain current, relevant, accurate and cost-effective data.

Despite having evolved in a scenario where it was crucial to learn from 'little data', I contend that Statistical Thinking and Methodology are still essential to guide the 'art and science' of learning from 'Big Data'.



Part V

**Short Course**



## MC1: Tools for missing data and sensitivity analysis

Geert Molenberghs<sup>1,2</sup>

<sup>1</sup> I-BioStat, Hasselt University, Diepenbeek, Belgium

<sup>2</sup> I-BioStat, KU Leuven, Belgium

**Email:** geert.molenberghs@uhasselt.be

**Abstract:** A brief review is given of longitudinal methodology, centering on the linear mixed model and generalized estimating equations. A framework to handle incomplete data is sketched. This includes: missing data patterns, missing data mechanisms, modeling frameworks, and inferential frameworks. It is argued why historic but overly simple methods should generally not be used. Standard methods under missingness at random are discussed, in particular direct (or ignorable) likelihood, weighted estimating equations, and multiple imputations. Aspects of sensitivity analysis are discussed. [4 hours]

---

## MC2: An introduction to modern statistical methods for analysis of gene-environment interactions

Bhramar Mukherjee<sup>1</sup>

<sup>1</sup> University of Michigan, USA

**Email:** bhramar@umich.edu

**Abstract:** In this two hour short course we will discuss analysis of gene-environment interaction in case-control and longitudinal studies. We will start with a general discussion regarding definition of statistical interaction versus biological interaction and then focus specifically on characterization of gene-environment interaction. Various methods for discovery of gene-environment interaction in genomewide studies including case-control, case-only, empirical Bayes, two-step and hybrid methods will be discussed. Methods will be compared in terms of simulation studies and illustrated with data examples. We will then discuss how to capture the temporal dynamics of gene-environment interaction in longitudinal studies with time-varying exposure. [2 hours]

## MC3: Modelos de regressão para dados de contagem com o R

Walmes Marques Zeviani<sup>1</sup>, Eduardo Elias Ribeiro Junior<sup>1</sup> e Cesar Augusto Taconeli<sup>1</sup>

<sup>1</sup> Universidade Federal de Pernambuco, Brasil

**Email:** walmes@ufpr.br

**Resumo:** Dados de contagens configuram variáveis aleatórias que assumem valores inteiros não negativos, sendo presentes em praticamente todas as áreas da ciência. No entanto, métodos de inferência, e, particularmente, modelos de regressão para dados de contagens estão bem aquém em relação à quantidade disponível para a análise de dados contínuos. Nesse contexto, pretendemos neste minicurso apresentar diferentes modelos de regressão para análise de dados de contagens, discutir suas principais propriedades e ilustrar suas aplicações em dados reais por meio do software R. Dentre os modelos considerados estão o modelo Poisson, algumas extensões para acomodação de super (ou sub) dispersão (Quasi-Poisson, COM-Poisson, Count-Gamma, Binomial-Negativo), modelos para dados inflacionados de zeros e com a inclusão de efeitos aleatórios. [4 horas]

---

## MC4: Como você compara seus modelos? Quais os critérios você conhece? Saiba quais são suas opções e conheça a medida L

Elayne Penha Veiga<sup>1</sup> e Mario Javier Ferrua Vivanco<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, Brasil

**Email:** elayne@posgrad.ufla.br

**Resumo:** Serão apresentados os critérios mais utilizados, tais como AIC (Akaike, 1974), AICc (Sugiura, 1978) e BIC (Schwarz, 1978), bem como uma opção, ainda pouco difundida, baseada no preditivismo: a Medida L (Ibrahim e Laud, 1994), cuja preocupação é a predição de valores advindos de um mesmo experimento ou de experimentos semelhantes usando de conceitos como Densidade Preditiva em sua definição, e com este, comparando o que é predito ao que é observado para fazer a escolha entre modelos. Serão apresentados também, resultados de trabalhos (Emiliano et al, 2013) quanto à performance destes critérios em diferentes modelos. Ainda, exemplos de comparações de modelos (no software R) com os diferentes critérios serão apresentados. [4 horas]

## MC5: Pesquisa reproduzível com o R: de documentos dinâmicos a pacotes

Fernando Mayer<sup>1</sup> e Walmes Zeviani<sup>1</sup>

<sup>1</sup> Universidade Federal de Pernambuco, Brasil

**Email:** fernandomayer@gmail.com

**Resumo:** Pesquisa reproduzível é a ideia geral onde análise de dados, e de maneira mais geral, descobertas científicas, devem ser publicadas contendo os dados e o código de análise, para que outras pessoas possam verificar os resultados e até mesmo continuar a construir ideias sobre eles. A necessidade da reprodutibilidade vem crescendo drasticamente, ao mesmo tempo que as análises de dados estão se tornando mais complexas, envolvendo grandes bases de dados e alto processamento computacional. A reprodutibilidade torna um resultado de uma pesquisa mais útil para qualquer pessoa, pois os dados e o código que realmente geraram os resultados estão disponíveis. Este curso será focado nas ferramentas para análise estatística documentada, que permitem que cientistas publiquem suas análises em um único documento que irá permitir que outros cientistas executem a mesma análise e obtenham os mesmos resultados. O objetivo é fornecer os conceitos e ferramentas por trás da comunicação de análises de dados modernas, de maneira reproduzível. A importância de reprodutibilidade na ciência é altamente reconhecida hoje em dia, mas ainda não é largamente praticada como deveria ser. Um dos motivos é que muitos cientistas não adotaram ainda as ferramentas necessárias para a pesquisa reproduzível. Neste curso serão discutidos os princípios gerais para a pesquisa reproduzível, mas o foco será prioritariamente no uso das ferramentas relevantes. Particularmente, será demonstrado como gerar relatórios dinâmicos com o R, e de maneira mais aprofundada, como criar pacotes do R contendo dados, análises e funções, com a intenção de facilitar a distribuição de comunicações científicas. Para isso, serão utilizados alguns pacotes modernos e auxiliares do R, como ‘knitr’, ‘rmarkdown’, ‘devtools’, e ‘roxygen2’. A intenção é que os atendentes do curso saiam preparados para utilizar essas ferramentas em suas próprias pesquisas, ajudando a ampliar a forma como se faz pesquisa reproduzível atualmente. [4 horas]

## MC6: Geoestatística para análise e modelagem de dados biométricos

Mário Sebastiao Tuzine<sup>1</sup>, Anderson Pedro Batista<sup>2</sup> e José Márcio de Mello<sup>2</sup>

<sup>1</sup> Instituto Superior Politécnico De Gaza, Moçambique

<sup>2</sup> Universidade Federal de Lavras, Brasil

**Email:** tuzine@yahoo.com.br

**Resumo:** Um dos princípios que a teoria de amostragem clássica se baseia é a “casualização”, cabe a este princípio neutralizar os efeitos da correlação espacial. Os objetivos deste curso são de apresentar os principais usos de técnicas geoestatísticas para análise e modelagem de informações obtidas em levantamentos biométricos, através da espacialização dos mensuradas em campo. O conteúdo proposto versa sobre conceitos básicos de geoestatística para análise e modelagem de dados biométricos, dando ênfase a levantamentos agrônômicos, florestais, zootécnicos e saúde. Os temas abordados são: Análise exploratória de dados ; Análise gráfica de Dados; Correlação entre variáveis; Avaliação da dependência espacial, Modelagem variográfica, Krigagem e Inferência com base na estrutura de dependência espacial. Toda conceituação estatística e o uso de técnicas geoestatística, dará ênfase a problemas biometrico com recurso ao pacote estatístico R extensão geoR. [4 horas]

---

## MC7: Diagrama de Hasse no planejamento de experimentos e na análise de dados

Clarice G.B. Demétrio<sup>1</sup> e Renata Alcarde Sermarini<sup>1</sup>

<sup>1</sup> ESALQ, Universidade de São Paulo, Brasil

**Email:** clarice.demetrio@usp.br

**Resumo:** O interesse em um considerável número de tratamentos somado a diversas estruturas de controle local pode dar origem a delineamentos complexos, dificultando o planejamento do experimento, bem como sua análise. Nesse ponto, vê-se a necessidade de uma ferramenta que proporcione uma melhor compreensão do experimento, facilitando seu procedimento como um todo. Uma ferramenta que se encaixa em tais qualidades recebe o nome de diagrama de Hasse. Trata-se da representação gráfica das estruturas dos tratamentos e das parcelas, que podem ser construídas por meio de conversas com o pesquisador ou simplesmente da leitura do problema. Ainda, o diagrama auxilia na obtenção do número de graus de liberdade de cada fator, tendo como pressuposto a ortogonalidade entre os fatores, auxilia também na obtenção das somas de quadrados, da razão apropriada para o teste F e do modelo linear estatístico. [4 horas]

## MC8: Análise de dados genéticos: Um problema de Big Data a cada novo paciente

Marcus Nunes<sup>1</sup>

<sup>1</sup> Universidade Federal do Rio Grande do Norte, Brasil

**Email:** marcus.nunes@gmail.com

**Resumo:** Ferramentas de sequenciamento genético tem se tornados mais e mais baratas com o passar do tempo. Devido a isto, experimentos genéticos tem ficado mais comuns. A quantidade de dados gerada por cada sujeito destes experimentos é imensa, fazendo com que um conjunto de problemas de Big Data seja gerado a cada novo experimento. Esta característica dos experimentos genéticos cria novas oportunidades de contribuição para quem trabalha com estatística. Desta forma, a área necessita de pessoal especializado e com conhecimento multidisciplinar para que estes dados sejam analisados corretamente. Neste curso veremos alguns dos principais desafios estatísticos desta área, como a não-normalidade dos dados e o problema das comparações múltiplas. Apresentaremos algumas das características das tecnologias de sequenciamento genético disponíveis no mercado. Também veremos como delinear experimentos nesta área e quais modelos matemáticos utilizar para realizar testes de diferenciação de expressão gênica. Por fim, veremos como resolver problemas genéticos na prática, utilizando o R e outros programas de computador específicos para este tipo de análise. [4 horas]

---

## MC9: Redução de dimensionalidade – Aplicações à genômica

Júlia M. Pavan Soler<sup>1</sup> e Adèle H. Ribeiro<sup>1</sup>

<sup>1</sup> IME, Universidade de São Paulo, Brasil

**Email:** pavan@ime.usp.br

**Resumo:** Muitos problemas de redução de dimensionalidade em Estatística têm sido solucionados por meio de métodos baseados na decomposição em valores singulares de matrizes, o caso mais comum sendo a Análise de Componentes Principais. Contudo, em espaços restritos a  $n \ll p$  (centenas de milhares de variáveis genéticas observadas em um número reduzido de unidades amostrais), bem como em espaços estruturados envolvendo amostras dependentes e/ou amostras pareadas, soluções clássicas não se aplicam. Além disso, o problema de redução de dimensionalidade aparece na teoria de inferência causal. Os modelos gráficos probabilísticos direcionados, que representam independências e dependências condicionais entre variáveis a partir de grafos direcionados, têm sido particularmente usados para a construção de redes causais de genótipos-fenótipos em estudos observacionais. O aprendizado dessa estrutura causal se baseia em princípios como d-separação e aleatorização Mendeliana e em variáveis instrumentais. Neste minicurso trataremos de situações como estas trazendo motivações da área de mapeamento genético. Dados serão gerados e analisados usando os recursos do R. [4 horas]



Part VI

**Invited Paper Sessions**



## **IPS2: Twenty years of P-splines, and the future is bright**

### **Organizer and Chair:**

*María Xosé Rodríguez-Álvarez, University of Vigo, Spain*

**Email:** *m.xrodriguez@uvigo.es*

### **IPS2.1: The analysis of Big Data using P-splines and sparse matrix algebra**

**Martin Boer<sup>1</sup>**

<sup>1</sup> Wageningen University & Research, The Netherlands

**Email:** martin.boer@wur.nl

**Abstract:** P-splines are computationally very attractive, because the corresponding matrices are sparse. However, the sparsity of the P-splines matrices is not always fully exploited, which becomes quite important in the era of Big Data. Several examples of how the efficiency of P-splines models can be further improved by using sparse matrix algebra will be presented.

First, the B-spline basis can be efficiently calculated and stored in a sparse format, which is quite important for big data sets. Second, a fast algorithm to calculate the effective dimension, a measure for the complexity of a P-splines model, will be presented. In standard expressions for the effective dimension the inverse of a sparse matrix is calculated, which can be quite inefficient in terms of computer time and memory. The fast and sparse method to calculate the effective dimension is based on automatic differentiation of the Cholesky algorithm. Finally, as a last example how efficiency can be improved, a sparse transformation of P-splines to mixed models will be presented. The computation time for the new transformation is linear in the number of B-splines, while computation time scales cubically for existing transformations.

The sparse methods will be illustrated with an example from plant breeding. In high-throughput phenotyping experiments typically hundreds of genotypes are measured over a long time range, and with thousands of time points per genotype, resulting in very big data sets.

---

### **IPS2.2: Spatial and spatio-temporal modelling of aggregated Counts with applications in epidemiology: Methods and tools**

**Dae-Jin Lee<sup>1</sup>**

<sup>1</sup> Basque Center for Applied Mathematics, Spain

**Email:** dlee@bcmath.org

**Abstract:** Mortality data provide valuable information for the study of the spatial distribution of mortality risk, in disciplines such as spatial epidemiology and public health. However, they are frequently available in an aggregated form over irregular geographical units, hindering the visualization of the un-

derlying mortality risk. Also, it can be of interest to obtain mortality risk estimates on a finer spatial resolution, such that they can be linked to potential risk factors that are usually measured in a different spatial resolution.

In this talk, we propose the use of the penalized composite link model and its mixed model representation. This model considers the nature of mortality rates by incorporating the population size at the finest resolution, and allows the creation of mortality maps at a finer scale, thus reducing the visual bias resulting from the spatial aggregation within original units. We also extend the composite link model to the spatio-temporal case where counts are aggregated spatially and temporally. We illustrate our proposal with several data sets taking advantage of new algorithms and tools in order to facilitate the visualization of very big data sets across space and time.

---

### IPS2.3: Spatial models for field trials using P-splines

María Xosé Rodríguez-Álvarez<sup>1</sup>

<sup>1</sup> University of Vigo, Spain

**Email:** mxrodriguez@uvigo.es

**Abstract:** We present models for large complicated field trials in which the spatial component consists of tensor product P-splines. They allow modelling the random spatial variation typically presented in a field experiment, providing, at the same time, an explicit estimate of the spatial random field. On top of this spatial field, effects of genotypes, blocks, replicates, and/or other sources of spatial variation are described by a mixed model in a standard way. Our proposal allows expressing the generalized heritability on the basis of effective dimensions. This equivalence brings new insights to the interpretation of both quantities, providing a common definition of heritability that can be used for a broad range of statistical models for field trials analysis. We show that this result can be generalized to all random effects in a mixed model. This suggests moving the interpretation of the results obtained by fitting a mixed model from estimated variances components to estimated effective dimensions. Our approach has several other advantages over other mixed models for field trial analysis. First, it is very robust, as testing on many real field trials have shown. A second advantage is that model selection is simplified in comparison with other approaches. Furthermore, the method can easily deal with irregular field layouts, and local patches with special behaviour become easily visible. All these advantages are illustrated with large-scale field trial experiments, and an R-package, known as SpATS (Spatial Analysis of Trials using Splines) is presented.

## IPS3: Modelos Mistos e Aplicações

### Organizer and Chair:

*Luiza Aparecida Trinca, Universidade Estadual Paulista, Botucatu, Brasil*

**Email:** *ltrinca@ibb.unesp.br*

### **IPS3.1: Modelos mistos potencialidades e desafios nas ciências biológicas**

**Viviana Giampaoli<sup>1</sup>**

<sup>1</sup> IME, Universidade de São Paulo, Brasil

**Email:** *vivig@ime.usp.br*

**Resumo:** Os modelos mistos sob a suposição de diferentes tipos de distribuição da variável resposta têm sido tema de pesquisas relativamente recentes. Várias aplicações podem servir como motivação para apresentar suas potencialidades. O foco desta apresentação é mostrar, por meio de exemplos das áreas da biologia, medicina, nutrição e fisioterapia, como estes modelos podem ser aplicados, incluindo métodos de predição e análises de resíduos. Pretende-se também suscitar a reflexão sobre tópicos de pesquisa neste assunto.

---

### **IPS3.2: A mixed model for positive responses augmented by zeros**

**Mariana Motta Rodrigues<sup>1</sup>**

<sup>1</sup> IME, Universidade Estadual de Campinas, Brasil

**Email:** *marianar@ime.unicamp.br*

**Abstract:** Semi-continuous response variables, such as food intake, are typically characterized by the presence of zeros and positive continuous outcomes that are often right-skewed. Linear regression models may be used for the transformed data as long as normality is supported. However, some appropriate transformations for skewed data may be inappropriate in the presence of zeros observations, as it is the case of the logarithmic transformation. Alternatively, we can analyse nonzero observations by ignoring zeros. However, it is impossible to predict the probability of zero and it may overestimate the mean. We propose a mixed regression model for positive and zero responses. We model the positive and zero responses using a mixture of a continuous distribution and a point mass in zero. We are particularly interested in studying positive responses whose distribution accommodates skewness. We model the mean of the positive response in a logarithm scale and the mixture probability in a logit scale, both as a function of fixed and random effects. The random effects link the two random components through their joint distribution and incorporate within subject correlation due to repeated measurements and between-subject heterogeneity. A Bayesian approach is adopted and a MCMC algorithm is tailored to obtain posterior distributions of the unknown quantities of interest. A case-deletion influence diagnostics based on appropriate divergence measures is performed. We motivate and illustrate the proposed methodology

by means of a data set from a 24 hours dietary recall study obtained in the city of São Paulo, Brazil, and present a simulation study to evaluate the performance of the proposed methods.

---

**Debatedor:** Mariza Andrade, Mayo Clinics, USA (email: mandrade@mayo.edu)

## IPS4: Robust statistics in the 21st century: New developments and applications

### Organizer and Chair:

*Vanda Lourenço, NOVA University of Lisbon, Portugal*

**Email:** *vmml@fct.unl.pt*

### IPS4.1: Functional outlier detection with applications to MRI image data

Mia Hubert<sup>1</sup>, Peter J. Rousseeuw<sup>1</sup>, Jakob Raymaekers and Pieter Segaert<sup>1</sup>

<sup>1</sup> KU Leuven, Belgium

**Email:** mia.hubert@wis.kuleuven.be

**Abstract:** The most common type of functional data are curves, typically measured over time. When several measurements are taken at each time point, we obtain multivariate functional data, such as height and weight curves of children. Whereas the domain of these data is always univariate, the response can thus be multivariate.

First, we set up a taxonomy of functional outliers and we construct new numerical and graphical techniques for the detection of outliers in multivariate functional data. Our tools include statistical depth functions and distance measures derived from them.

Next, we extend our methodology to functional data with a bivariate domain, such as images and surfaces. We illustrate our results on MRI image data and fluorescence excitation-emission surface data.

---

### IPS4.2: Detecting anomalous data cells

Peter J. Rousseeuw<sup>1</sup> and Wannes van den Bossche<sup>1</sup>

<sup>1</sup> KU Leuven, Belgium

**Email:** Peter@rousseeuw.net

**Abstract:** A multivariate dataset consists of  $n$  cases in  $d$  dimensions, and is often stored in an  $n$  by  $d$  data matrix. It is well known that real data may contain outliers. Depending on the circumstances, outliers may be (a) undesirable errors which can adversely affect the data analysis, or (b) valuable nuggets of unexpected information.

In statistics and data analysis the word outlier usually refers to a row of the data matrix, and the methods to detect such outliers only work when at most 50% of the rows are contaminated. But often only one or a few cells (coordinates) in a row are outlying, and they may not be found by looking at each variable (column) separately.

We propose the first method to detect cellwise outliers in the data which takes the correlations between the variables into account. It has no restriction on the number of contaminated rows, and can deal with high dimensions. Other advantages are that it provides estimates of the 'expected' values of the outlying

cells, while imputing the missing values at the same time. We illustrate the method on several real data sets, where it uncovers more structure than found by purely columnwise methods or purely rowwise methods.

---

## IPS4.3: An R package for robust orthogonal regression for compositional data

Valentin Todorov<sup>1,2</sup>, Klára Hružová<sup>3</sup>, Karel Hron<sup>3</sup> and Peter Filzmoser<sup>4</sup>

<sup>1</sup> United Nations Industrial Development Organization (UNIDO)

<sup>2</sup> Vienna International Centre, Vienna, Austria

<sup>3</sup> Palacký University, Czech Republic

<sup>4</sup> Vienna University of Technology, Vienna, Austria

**Email:** valentin.to@gmail.com

**Abstract:** In the context of building a regression model on compositional data, orthogonal regression is appropriate since all compositional parts - also the explanatory variables - are measured with errors. Moreover, in the presence of outlying observations in compositional data the orthogonal regression (that is able to handle the regression problem statistically) should be replaced by its robust counterpart. Statistical inference for the regression parameters is obtained by bootstrap. The methodology is illustrated with a data set from macroeconomics.

## IPS5: Innovative statistical methods with applications in medicine and biology

### Organizer and Chair:

*Paulo Henrique Ferreira, Universidade Federal da Bahia, Brasil*

**Email:** paulohenri@ufba.br

### **IPS5.1: Bivariate Clayton copula-based SUR Tobit model: A modified inference function for margins, interval estimation and an application to food consumption data**

**Paulo Henrique Ferreira<sup>1</sup> and Francisco Louzada<sup>2</sup>**

<sup>1</sup> Universidade Federal da Bahia, Brasil

<sup>2</sup> ICMC, Universidade de São Paulo, Brasil

**Email:** paulohenri@ufba.br

**Abstract:** This work extends the analysis of the bivariate Seemingly Unrelated Regression (SUR) Tobit model by modeling its nonlinear dependence structure through the Clayton copula. The ability to capture/model the lower tail dependence of the SUR Tobit model where some data are censored (generally, left-censored at zero) is an useful feature of the Clayton copula. We propose a modified version of the (classical) Inference Function for Margins (IFM) method by Joe and Xu (1996), which we refer to as Modified Inference Function for Margins (MIFM) method, to obtain the (point) estimates of the marginal and Clayton copula parameters. More specifically, we employ a (frequentist) data augmentation technique at the second stage of the IFM method (the first stage of the MIFM method is equivalent to the first stage of the IFM method) to generate the censored observations and then estimate the Clayton copula parameter. This process (data augmentation and copula parameter estimation) is repeated until convergence. Such modification at the second stage of the usual estimation method is justified in order to obtain continuous marginal distributions, which ensures the uniqueness of the resulting Clayton copula, as stated by Sklar (1959)'s theorem; and also to provide an unbiased estimate of the association parameter (the IFM method provides a biased estimate of the Clayton copula parameter in the presence of censored observations in both margins). Since the usual asymptotic approach, that is the computation of the asymptotic covariance matrix of the parameter estimates, is troublesome in this case, we also propose the use of resampling procedures (bootstrap methods) to obtain confidence intervals for the model parameters. We illustrate our procedure using bivariate data on consumption of salad dressings and lettuce by U.S. individuals.

## IPS5.2: SISGERES: A statistical system for urban solid waste management

Francisco Louzada<sup>1</sup>

<sup>1</sup> ICMC, Universidade de São Paulo, Brasil

**Email:** louzada@icmc.usp.br

**Abstract:** Nowadays more than 165 million people, 85% of Brazilians, live in cities. Their quality of life depends to a large extent on public policies in different sectors of the administration, taking into account environmental aspects. In this context, the issues related to management of urban solid waste is of primary concern, especially in view of the need to promote the reverse logistics of different materials disposed in order to promote the conservation of natural resources, human health, environmental health, contributing to the implementation of the National Solid Waste Policy (Federal law 12.305/2010). The management of waste arriving at the landfill through conventional collection is a priority, as well as those coming to the screening centers through selective collection. With the focus on the development of various metrics statistics to help solid waste management, we developed the Management System of Municipal Solid Waste (SISGERES). The system allows the study of generation and gravimetric composition of the waste. Supported by SISGERES, it is possible to know the relationship between demographics and waste generation; factors, and time related to the duration or obsolescence of a product; inventories of natural resources, towards consumption and the consequent generation of waste; forms of optimization-collecting, transport, storage, treatment and disposal of solid waste. This project is a partnership between Reevolta Cooperative and CEPID-CeMEAI.

---

## IPS5.3: Survival modelling for prediction of intrapartum interventions

Gleici da Silva Castro Perdoná<sup>1,2</sup>

<sup>1</sup> Ribeirão Preto School of Medicine, Brasil

<sup>2</sup> University of São Paulo, Brasil

**Email:** pgleici@fmrp.usp.br

**Abstract:** For many women, care during labor consists of the expectant monitoring and a hands-off approach to support. While, other women may require interventions to avoid complications or expedite labour and delivery. During this process, health professionals are frequently acquiring information, processing it and making decision to keep monitoring as it is, intensifying the monitoring or intervening. The important action is to identify the essential elements (including thresholds and interactions) of intrapartum monitoring that trigger the decision to consider interventions (e.g. augmentation, rupture of membranes, caesarean section and so on), as well as, to predict the probability of interventions. In order to model this process we proposed a multistate model where the states are considered as interventions. Our modeling is then embedded in a web-based Simplified, Effective, Labor Monitoring Assistant (SELMA) tool for simplify labor monitoring and support health workers decision-making. SELMA enables better, more consistent and timely decisions leading to safer labor management and better birth outcomes. This work is part of BOLD, the project for Better Outcomes in Labour Difficulty project to address weaknesses in labour care processes and better connecting health systems and communities. African data illustrate the applicability of our modeling. This work is co-authored by João Paulo Dias de Souza.

## **IPS6: New approaches for big and complex data in biostatistics**

### **Organizer and Chair:**

*Paulo Canas Rodrigues, Universidade Federal da Bahia, Brasil*

**Email:** paulocanas@gmail.com

### **IPS6.1: On the robust estimation of trait heritability**

**Vanda M. Lourenço<sup>1</sup>, Miguel Fonseca<sup>1</sup>, Ana M. Pires<sup>2</sup> and Paulo C. Rodrigues<sup>3</sup>**

<sup>1</sup> FCT & CMA, NOVA University of Lisbon, Portugal

<sup>2</sup> IST & CEMAT, University of Lisbon, Portugal

<sup>3</sup> Federal University of Bahia, Brazil

**Email:** vmml@fct.unl.pt

**Abstract:** Trait heritability, which is the proportion of phenotypic variation in the population that is caused by genetic differences, plays an important role in plant breeding studies since it reflects the breeding potential of a trait. Recently, there has been growing interest in marker-based estimation of heritability which is done via the fit of a linear mixed model, considering individual SNPs or other markers as fixed effects and an unobserved polygenic effect, whose covariance structure is estimated from genetic markers, in the random effects component of the model. In these circumstances, and assuming marker information is complete, both the coefficient of determination and broad-sense heritability are equivalent. Since the former is known to be biased when, in particular, the normality assumption of the error distribution is violated, the accurate estimation of heritability is therefore also compromised. Considering that data contamination is a major contributor to the infraction of the normality premise and that in plant studies phenotypic data are prone to several types of contamination, we propose a robust derivative-free restricted maximum likelihood (REML) framework for the linear mixed model and a robust coefficient of determination, based on a classical counterpart, upon which to estimate broad-sense heritability. The performances of both approaches in the estimation of broad-sense heritability are compared via simulation and an example of application of the robust methodology to a real maize dataset is presented in order to validate the usefulness of the proposed robust approach.

---

### **IPS6.2: Set-based interaction analysis: Some new observations**

**Bhramar Mukherjee<sup>1</sup>**

<sup>1</sup> University of Michigan, USA

**Email:** bhramar@umich.edu

**Abstract:** Most interaction searches in the post-genome-wide association study (GWAS) era have focused on single nucleotide polymorphisms (SNPs) as the unit of analysis. Given the fact that SNPs in a gene/pathway may function in coordination with others, there have been some recent advances in the area of set-based interaction tests. In this talk, we introduce a statistical framework for set-based interac-

tion analysis that achieves robust inference under misspecification of the main effects of genetic variants and environmental exposures. The more well-established theory of set-based marginal genetic association tests takes a twist here as the null model for testing interaction is also potentially high-dimensional, containing main effects of many SNPs. We describe a selection strategy for genetic variants to be able to fit the null model. We also delineate the role of gene-environment independence and misspecification of main effects in this context. We illustrate the methods through extensive simulation studies and analysis of data from the Multi-Ethnic Study of Atherosclerosis (MESA).

---

### IPS6.3: Pseudo-likelihood and split-sample methods in small and very large trials

Geert Molenberghs<sup>1,2</sup>

<sup>1</sup> I-BioStat, Hasselt University, Diepenbeek, Belgium

<sup>2</sup> I-BioStat, KU Leuven, Belgium

**Email:** geert.molenberghs@uhasselt.be

**Abstract:** Hierarchical and otherwise clustered data can easily be analyzed using maximum likelihood or Bayesian methods in moderate to large samples. This may well be different in very small populations (e.g., clinical trials in rare diseases) or, on the other side of the spectrum, when samples become very large.

We explore a few techniques that can be used in such settings. One method is pseudo-likelihood (or composite likelihood), where a cumbersome likelihood is replaced by a simpler function, which is easier to maximize and still produces consistent and asymptotically normal estimates. On the other hand, the sample can be split into sub-samples, each of which is analyzed separately (ideally in parallel), after which the results are funneled into a single set of inferences using appropriate combination rules. These techniques can also be combined.

Using a set of examples, we illustrate how the methods work, and what the computational gains are.

## **IPS7: Estatística Experimental**

### **Organizer and Chair:**

*Roseli Aparecida Leandro, ESALQ, Universidade de São Paulo, Brasil*

**Email:** *roseli.ap.leandro@gmail.com*

## **IPS7.1: On compound design criteria**

**Luzia Aparecida Trinca<sup>1</sup>**

<sup>1</sup> Universidade Estadual Paulista, Botucatu, Brasil

**Email:** ltrinca@ibb.unesp.br

**Abstract:** For practical purposes an experimental design should present several good properties as the list highlighted by Box and Draper in 1975, further emphasised and enlarged by many other authors. Optimum design theory allows the construction of very efficient and economical designs of experiments. However such designs are usually optimal for the specific property optimised and under the correctness of the statistical model supposed at the planning phase. Advances are possible by using multiple criteria or composite criteria of the type proposed by Gilmour and Trinca in 2012.

In this talk a range of properties is explored in order to construct efficient and flexible factorial designs for response surface studies.

The approach provides efficient designs for parameter estimation, under practical experimental restrictions, that allow inferences to be carried out while preserving good performances on several desired aspects. Illustrations are motivated from practical problems.

---

## **IPS7.2: Variance component estimation on discrete time competing risks analysis with masked causes**

**Rafael Pimentel Maia<sup>1</sup>**

<sup>1</sup> Universidade Estadual de Campinas, Brasil

**Email:** rafaelpmaia@ime.unicamp.br

**Abstract:** In longevity studies often the interest lies in modeling the time until death of a group of individuals that might die of one of different possible specific causes (competing risks). The cause of death is said to be masked when the when the time to death of this individual is observed but not the death cause. This work will study some techniques, based on suitable variants of the EM-algorithm, to perform statistical inference in a competing risk scenario in the presence of partial masking and right censoring. The goal is to extent a class of multivariate proportional hazard models for competing risks containing suitably gaussian random components to characterize the quantitative genetic determination of longevity in large scale animal production systems. Moreover, it is provided a simulation study evaluating the performance of the proposed inference procedures with respect the estimation of variance-covariance components.

## IPS7.3: Case study of sugarcane trials with commercial varieties and test lines

Alessandra dos Santos<sup>1</sup>, Chris J. Brien<sup>2</sup>, Clarice Garcia Borges Demétrio<sup>1</sup> and Renata Alcarde Sermarini<sup>1</sup>

<sup>1</sup> ESALQ, Universidade de São Paulo, Brasil

<sup>2</sup> University of South Australia, Australia

**Email:** ralcarde@usp.br

**Abstract:** Breeding companies continuously pursue lower production costs and better crop yield. The release and recommendation of new commercial varieties requires several breeding trials with potential genotypes obtained from crossing of known varieties. At the initial phases of the breeding program there are many genotypes and little material for planting in experiments, which are usually in large fields, depending on the studied crop (sugarcane, soybean, corn, etc.). To cope with this, all the blocks have commercial varieties which are considered as standard and new varieties can appear in more than one block. In this study, a full-sib progeny was obtained from a bi-parental cross between the elite clone IACSP95-3018 (female parent) and the variety IACSP93-3046 (male parent), which were developed by the Sugarcane Breeding Program at the Instituto Agronômico de Campinas (IAC). IACSP95-3018 is a promising clone used in the IAC Sugarcane Breeding Program, whereas IACSP93-3046 has a high level of sucrose, good tillering and an erect stool habit and is recommended for mechanical harvest. The full-sib progeny and the sugarcane varieties RB835486 and SP81-3250, used as checks, were planted in 2011 in the Ribeirão Preto municipality, São Paulo, Brazil, in a randomized complete block design, with three replicates and plots of 2 meter rows spaced 1.5 meter apart. This trial contained 708 plots in a rectangular array with 59 rows and 12 columns. Both parents and two varieties (SP81-3250 and RB835486) were included four times in each replicate as checks and there were 220 crosses that occurred once in each replicate. The aim is to review the theoretical principles related to block designs, the mixed model approach and spatial analysis of experiments. The experiment will be analyzed as a mixture of treatments as fixed effects (control) and others as random effects (new treatments) and will verify if there is spatial dependence and competition between individuals. The field experiment was carried out in 2012 and the response evaluated was plant weight (kg), where each measurement was obtained from a sample of 10 stalks harvested from each individual plot, according to the methods described in Consecana (2006), for both sugar cane and other ratoon crops. For model selection the REML likelihood ratio test or/and Bayesian Information Criteria (BIC) has been used for each model.

---

**Debatedor:** Clarice Garcia Borges Demétrio, ESALQ, Universidade de São Paulo, Brasil  
(email: clarice.demetrio@usp.br)

## **IPS8: Sessão ABE – Projetos de divulgação e educação estatística**

### **Organizer and Chair:**

*Vera Tomazella, Presidente da Associação Brasileira de Estatística / UFSCar, Brasil*

**Email:** *veratomazella@gmail.com*

### **IPS8.1: Literacia, capacitação e divulgação estatística no ISI**

**Pedro Luis do Nascimento Silva<sup>1,2</sup>**

<sup>1</sup> Instituto Brasileira de Geografia e Estatística, Brasil

<sup>2</sup> President do International Statistical Institute

**Email:** pedronsilva@gmail.com

**Resumo:** O ISI - Instituto Internacional de Estatística tem como missão “promover a compreensão, desenvolvimento e a boa prática da Estatística no mundo todo”. Como parte dessa missão desenvolve atividades em várias áreas. Apresentarei as atividades voltadas ao desenvolvimento da literacia estatística, à capacitação e à divulgação da Estatística desenvolvidas pelo instituto.

---

### **IPS8.2: Reflexões sobre o potencial do ensino de estatística na educação básica**

**Irene Mauricio Cazorla<sup>1</sup>**

<sup>1</sup> Universidade Estadual de Santa Cruz, Brasil

**Email:** icazorla@uol.com.br

**Resumo:** A importância da Estatística na formação de cidadãos fez o Brasil recomendar seu ensino na Educação Básica por meio dos Parâmetros Curriculares Nacionais, cabendo a Matemática a formalização dos seus conceitos e procedimentos e, a outros componentes curriculares como Ciências, Biologia, Geografia e Sociologia, o seu uso como ferramenta de apreensão e compreensão dos fenômenos da natureza e sociais. A Estatística por ser ferramenta do método científico permite o planejamento, organização, coleta e tratamento de dados nas variadas áreas do conhecimento; informações que serão veiculadas pelos grandes meio de comunicação e, portanto, crucial para a formação científica e cidadã dos estudantes da Educação Básica. Tecemos reflexões sobre o potencial papel do ensino de Estatística, a partir das ideias de Gal sobre o letramento estatístico e das fases do ciclo investigativo de Wild e Pfannkuch. Focamos o distanciamento dos currículos dos cursos de licenciatura que formam os professores da Educação Básica, em especial, o de Matemática, que não preparam os futuros professores para o uso das ferramentas estatísticas no contexto da investigação científica e do levantamento de dados para a tomada de decisões. Por fim, indicamos a necessidade da comunidade de estatísticos, matemáticos e educadores para envidar esforços para intervir, de forma institucional, na formação estatística de professores da Educação Básica, bem como dos formadores desses professores na Educação Superior, a fim de propiciar um ensino que

possibilite os professores e estudantes da Educação Básica utilizar a Estatística como uma ferramenta ao serviço de sua formação cidadã.

---

### **IPS8.3: Projeto Bussab de divulgação e educação estatística**

**Francisco Louzada<sup>1</sup>**

<sup>1</sup> ICMC, Universidade de São Paulo, Brasil

**Email:** louzada@icmc.usp.br

**Resumo:** O “Projeto Bussab de Divulgação e Educação Estatística”, tem como objetivo aumentar o conhecimento da Estatística e a sua inserção dentro da comunidade e aumentar o número de candidatos para cursos de graduação e reduzir a evasão nos cursos de bacharelado.

**IPS9: Sessão EMBRAPA****Organizers and Chairs:**

*Carlos Alberto da Silva Ledo, Embrapa Mandioca e Fruticultura, Brasil*

**Email:** *carlos.ledo@embrapa.br*

*Waldomiro Barioni Jr., Embrapa Pecuária Sudeste, Brasil*

**Email:** *waldomiro.barioni@embrapa.br*

**IPS9.1: Embrapa: “Quem somos nós?”**

Waldomiro Barioni Jr.<sup>1</sup> e Carlos Alberto da Silva Ledo<sup>2</sup>

<sup>1</sup> Embrapa Pecuária Sudeste - São Carlos/SP, Brasil

<sup>2</sup> Embrapa Mandioca e Fruticultura, Brasil

**Email:** waldomiro.barioni@embrapa.br; carlos.ledo@embrapa.br

**Resumo:** Apresentação Institucional da Empresa Brasileira de Pesquisa Agropecuária – Embrapa, fundada em 1973, vinculada ao Ministério da Agricultura Pecuária e Abastecimento. Mostrando a sua Missão, Visão e Valores e a sua atuação com a distribuição geográfica no Brasil e no exterior. A Embrapa é uma empresa de inovação tecnológica focada na geração de conhecimento e tecnologia para a agricultura brasileira com programas de pesquisa que abrangem várias cadeias produtivas: animal, vegetal, energia e fibras.

**IPS9.2: Métodos quantitativos e computação científica na Embrapa**

Geraldo da Silva e Souza<sup>1</sup>

<sup>1</sup> Embrapa SEDE/DF, Brasil

**Email:** geraldo.souza@embrapa.br

**Resumo:** Os Métodos Quantitativos e a Computação Científica (MQCC) definem áreas de estudos que dizem respeito à construção de modelos matemáticos, ao estudo de técnicas de análise quantitativa e ao uso de computadores na análise e solução de problemas científicos. Muito da pesquisa aplicada em geral, e na agricultura em particular, baseia-se em processos de decisão sob condições de incerteza. A modelagem neste contexto depende de métodos quantitativos e em última análise de computação científica. Ciências como Estatística, Pesquisa Operacional, Econometria, Biometria e outras são de muita importância neste contexto. A Empresa Brasileira de Pesquisa Agropecuária (Embrapa), desde sua fundação, no início dos anos 70, tem reconhecido a importância dessas áreas na execução da pesquisa agropecuária no país. Nesta exposição apresenta-se a Embrapa, a história de MQCC e aplicações correntes da área na empresa.

## IPS9.3: Armazenamento e processamento de grandes volumes de dados na Embrapa

Roberto Hiroshi Higa<sup>1</sup>

<sup>1</sup> Embrapa Informática - Campinas/SP, Brasil

**Email:** roberto.higa@embapa.br

**Resumo:** A sociedade passa, atualmente, por grandes transformações devido à uma crescente capacidade de gerar, medir, coletar e armazenar grandes volumes de dados e, então, utilizá-los para geração de conhecimentos e tomada de decisão. É a chamada era do “Big Data”. Na pesquisa científica o “Big Data” se manifesta a partir do desenvolvimento de tecnologias capazes de gerar conjuntos de dados da ordem de terabytes/petabytes como, por exemplo, dados de sequências genômicas, imagens de satélite em alta resolução ou dados obtidos de redes de sensores. Na agricultura, são esperados impactos no melhoramento genético, na previsão de clima, na agricultura de precisão e no entendimento da dinâmica dos mercados, entre outros. Nesta exposição são apresentados esforços atualmente em desenvolvimento na Embrapa para se estruturar para esse cenário, com foco no armazenamento e organização de seus dados e na criação de uma infraestrutura institucional de processamento de alto desempenho para suporte à análise desses dados.

Part VII

**Round Tables**



## RT1: Meet the Presidents

Paulo Canas Rodrigues<sup>1</sup> and John Hinde<sup>2</sup>

<sup>1</sup> Federal University of Bahia, Brazil

<sup>2</sup> National University of Ireland Galway, Ireland

**Email:** paulocanas@gmail.com

**Abstract:** In this roundtable the presidents of three of the most important worldwide statistical societies will discuss the current state of statistical sciences and share some of their experience with the attendees.

### Participants:

- Peter J. Diggle, President of the Royal Statistical Society  
Email: p.diggle@lancaster.ac.uk
- Pedro Luis Nascimento Silva, President of the International Statistical Institute  
Email: pedronsilva@gmail.com
- Elizabeth A. Thompson, President of the International Biometrical Society  
Email: eathomp@u.washington.edu

## RT2: Tecnologia e estatística aplicadas a políticas de saúde

Doris Fontes<sup>1</sup>

<sup>1</sup> Presidente do CONRE-3, SP, Brasil

**Email:** dsfontes@gmail.com

**Resumo:** Com o aumento do uso de grandes bancos de dados, de características, fontes e formatos diversos, o mercado de trabalho assiste também a um crescimento contínuo no número de vagas para estatísticos. Em todas as áreas, seja na pesquisa acadêmica ou mercado fora das universidades, há grande interesse em utilizar dados já disponíveis e, para tanto, o estatístico precisa estar preparado para entender, organizar, estruturar e analisá-los de forma rápida e útil. Na área da saúde, o uso da estatística sempre foi muito grande, mas tem crescido a preocupação pelo uso inadequado, insuficiente ou incorreto de técnicas estatísticas em pesquisas científicas. Revisões de artigos científicos da área médica têm sido cada vez mais constantes. As reclamações sobre notícias distorcidas pelos jornais têm crescido muito também, como no caso do Zika vírus. No mercado de trabalho, público ou privado, a saúde tem uma importância e peso muito grandes. A escolha pelas técnicas estatísticas corretas para subsidiar uma tomada de decisão pode significar uma grande economia para uma seguradora/plano de saúde, um hospital ou um paciente. Assim, essa mesa redonda propõe discutir algumas técnicas estatísticas já utilizadas por hospitais e pesquisadores da área médica, além de discutir a importância da qualificação dos estatísticos que lidam com esse tipo de dados.

### Participantes:

- Ângela Tavares Paes (UNIFESP e Presidente da Comissão de Educação e Aperfeiçoamento do CONRE-3): *O papel da Bioestatística nas práticas em saúde*  
Email: atpaes@gmail.com
- Jackson Santos da Conceição (Presidente do CONRE-5): *Analytics na administração hospitalar*  
Email: jsantos1013@gmail.com
- Luís Claudio Lemos Correia (Médico, Professor adjunto da Escola Bahiana de Medicina e Saúde Pública): *Zika e microcefalia: Fato ou ficção científica?*  
Email: lccorreia@cardiol.br

Part VIII

**Contributed Papers**



---

## CPS1: Recent advances in survival analysis

---

### CPS1.1: The exponentiated Kumaraswamy inverse Weibull distribution

Jailson A. Rodrigues<sup>1</sup>, Gholamhossein Hamedani<sup>2</sup>, Ana Paula Silva<sup>3</sup> and Angela Silva<sup>4</sup>

<sup>1</sup> Instituto Federal da Bahia, BA, Brasil

<sup>2</sup> Marquette University, USA

<sup>3</sup> Universidade Federal São João del-Rei, MG, Brasil

<sup>4</sup> Universidade Federal da Bahia, BA, Brasil

**Email:** jailsondearaujo@yahoo.com.br

**Abstract:** A new distribution called the exponentiated Kumaraswamy inverse Weibull is proposed. This distribution includes as special cases the inverse exponential, inverse Weibull, inverse Rayleigh and exponentiated inverse Weibull distributions. We study the main properties of this distribution, with special emphasis on its moments and characterizations of model. We also discuss parameter estimation considering the methods of moments and maximum likelihood. An application reveals that the model proposed can be very useful in fitting real data.

---

### CPS1.2: Regression models for the restricted mean residual time

Giuliana Cortese<sup>1</sup>, Stine Holmboe<sup>2</sup> and Thomas H. Scheike<sup>3</sup>

<sup>1</sup> Department of Statistical Sciences/University of Padova, Itália

<sup>2</sup> Department of Growth and Reproduction/Rigshospitalet, Copenhagen, Dinamarca

<sup>3</sup> Department of Biostatistics/University of Copenhagen, Copenhagen, Dinamarca

**Email:** gcortese@stat.unipd.it

**Abstract:** Hazard functions are typically studied via Cox's regression survival models. The resulting hazard ratios estimate the relative risk, and these measures are difficult to interpret and hard to be translated into clinical benefits in terms of increased survival time (Uno et al., J Clin Oncol, 2014). The main objective is often to study survival functions, aiming at a global summary over a specific time period. Therefore, there has been increasing interest in summary measures based on the survival function, easily to interpret than the hazard ratio. For the survival time  $T$ , we consider the mean residual time  $E(T - t|T \leq t)$ , which has recently received increasing interest in literature (Sun et al., 2012). This quantity represents a partial area under the survival function and is interpreted as the residual life expectancy of individuals who had survived up to a certain time  $t$ . However, due to the presence of right censoring, the tail of the survival distribution is often difficult to be correctly estimated. As a solution, we propose to study the restricted mean residual time  $E(\min(T, \tau) - t|T \leq t)$ , for any  $\tau > 0$ . We present regression models for this new measure, based on weighted estimating equations and the

inverse probability of censoring weighted estimator to model potential right censoring. We also show how to extend the models to deal with delayed entries. Estimation performance is investigated by simulation studies. Using real data about the Danish Monitoring Cardiovascular Risk Factor Survey, we illustrate an application of these regression models when the link function is identity or exponential.

---

### CPS1.3: A flexible cure rate model with frailty term in latent risks: a Bayesian approach

Agatha Rodrigues<sup>1,2</sup> and Vinicius Calsavara<sup>2</sup>

<sup>1</sup> IME/ Universidade de São Paulo, SP, Brasil

<sup>2</sup> A.C. Camargo Cancer Center, SP, Brasil

**Email:** cokin15@hotmail.com

**Abstract:** In this paper we propose a flexible cure rate model with a frailty term in latent risks, which is obtained of Cancho et al. 2011 models by incorporating a frailty term in risk function of latent competing causes. We consider a cure rate model based on the negative binomial distribution for the number of competitive causes and power variance function for the frailty variable, encompassing several special models. It is more flexible in terms of dispersion and by quantify the unobservable heterogeneity. Bayesian analysis is based on Markov chain Monte Carlo (MCMC) methods. We performed several simulation studies and the practical relevance of the proposed model is demonstrated in a real data set.

---

### CPS1.4: Subconjuntos de validação em análise de sobrevivência com censura intervalar e erros de classificação

Magda Pires<sup>1</sup>, Enrico Colosimo<sup>1</sup>, Guilherme Veloso<sup>1</sup> e Raquel Ferreira<sup>1</sup>

<sup>1</sup> Universidade Federal de Minas Gerais, MG, Brasil

**Email:** magda@est.ufmg.br

**Resumo:** Em estudos de análise de sobrevivência, muitas vezes o teste diagnóstico utilizado para detectar a ocorrência do evento de interesse está sujeito a erros de classificação, que podem ser quantificados por medidas como sensibilidade e especificidade. Incorporar essas medidas em um modelo de sobrevivência é fundamental para se obter estimativas não viciadas dos parâmetros da distribuição do tempo de falha. Quando não se tem informações sobre a magnitude dos erros de classificação, o modelo é não identificável, mas esse problema pode ser contornado utilizando uma abordagem Bayesiana associada ao uso de subconjuntos de validação. Estudos de simulação demonstraram que a metodologia proposta apresentou resultados satisfatórios quando comparados ao modelo convencional, com estimativas menos viciadas.

---

## CPS2: Recent advances in mixed models and applications

---

### CPS2.1: Modelos mistos de regressão segmentada: Um exemplo

Julio Singer<sup>1</sup>, Francisco Marcelo Monteiro-Da-Rocha<sup>1</sup>, Antonio Carlos Pedroso-De-Lima<sup>1</sup>, Giovani Silva<sup>2</sup>, Giuliana Coatti<sup>1</sup> e Mayana Zatz<sup>1</sup>

<sup>1</sup> Universidade de São Paulo, SP, Brasil

<sup>2</sup> Universidade de Lisboa, Portugal

**Email:** jmsinger@ime.usp.br

**Resumo:** Consideramos um conjunto de dados proveniente de um estudo realizado no Instituto de Biociências da Universidade de São Paulo com o objetivo de avaliar o efeito da utilização de células-tronco para tratamento de camundongos transgênicos (SOD1-G93A). Esses animais servem de modelo para o estudo da esclerose lateral amiotrófica, uma doença neurodegenerativa bastante severa e atualmente sem cura. De maneira semelhante à forma clínica que acomete os pacientes, nos camundongos, os sintomas iniciais são tremores nos membros evoluindo progressivamente até a paralisia total com a consequente morte. Neste modelo, os sintomas da doença tendem a se manifestar a partir da décima segunda semana de vida quando se observa uma diminuição da capacidade motora dos animais. Um conjunto de 55 animais (34 fêmeas e 21 machos) foi subdividido em três grupos e cada um deles foi tratado semanalmente ou com células mesenquimais ou com pericitos ou com solução de Hank (o veículo utilizado nos demais tratamentos) a partir da oitava semana de vida até a morte natural. Quatro variáveis associadas ao desenvolvimento motor foram observadas semanalmente. Utilizamos técnicas de identificação de pontos de mudança no valor esperado das respostas e modelos mistos de regressão segmentada para avaliar o efeito de tratamento e sexo nas taxas esperadas de variação das respostas correspondentes ao longo da vida dos animais. Fazemos um diagnóstico do ajuste dos modelos e identificamos tópicos para futuras pesquisas.

---

### CPS2.2: Análise de dados longitudinais da evapotranspiração de alface em função de níveis de água

Omar Pereira<sup>1</sup>, Altair Bertonha<sup>1</sup> e Isolde Previdelli<sup>1</sup>

<sup>1</sup> Universidade Estadual de Maringá, PR, Brasil

**Email:** omarcnpereira@gmail.com

**Abstract:** The objective of this study was to evaluate the evapotranspiration of lettuce *Lactuca sativa* cv. Veronica according to four levels of irrigation, planted in pots under greenhouse over a period of 23 days through a nonlinear mixed model. When the humidity of the substrate reached 50% of its maximum capacity retention, we added water to raise the humidity according to the four treatments: W1=62.5%; W2=75.0%; W3=87.5% and W4=100.0%. The non-linear mixed-effects model adjusted very well to the

data and was able to show that the increase of agricultural production is subject to the availability of water in the soil and, in some cases, high levels of water are not required for maximized over there.

---

### **CPS2.3: The longitudinal approach to evaluating changes behavioural in rats with cerebral ischemia induced by chronic cerebral hypoperfusion**

**Matheus Henrique Dal Molin Ribeiro<sup>1</sup>, Omar Pereira<sup>2</sup>, Humberto Milani<sup>2</sup> and Isolde Previdelli<sup>2</sup>**

<sup>1</sup> Universidade Tecnológica Federal do Paraná, PR, Brasil

<sup>2</sup> Universidade Estadual de Maringá, Paraná, Brasil

**Email:** matheus.dalmolinribeiro@gmail.com

**Abstract:** Longitudinal studies associated with animal models have been frequently made in the field of pharmacology in order to understand the cognitive behavior of animals subjected to cerebral ischemia through different methods, in the presence of different treatments. Thus, the aim of this study was to apply the methodology of Generalized Estimating Equations (GEE) for the evaluation of an experiment whose purpose is to evaluate the effect of different treatment times with Atorvastatin in animals subjected to cerebral ischemia by chronic cerebral hypoperfusion. Accordingly, a longitudinal negative log-binomial model is used to model the response variable error of reference (Er) of this experiment, in which by modeling a work correlation matrix, the intra-individual variations were accommodated. From this methodology, it was found that animals treated with Atorvastatin showed a restoration of cognitive abilities compared to placebo-treated animals, independently of the time of treatment.

---

### **CPS2.4: Modelo linear misto: Aplicações no mapeamento de determinantes genéticos associados com depressão e ansiedade**

**Tâmara P. Taporoski<sup>1</sup>, André B. Negrão<sup>1</sup>, Andréa R. V. R. Horimoto<sup>2</sup>, Nubia E. Duarte<sup>2</sup>, Mariza de Andrade<sup>3</sup>, José E. Krieger<sup>2</sup>, Malcolm von Schantz<sup>4</sup>, Homero Vallada<sup>1</sup> e Alexandre C. Pereira<sup>2</sup>**

<sup>1</sup> Departamento e Instituto de Psiquiatria (LIM23) - Faculdade de Medicina da Universidade de São Paulo, SP, Brasil

<sup>2</sup> Laboratório de Genética e Cardiologia Molecular/Instituto do Coração (Incor) - Faculdade de Medicina da Universidade de São Paulo

<sup>3</sup> Divisão de Estatística Biomédica e Informática - Clínica Mayo, Rochester, MN, USA

<sup>4</sup> Departamento de Bioquímica e Fisiologia - Faculdade de Saúde e Ciências Médicas, Universidade de Surrey, Guildford, Reino Unido

**Email:** nued03@gmail.com

**Resumo:** O Modelo Linear Misto tem ampla aplicabilidade nas mais diversas áreas, principalmente por permitir a modelagem de diferentes estruturas de dependência entre observações. Especificamente para análises de associação em dados em família este modelo leva em consideração o parentesco existente entre indivíduos e tem a flexibilidade para modelar efeitos fixos e efeitos genéticos aleatórios. Neste trabalho

aplica-se o modelo linear misto, Univariado e Multivariado, em um estudo cujo intuito é identificar variantes genéticas relacionadas a escores de Depressão e Ansiedade e discriminar as variantes que regulam os dois escores simultaneamente ou seja achar variantes genéticas com efeitos de Pleiotropia. Como resultado, do ponto de vista estatístico foi possível discriminar variáveis preditoras genéticas associadas individualmente aos escores e também resultados que fornecem forte evidência do compartilhamento genético entre sintomas de Depressão e Ansiedade o qual tem relevância para o entendimento de bases biológicas desses transtornos. A respectiva aplicação é realizada em dados de família e plataformas de marcadores moleculares do tipo SNP (Single Nucleotide Polimorphism) do Projeto “Corações de Baependi” (Processo Fapesp 2007/5850-7) do Laboratório de Genética e Cardiologia Molecular, INCOR - USP.

---

## CPS3: Applications in biometry and biostatistics

---

### CPS3.1: Accuracy in the estimates of zucchini production

Alessandro D. Lúcio<sup>1</sup> and Vilson Benz<sup>1</sup>

<sup>1</sup> Universidade Federal de Santa Maria, RS, Brasil

**Email:** adlucio@ufsm.br

**Abstract:** The productive variability in horticultural crops affects the planning and quality of the experiments, inflating the error and leading to wrong conclusions with low experimental accuracy and reliability. The objective of this study was to dimension the plot size and number of harvests that are necessary so that there is improvement in the accuracy of zucchini production. The production data used was the one from the uniformity trials where the basic units (BU) were identified with the number of crop row and were numbered according to the position inside the row. With the values of individual production, (fresh fruit biomass) different sizes of parcels were simulated by summing up the adjacent BUs in the crop rows and the two forms of harvesting groups. For each parcel simulated size and their respective harvests, a variation coefficient (%) was estimated at the crop row. A response surface design was applied, with a dependent variable the variation coefficient and, with the independent variable, the plot size and the number of harvests. The best combination between the plot size and the number of harvests is seven plants and half the productive cycle in both growing seasons.

---

### CPS3.2: Sample size for estimating the concentration of organisms in ballast water

Eliardo Costa<sup>1</sup>, Rubens Lopes<sup>1</sup> and Julio Singer<sup>1</sup>

<sup>1</sup> IME e IO/ Universidade de São Paulo, SP, Brasil

**Email:** eliardo@ime.usp.br

**Abstract:** We consider the computation of sample sizes for estimating the concentration of organisms in ballast water. Given the possible heterogeneity of their distribution in the tank, we adopt a negative binomial model to obtain confidence intervals for the concentration. We show that the results obtained by Chen and Chen (2012) in a different set-up hold for the proposed model and use them to develop algorithms to compute sample sizes both in cases where the concentration is known to lie in some bounded interval or where there is no information about its range. We also construct simple diagrams that may be easily employed to decide for compliance with the D-2 regulation of the International Maritime Organization (IMO).

## CPS3.3: Distribuições zero-modificadas biparamétricas: Uma aplicação aos dados de variações de temperatura média global

Luisa Hebling<sup>1</sup> e Katiane Conceição<sup>1</sup>

<sup>1</sup> ICMC/Universidade de São Paulo, SP, Brasil

**Email:** katianeconceicao@gmail.com

**Resumo:** Este trabalho tem como objetivo apresentar o procedimento de estimação pelo Método dos Momentos para os parâmetros das distribuições Série de Potência Zero-Modificadas biparamétricas. Essa classe de distribuições, que é uma extensão da família de distribuições Série de Potência, permite ser ajustada a conjuntos de dados de contagem sem o conhecimento prévio da frequência de observações zero presentes na amostra. Para inferir sobre os parâmetros, foram utilizados os procedimentos Bootstrap paramétrico e não-paramétrico. Foi realizada uma aplicação com dados reais, obtidos a partir da variação anual da temperatura global.

---

## CPS3.4: Estudos, em termos estatísticos, das condições de moradia para pessoas com deficiências

Paulo Oliveira<sup>1</sup>

<sup>1</sup> IME/Universidade de São Paulo, SP, Brasil

**Email:** poliveir@ime.usp.br

**Resumo:** Acredita-se que as pessoas com deficiência se encontram em condições de desigualdades, violações de sua dignidade e com perda de sua autonomia. Em 1991, de acordo com a OMS, ficou estabelecido que moradia adequada é aquela que oferece abrigo contra as mudanças climáticas, condições de saúde, segurança e um tamanho mínimo para ser considerada habitável. Devem ser concedidas as facilidades sanitárias necessárias, servida por serviços públicos essenciais, incluindo água, esgoto, energia elétrica, iluminação pública, coleta de lixo, calçadas, transporte público e acesso a equipamentos sociais. Políticas públicas de inclusão estão sendo acionadas nos diversos movimentos sociais ligados a pessoas com deficiência como parte de sua mobilização pela busca de seus direitos básicos, ação que insta as autoridades a pensar e desenvolver projetos que tratam dessa população. Neste trabalho, pretende-se mostrar como encontra as condições de acessibilidade para as pessoas com as diferentes deficiências em relação às que não apresentam deficiência, considerando os dados obtidos a partir de entrevistados que compuseram a amostra do Censo 2010 do IBGE (Instituto Brasileiro de Geografia e Estatística), e, para cada um dos diferentes tipos de deficiência, vamos investigar quais as condições de vida considerando as variáveis que mais influência e melhores partições usando técnicas como o teste de homogeneidade, a AID, análise de regressão logística ordinal estereótipo, análise de risco, seleção de modelos e seleção de variáveis. Em moradia, incluindo a Constituição do Brasil a acessibilidade para residência e reabilitação das pessoas com deficiências e promovendo a sua integração à vida comunitária.

---

## CPS4: Regression and spatio-temporal modelling

---

### CPS4.1: On quantile residuals in beta regression

Gustavo Pereira<sup>1</sup>

<sup>1</sup> Universidade Federal de São Carlos, SP, Brasil

**Email:** gpereira@ufscar.br

**Abstract:** Beta regression is often used to model the relationship between a dependent variable that assumes values on the open interval  $(0,1)$  and a set of predictor variables. An important challenge in beta regression is to find residuals whose distribution is well approximated by the standard normal distribution. Two previous works compared residuals in beta regression, but the authors did not include the quantile residual. Using Monte Carlo simulation techniques, this paper studies the behavior of certain residuals in beta regression in several scenarios. Overall, the results suggest that the distribution of the quantile residual is better approximated by the standard normal distribution than that of the other residuals in most scenarios.

---

### CPS4.2: Previsão de preços para soja e café com base em variáveis climáticas: Aplicação da análise espectral multivariada

Carlos Pinheiro<sup>1,2</sup> e Valter Senna<sup>1</sup>

<sup>1</sup> Cimatec /Serviço Nacional de Aprendizagem Industrial, BA, Brasil

<sup>2</sup> Universidade do Estado da Bahia, BA, Brasil

**Email:** email@carlospinho.com.br

**Resumo:** Este estudo tem como objetivo aplicar a metodologia proposta por Pinheiro & Senna (2015) para previsão dos preços do café e da soja produzidos no Brasil, com base nas variáveis climáticas temperatura, precipitação e umidade relativa média. O caráter multivariado do modelo de análise espectral multivariada demonstrou-se adequado, quando comparado com o modelo de redes neurais, ao permitir melhores desempenhos preditivos. Os resultados obtidos no período fora da amostra, mediante o uso de medida de erro e teste estatístico confirmam isso. A pesquisa é útil na formulação e na aplicação de políticas direcionadas ao setor agrícola por conta da relevância da previsão dos preços ao especificar a tendência dos mesmos.

## CPS4.3: Skew scale mixtures of normal distributions and applications

Clécio Ferreira<sup>1</sup> and Victor Hugo Lachos<sup>2</sup>

<sup>1</sup> Universidade Federal de Juiz de Fora, MG, Brasil

<sup>2</sup> Universidade Estadual de Campinas, SP, Brasil

**Email:** clecio.ferreira@yahoo.com.br

**Abstract:** Ferreira (2011) proposed a new family of distributions called skew scale mixtures of normal distributions (SSMN), including the skew-t, skew-slash, skew-contaminated normal and the skew-exponential power distributions and the skew-normal as particular case. Ferreira (2016) extend these distributions to multivariate version. This family of distributions is attractive because it simultaneously takes into account the skewness and heavy tails, besides having a stochastic representation that allows easy implementation of the expectation-maximization (EM) algorithm, which facilitates the study of many of its properties. In this work, we presented some generalizations and applications of the SSMN distributions, as multivariate version, regression models, mean-shift method to detect outliers, nonlinear models.

---

## CPS4.4: Análise de fronteiras geográficas e limites no processo temporal para avaliação de conglomerados espaço temporais

Spencer Barbosa da Silva<sup>1</sup>, Anderson Ribeiro Duarte<sup>1</sup>, Fernando Luiz Pereira de Oliveira<sup>1</sup>, Alexandre Celestino Leite Almeida<sup>3</sup> e Luiz Duczmal<sup>4</sup>

<sup>1</sup> Universidade Federal de Ouro Preto, MG, Brasil

<sup>2</sup> Universidade São João Del-Rei, MG, Brasil

<sup>3</sup> Universidade Federal de Minas Gerais, MG, Brasil

**Email:** fernandoluizest@gmail.com

**Resumo:** A investigação de clusters espaço temporais tem se tornado um assunto de vasto interesse nas pesquisas científicas. A velocidade da propagação de diversos fenômenos tem impulsionado este interesse. Em particular, estudos associados à formação de clusters de doenças e propagação de surtos infecciosos são problemas práticos bastante relevantes nessa área. A função Intensidade já se mostrou eficiente no delineamento de clusters puramente espaciais, por outro lado, a análise puramente espacial tem se mostrado ineficaz para contornar diversas situações. Na prática, a busca pelo delineamento tanto no espaço quanto no instante de início da formação do cluster são de extrema valia. Propõe-se a extensão da função Intensidade para o problema espaço temporal e a introdução de uma nova tecnologia capaz de avaliar a intensidade dos possíveis instantes de alteração efetiva no processo temporal, ou seja, delinear o possível instante de surgimento de clusters (alarme para clusters emergentes). As técnicas em avaliação serão aplicadas em bases de dados reais, com o interesse em aferir qualidade e comparar com mecanismos de detecção e inferência já existentes.

---

## CPS5: Computer intensive methods and software development

---

### CPS5.1: A new likelihood method for three-parameter gamma distribution fitting

Simplice Dossou Gbete<sup>1</sup> and Quédraogo Etienne<sup>2</sup>

<sup>1</sup> Université de Pau et des Pays de l'Adour, Laboratoire de Mathématiques et leurs Applications UMR 5142, França

<sup>2</sup> Université de Ouagadougou, Burkina Faso

**Email:** simplice.dossou-gbete@univ-pau.fr

**Abstract:** This paper deals with a Maximum likelihood method to a three-parameter gamma distribution to data from an independent and identically distributed scheme of sampling. The likelihood hinges on the joint distribution of the  $n - 1$  largest order statistics and its maximization is done by resorting to a MM-algorithm. Monte Carlo simulations is performed in order to examine the behaviour of the bias and the root mean square error of the proposed estimator. The performances of our method is compared to those of two alternatives methods.

---

### CPS5.2: A new process for modeling count time series with overdispersion

Marcelo Bourguignon Pereira<sup>1</sup>

<sup>1</sup> Universidade Federal do Rio Grande do Norte, RN, Brasil

**Email:** m.p.bourguignon@gmail.com

**Abstract:** In this paper, we propose a new first-order non-negative integer-valued autoregressive process with Poisson-Geometric marginals based on binomial thinning for modeling integer-valued time series with overdispersion. The main properties of the model are derived, such as probability generating function, moments and conditional distribution. Estimators for the parameters of process are propose and their asymptotic properties are established. Some numerical results of the estimators are presented with a discussion of the obtained results. Application to real data set are given to show the potentiality of the new process.

## CPS5.3: O modelo de regressão Weibull Modificado Exponenciado de longa duração para dados oncológicos: Estudo de simulação e desenvolvimento de uma ferramenta Web

Hayala Souza<sup>1</sup>, Gleici Perdoná<sup>1</sup>, Francisco Louzada<sup>1</sup>, Ana Paula de Andrade<sup>1</sup>,  
Francisco Barbosa Junior<sup>1</sup> e Fernanda Peria<sup>1</sup>

<sup>1</sup> FMRP e ICMC/Universidade de São Paulo, SP, Brasil

**Email:** fernandaperia@fmrp.usp.br

**Resumo:** Na área de oncologia, modelos de sobrevivência de regressão são necessários para estimar o efeito na sobrevivência de fatores como idade ao diagnóstico, condição dos linfonodos axilares, tamanho do tumor, aspectos histológicos do tumor, marcadores biológicos, entre outros. Este modelo deve incluir duas propriedades importantes: flexibilidade na função de risco e proporção de cura. Neste trabalho apresentamos um estudo de simulação para o modelo de regressão Weibull Modificado Exponenciado de Longa duração, o qual é bastante flexível do ponto de vista da função de risco e permite a consideração de que uma fração dos indivíduos não apresentem o evento de interesse. Estimativas de máxima verossimilhança e intervalares (assintótico e bootstrap) foram obtidas. Vício e erro quadrático médio, assim como probabilidade de cobertura, amplitude média e o percentual de vezes que os intervalos referentes as covariáveis continham o valor zero foram analisados. O Modelo foi aplicado a um banco de dados real de pacientes com diagnóstico de câncer de mama em estadiamento II, apresentando boa performance. Este modelo foi considerado para na construção de uma ferramenta de análise em oncologia Web desenvolvida utilizando os softwares Drupal e Shiny RStudio.

---

## CPS5.4: Ferramenta para o cálculo dos limites de controle para variáveis, baseado em R utilizando interface gráfica TclTk

Jonathan Silva<sup>1</sup>, Flaviane Peccin Brevi<sup>1</sup> e Fernando Henrique Rosa<sup>1</sup>

<sup>1</sup> Souza Cruz, RS, Brasil

**Email:** flaviane\_brevi@souzacruz.com.br

**Resumo:** Controle Estatístico de Processo (CEP) é uma abordagem amplamente utilizada nos mais diversos campos onde busca-se melhorar eficiência, produtividade além de reduzir custos e desperdícios, entre as ferramentas de CEP, no que diz respeito ao acompanhamento de medidas variáveis podemos citar a utilização de gráficos de controle como uma ferramenta de suma importância para a tomada de decisões. No presente trabalho são apresentados os principais conceitos de gráficos de controle para variáveis além do desenvolvimento e aplicação de uma ferramenta para construção dos limites de controle utilizando a linguagem de programação R.



Part IX

**Contributed Posters 1**



---

## Sessão de Pôsters 1 / Poster Session 1

---

### Análise descritiva e exploratória e o uso de outras aplicações metodológicas

*[Descriptive and exploratory statistics and the use of other methodological applications]*

---

#### CP1 – SAS codes for quantifying influence of the El Niño/Southern oscillation phenomenon on climate variable distributions

Aline Maia<sup>1</sup>

<sup>1</sup> Embrapa - Empresa Brasileira de Pesquisa Agropecuária, DF, Brasil

**Email:** aline.maia@embrapa.br

**Abstract:** El Niño/Southern oscillation (ENSO) is an oceanic/athmospheric phenomenon related to warming or cooling of the Equatorial Pacific Ocean, known to influence climate variables worldwide via teleconnections, such as rainfall, temperature or wet season timing. ENSO is measured by standardized temperature anomalies of the ocean surface, a continuous variable, usually categorized into three phases: El Niño (warm phase), Neutral or La Niña (cold phase). Here we present SAST codes for quantifying lag-relationships between ENSO predictors and monthly rainfall distributions by using a novel approach for seasonal forecasting, the proportional hazards Cox Model with continuous predictors. We developed SAST codes for: (i) fitting the Cox model and producing associated parameter estimates with respective standard errors, (ii) testing linear hypothesis on model parameters, (iii) estimating predicted probabilities of exceedance for any predictor value and (iv) plotting predictive cumulative distribution functions (CDFs) or probability of exceedance functions (PEF=1-CDF) for any particular value of ENSO derived predictors with respective confidence bands. This information is critically important seasonal forecast output, useful for planning of climate-related economic activities such as agriculture, fisheries or hydroelectric sector. As application examples, we present studies for investigating the influence of ENSO predictors on seasonal rainfall in Quixeramobim, Ceará, Brazil.

---

#### CP2 – Análise estatística em saúde reprodutiva pré e pós-gestacional de adolescentes na cidade de Campina Grande, Paraíba

Alisson Brito<sup>1</sup>, Rodolfo Silva<sup>1</sup>, Marcio Albuquerque<sup>1</sup> e Walber Colaço<sup>1</sup>

<sup>1</sup> Universidade Estadual da Paraíba, PB, Brasil

**Email:** rcrystian@gmail.com

**Resumo:** Este estudo tem como objetivo caracterizar mudanças de saúde reprodutiva pré e pós-gestacional de mulheres que foram mães na adolescência, na cidade de Campina Grande, Paraíba. Estudo transversal, com amostragem aleatória por conglomerados de 438 mulheres de 10 a 24 anos que foram mães na

adolescência e frequentaram as Unidades Básicas de Saúde de cidade de Campina Grande, de agosto a dezembro de 2014. Na análise estatística, calculou-se média, desvio padrão e significância estatística, com nível de 5% ( $p < 0,05$ ), utilizando-se o teste de McNemar. Constatou-se que 12,6% relataram iniciação sexual precoce; 91,2% das mulheres não procuraram o serviço de planejamento familiar no período pré-gestacional, 60,9% não faziam uso de contraceptivos e das que engravidaram, 87,5% procuraram o serviço de pré-natal. No período pós-gestacional, essas proporções aumentaram de forma significativa, para 43,7%, o número das que passaram a frequentar o Serviço de Planejamento Familiar, 70,8% passaram a fazer uso de métodos contraceptivos; 44,5% relataram outros filhos, com maiores proporções naquelas de 10 a 16 anos, e 22,8% referiram aborto. Após a gravidez na adolescência, foi constatada a procura do planejamento familiar e o uso de contraceptivo, além da presença de novos filhos.

---

### CP3 – Detecção sensorial de grãos precoces no café torrado e moído

Allan Fernandes<sup>1</sup>, Paulo Lima<sup>1</sup>, Mariane Rabelo<sup>1</sup>, Rosemary Pereira<sup>1</sup> e Renato R. de Lima<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil  
Email: alanfernandes538@yahoo.com.br

**Resumo:** Nas últimas décadas, o setor da cafeicultura passou por altos e baixos, tornando imprescindível a busca pela melhoria da qualidade, procurando ampliar não só o consumo de café no mercado brasileiro, mas, também, as exportações. O sabor característico do café como bebida é proveniente do grão, sendo que, um dos fatores que alteram o sabor da bebida, dando a ela um sabor herbáceo, é a mistura de grãos imaturos na torrefação. Diante disso, com o objetivo de detectar a presença de grãos precoces adicionados no café torrado e moído, foi realizado um experimento na Universidade Federal de Lavras, sendo que a detecção foi feita através de análise sensorial. Foram avaliadas as porcentagens 0, 1.54, 6.15, 10.77, 15.38, 20.00 e 24.62, correspondentes a 0, 1, 4, 7, 10, 13 e 16 grãos em 65, utilizando um painel composto por 5 provadores devidamente treinados. Os dados foram submetidos à análise de variância e em seguida, procedeu-se a análise de regressão para avaliar o comportamento das notas em função do aumento da porcentagem de grãos precoces e, além disso, foi aplicado o teste de Skott-Knott às notas médias de cada porcentagem de grãos precoces. A curva ajustada aos dados mostrou que à medida que a porcentagem de grãos precoces adicionados ao café aumenta as notas dadas pelos provadores diminui. Além disso, o teste de Skott-Knott mostrou que uma porcentagem até 6,15% não prejudicou a qualidade da bebida devido ao fato de que as porcentagens até esse valor se igualaram à Testemunha (0%).

---

### CP4 – Impact evaluation using propensity score matching for analysis of school enrollment rate data from rural Burkina Faso

Ana Clara Campos<sup>1</sup> and Leila D.A.F. Amorim<sup>1</sup>

<sup>1</sup> IM e ISC/Universidade Federal da Bahia, BA, Brasil  
Email: acpcampos27@gmail.com

**Abstract:** With increasing availability of large amounts of data, there has also been an increase in the need for and interests in the impact evaluation of various political, social and economical development

programs to assist in the political decision-making about effectiveness of public programs, to promote better allocation of resources among different programs and to provide knowledge about its functioning. In observational studies, the counterfactual causality modelling has been a strategy adopted to set assumptions about the logical structures of transmission and effects of interventions, having the main objective of estimating the causal effect (impact) of the intervention in a response variable. Propensity Score Matching (PSM) is one of the methods traditionally used to find an ideal counterfactual group in order to eliminate / reduce the confounding effects (selection bias). This paper summarizes the idea of the PSM approach and illustrates its use through an application with data regarding the school enrolment rate of girls in rural Burkina Faso. The results conclude that the BRIGHT program had a positive impact on school enrolment, indicating an increase of 10% to 16% in the number of girls enrolled in school depending on the adopted criteria and the causal effect estimator. PSM is a useful tool when it's assured that only observed factors are responsible for allocation to treatment. Potential bias problems can occur when assumptions are not followed or when the choice of observed variables to be included in the model is not made cautiously by the researcher.

---

## CP5 – Explorando tópicos de estatística básica com o GeoGebra

Ana Maria Lemes<sup>1</sup> e Adriana Strieder Philippsen<sup>1</sup>

<sup>1</sup> Universidade Estadual do Paraná, PR, Brasil

**Email:** adristrieder@yahoo.com.br

**Resumo:** Apresentamos, neste trabalho, o software GeoGebra juntamente com alguns de seus recursos para analisar estatisticamente um conjunto de dados. Consideramos um desafio explorar e empregar os recursos e potencialidades deste software na Estatística, uma vez que as experiências que tínhamos deste software eram no campo da Matemática. Como exemplos serão construídas tabelas de frequências, gráficos e cálculo das principais medidas estatísticas utilizando os recursos disponíveis do programa.

---

## CP6 – Fatores que impactam o desempenho dos alunos do 9º ano do estado de São Paulo nas provas do SAEB 2013

Angelica Cristina Custodio Cruz<sup>1</sup>, Pedro Ferreira Filho<sup>1</sup> e Estela Maris Pereira Bereta<sup>1</sup>

<sup>1</sup> DES/Universidade Federal de São Carlos, SP, Brasil

**Email:** angelica-custodio@hotmail.com

**Resumo:** O SAEB pode ser então entendido como uma parte de um processo de avaliação educacional que procura contribuir, por um lado, para a universalização do acesso à escola e, por outro, para a ampliação da equidade e da eficiência do sistema educacional brasileiro. Neste sentido, a análise de seus resultados, fornece subsídios à formulação de políticas públicas e diretrizes adequadas à diversidade de situações presentes nos estados e regiões brasileiras, constituindo-se como uma referência no que diz respeito ao desempenho escolar dos estudantes. Neste trabalho a partir do ajuste de modelos multiníveis com coeficientes aleatórios (Kreft e Leeuw, 1988) investigamos fatores que podem impactar o desempenho dos alunos do 9º ano do estado de São Paulo na prova de Língua Portuguesa SAEB 2013. Resultados

apontam que o sexo dos alunos, interesse por língua portuguesa, o tipo de escola e o acesso a recursos computacionais são alguns dos fatores que contribuem para um melhor desempenho nas provas.

## CP7 – Comparação de métodos de agrupamento em acessos de tabaco

Antonio Leandro da Silva Conceição<sup>1</sup>, Carlos Alberto da Silva Ledo<sup>2</sup>, Mauricio dos Santos da Silva<sup>1</sup>, Sandra Domingos João Afonso<sup>1,3</sup> e Ricardo Franco Cunha Moreira<sup>1</sup>

<sup>1</sup> Universidade Federal do Recôncavo da Bahia, BA, Brasil

<sup>2</sup> Embrapa Mandioca e Fruticultura, Bahia, Brasil

<sup>3</sup> Instituto Superior Politécnico do Kwanza Sul, Angola

**Email:** leandrosilvaufbr@hotmail.com

**Resumo:** O objetivo deste trabalho foi examinar diferentes metodologias de análise de agrupamento com base na análise simultânea e individual de descritores quantitativos e qualitativos selecionados de 15 acessos de tabaco. Os métodos de agrupamento utilizados foram o UPGMA e WARD, e a consistência do ajuste desses agrupamentos foram medidos pelo coeficiente de correlação cofenética. Foram utilizadas para as análises individuais as distâncias de Mahalanobis, euclidiana média e de Cole-Rodgers. Para as análises simultâneas dos dados quantitativos e qualitativos foram testadas três metodologias distintas: O algoritmo de Gower; Soma algébrica de matrizes individuais e a transformação de caracteres quantitativos em multicategóricos por duas estratégias distintas (Regra de Sturges e Regra da raiz quadrada). A estratégia da raiz quadrada foi a mais indicada, com correlação de 0,75 e 0,82 entre as matrizes de dissimilaridade dos dados codificados e quantitativos originais. Como critério para definição do número ótimo de grupos foi usado o índice Pseudo-t2, com este foi possível a formação de 3 grupos pelo método UPGMA para todas as metodologias de análise simultânea utilizadas. Todas as matrizes de análise conjunta foram comparadas, mostrando grande correspondência entre as mesmas, com correlações que variaram de 0,824 a 0,998. Os resultados mostraram que as metodologias de análise simultânea foram eficazes em revelar a existência de divergência genética entre acessos de *Nicotiana tabacum* L. tipo Sumatra e mostram a importância da combinação de métodos, uma vez que puderam otimizar de forma considerável a interpretação dos resultados para maior conhecimento do germoplasma em estudo.

## CP8 – Percepção ambiental dos moradores da cidade de Coremas-PB em relação ao esgotamento sanitário

Carlos Sérgio Araújo dos Santos<sup>1</sup>, Daniel Jackson Andrade de Sousa<sup>1</sup>, Gabriel Carlos Moura Pessôa<sup>1</sup>, Alan Dél Carlos Gomes Chaves<sup>1</sup> e Ricardo Ricelli P. de Almeida<sup>1</sup>

<sup>1</sup> Universidade Federal de Campina Grande, PB, Brasil

**Email:** carlos.santos@ccta.ufcg.edu.br

**Resumo:** O esgotamento sanitário e o termo usado para as águas que, após a utilização humana, apresentam as suas características naturais alteradas. O objetivo do estudo foi avaliar a percepção ambiental da população sobre o acesso aos serviços de esgotamento sanitário na cidade de Coremas-PB. Os dados referentes a percepção ambiental dos atores sociais foram coletados na cidade de Coremas por meio de questionários, previamente estruturados com perguntas de múltipla escolha afim de conhecer o perfil dos entrevistados através das seguintes variáveis como sexo, faixa etária, escolaridade, estado civil e renda. Conclui-se que a população da cidade de Coremas-PB, apresentaram níveis de conhecimento baixo em relação a percepção ambiental sobre o significado da importância do saneamento ambiental.

---

## CP9 – Comparação de métodos para avaliação de impacto do programa KALAHÍ na redução da pobreza nas Filipinas

Cristiane N.P. Souza<sup>1</sup> e Leila D.A.F. Amorim<sup>1</sup>

<sup>1</sup> Universidade Federal da Bahia, BA, Brasil

**Email:** crissouza.ufpa@gmail.com

**Resumo:** A percepção da pobreza passou a ser entendida como padrão de vida abaixo do que é socialmente aceitável e sua redução implica em um regime alimentar adequado, um nível de conforto e o desenvolvimento de papéis e de comportamentos socialmente adequados. Vários programas de governo têm sido propostos e implementados visando a redução da pobreza em diversos países e para sua avaliação têm sido discutidas muitas metodologias para análise do impacto de intervenções, sendo que sua escolha depende do objetivo central da análise e da disponibilidade de dados. Neste contexto, este trabalho tem como objetivo realizar uma comparação de métodos que visem a avaliação de impacto do programa KALAHÍ implementado nas Filipinas. Para tanto, utilizaram-se metodologias de pareamento, que incluem técnicas amplamente conhecidas para a estimação dos efeitos causais e o método da diferença dupla, usado para comparar as mudanças da variável resposta nos grupos ao longo do tempo em relação ao desfecho mensurado no baseline. Com base nos resultados destas análises, pode-se observar que o impacto do programa KALAHÍ nas despesas com alimentação difere segundo o método usado e que seu impacto no aumento no logaritmo das despesas com alimentos foi observado sobretudo em famílias pobres. Desta forma, a avaliação de impacto dos programas é de suma importância no sentido de subsidiar gestores públicos na inclusão social, o que contribui para a autonomia das famílias beneficiárias e condições para que as mesmas saiam da situação de vulnerabilidade em que se encontram.

## CP10 – Programas governamentais e seus efeitos na mortalidade infantil do semiárido Brasileiro

Everlane Araújo<sup>1</sup> e Neir A. Paes<sup>1</sup>

<sup>1</sup> Universidade Federal da Paraíba, PB, Brasil

**Email:** everlanesuane@hotmail.com

**Resumo:** A mortalidade infantil é considerada como um dos principais indicadores de saúde, que reflete a qualidade de vida de uma população, tendo como base o nível de desenvolvimento e acesso aos serviços de saúde. Estudos sugerem que programas governamentais são fatores adicionais importantes em relação à redução da mortalidade infantil. Assim, o principal objetivo deste estudo consistiu em avaliar o impacto dos programas sociais, Programa Bolsa Família (PBF) e Estratégia Saúde da Família (ESF), e da fecundidade sobre a mortalidade infantil da região Semiárida brasileira, no período 2005-2010. Para tanto, foi aplicado o modelo de regressão linear multivariado de dados em painel com efeitos fixos. O modelo utilizou a Taxa de Mortalidade Infantil como variável dependente e fatores associados à atenção à saúde representaram as variáveis independentes: PBF, ESF, condicionalidades de saúde do PBF e taxa de fecundidade. As evidências estatísticas foram claras ao detectar a importância motivada pelas ações públicas que objetivam auxiliar o combate de dois agravos importantes no País: pobreza e elevados níveis de mortalidade infantil. Ressalta-se que os gestores governamentais devem propiciar os meios adequados para a manutenção do acesso da população aos programas do governo.

---

## CP11 – Testes de não-inferioridade: Uma proposta para validação de equipamentos

Flaviane Peccin<sup>1</sup>, Fernando Henrique Rosa<sup>1</sup> e Jonathan Silva<sup>1</sup>

<sup>1</sup> Souza Cruz, RS, Brasil

**Email:** flavipeccin@gmail.com

**Resumo:** Os testes de não-inferioridade são utilizados na área clínica para testar se uma droga pode ser considerada não-inferior a outra, dentro de determinada margem aceitável. Nesse artigo é analisada a possibilidade de aplicação dos estudos de não-inferioridade para validação de novos equipamentos a serem introduzidos em um processo produtivo. Além disso, foram comparados quatro diferentes critérios para definição da Margem de não-inferioridade M. Os resultados analisados mostraram que os estudos de não-inferioridade podem ser considerados adequados para validar um novo equipamento a ser inserido em um processo produtivo. Dentre os quatro critérios considerados para definição de M, o que se mostrou mais adequado foi aquele conhecido como “50% rule”, que é endossado pelo FDA e consiste em assumir M como sendo 50% do limite inferior do intervalo com 95% de confiança para a diferença entre as médias de dois equipamentos em uso no processo atual.

## CP12 – Precipitação pluviométrica mensal em Uberlândia-MG: Análise descritiva e teste de normalidade

Gabriela Maria Rocha Bolaina<sup>1</sup>, Matheus Ferreira do Amaral<sup>1</sup> e Ednaldo Carvalho Guimarães<sup>1</sup>

<sup>1</sup> Universidade Federal de Uberlândia, MG, Brasil

**Email:** gabibolaina@gmail.com

**Resumo:** O estudo das distribuições de probabilidades de variáveis climáticas com a finalidade de compreender os fenômenos meteorológicos, buscando assim compreender seus padrões de ocorrência em uma determinada região é uma ferramenta de grande importância em diversas áreas de atuação humana. O objetivo deste trabalho foi descrever o comportamento da precipitação pluviométrica mensal na cidade de Uberlândia-MG, utilizando-se da análise descritiva e do teste de normalidade, sendo este o trabalho prévio e, posteriormente, serão utilizados modelos da família exponencial de distribuição de probabilidades para o ajuste aos dados da série de precipitações pluviométricas. A série utilizada corresponde ao total mensal de precipitação pluviométrica (mm) da cidade e corresponde ao período de janeiro de 1981 a 2013. As informações foram obtidas no Laboratório de Estudos Climatológicos, do Instituto de Geografia, da Universidade Federal de Uberlândia (UFU). As informações mensais foram submetidas à análise descritiva e ao teste de normalidade de Kolmogorov-Smirnov, utilizando-se o programa R. Os resultados indicaram variabilidade relativamente alta para todos os meses, sendo as maiores variações observadas nos meses de secas. O teste de normalidade indicou que os meses chuvosos tendem a distribuição normal, entretanto devido a variável ser truncada no zero posteriormente serão testados novos modelos. O conhecimento do comportamento das variáveis climáticas pode contribuir com o melhor aproveitamento dos recursos hídricos, na cidade.

## CP13 – IPESE: Uma análise da performance socioeconômica dos municípios Baianos

Jonatas Espirito Santo<sup>1</sup>, Alex Santos<sup>1</sup>, Jadson Silva<sup>1</sup> e Urandi Freitas<sup>1</sup>

<sup>1</sup> SEI /SEPLAN/ Governo da Bahia, BA, Brasil

**Email:** jonates@gmail.com

**Resumo:** No Brasil, o desenvolvimento de indicadores socioeconômicos disseminou-se a partir da segunda metade da década de 1960, para atender ao planejamento das políticas públicas durante os governos militares. A estratégia era produzir informações para acompanhar o desempenho dos programas do Governo Federal e seus desdobramentos para estados e municípios. Informações municipais só eram produzidas por meio do censo demográfico, realizado pelo IBGE com periodicidade decenal. Atualmente, com o avanço do uso dos registros administrativos, a possibilidade de criar indicadores com periodicidade anual ou bianual para avaliar e monitorar políticas públicas municipais. A Superintendência de Estudos Econômicos e Sociais da Bahia (SEI), apresenta a proposta de um novo indicador com a finalidade de ser um instrumento de monitoramento e avaliação de políticas públicas dos municípios baianos, medindo a capacidade e a qualidade com que são ofertados certos serviços básicos à população. O indicador tem como objetivo nortear o bom emprego de recursos públicos, atentando para as prioridades a serem atendidas

no recorrente comprometimento de mitigar as disparidades econômicas e sociais.

---

## CP14 – Doença de chagas aguda: Aspectos associados a confirmação do caso

José Luiz De Carvalho Lisboa<sup>1</sup>, Cristiane Nazaré Pamplona de Souza<sup>2</sup>, Edson Marcos Leal Soares Ramos<sup>1</sup> e Silvia dos Santos de Almeida<sup>1</sup>

<sup>1</sup> Universidade Federal do Pará, PA, Brasil

<sup>2</sup> Universidade Federal da Bahia, BA, Brasil

**Email:** 07joseluiz@gmail.com

**Abstract:** In Brazil reported cases of acute Chagas disease transmitted orally are related to the consumption of fresh acai juice, considered an essential food in the diet of the Northern Region's population. In view of this, this study aims to identify some signs / symptoms that are related to the classification of the case, more particularly with the confirmation of the patient with Chagas disease. Therefore, we used the binary logistic regression model. Among the main results we highlight that patients reported with acute Chagas disease is about 3 times more likely to have the confirmed case when it presents possible transmission via oral compared to those without this possible means of transmission. It was also observed that patients possibly infected by transmission via oral and showed signs / symptoms of edema, asthenia and chagoma inoculation / Romana's sign is likely to 52.22% of the case have confirmed with Chagas disease. Thus, it is clear that contamination oral and symptoms of edema, asthenia and chagoma inoculation / Roman signal are some of the factors that need the attention of health managers for the treatment of patients at an early stage, with a view the presence of these signals increases the likelihood the patient has confirmed the case of acute Chagas disease.

---

## CP15 – Relação entre transtorno de personalidade antissocial, usuários de drogas e criminalidade

Joyce Andrade<sup>1</sup>, Thomas Barbosa<sup>1</sup> e Elizeu Alves<sup>1</sup>

<sup>1</sup> Universidade Federal de Sergipe, SE, Brasil

**Email:** joyce.andrade7@gmail.com

**Resumo:** O abuso de drogas de abuso relaciona-se a autodestruição dos seus usuários, com consequentes alterações mentais que podem levar a criminalidade e violência. Adicionalmente, o seu uso prolongado pode levar ao desenvolvimento do fenômeno de transtornos de comportamento. Dessa forma, o presente estudo objetivou avaliar a ocorrência do transtorno de personalidade antissocial em usuários de drogas. Assim, este trabalho investigou a prevalência de transtorno de personalidade antissocial pacientes atendidos nos Centros de Atenção Psicossocial Álcool e Drogas (CAPS AD) localizados em Sergipe. Trata-se de um estudo regional, multicêntrico, de corte transversal. A amostra total foi composta por 103 pacientes distribuídos com alocação proporcional. Como resultado foi observado uma correlação significativa ( $p < 0,05$ ) entre o diagnóstico de transtorno de personalidade antissocial (43,69%) e um padrão de comportamento voltado a atos ilícitos nos usuários de substâncias psicoativas (63,10%). O transtorno de personalidade antissocial é considerado um dos fatores responsáveis pela associação entre uso de substâncias e crimi-

nalidade, principalmente devido ao comportamento impulsivo desses indivíduos, que apresentam prejuízo na capacidade de inibir condutas e impulsos.

---

## CP16 – Estudos probabilísticos do índice de desenvolvimento humano municipal (IDH-M) do norte de Minas Gerais (1991 a 2010) - análise via Krigagem ordinária e indicativa

Maria de Fátima Ferreira Almeida<sup>1</sup>, Jacqueline Domingues<sup>1</sup>, Vivian Aparecida Brancaglioni<sup>1</sup> e José Sílvio Govone<sup>2</sup>

<sup>1</sup> Universidade Estadual Paulista, Botucatu, SP, Brasil

<sup>2</sup> Universidade Estadual Paulista, Rio Claro, SP, Brasil

**Email:** jacqueline Domingues.unesp@gmail.com

**Resumo:** Analisou-se a variável IDH dos municípios do norte do Estado de Minas Gerais com base nos resultados dos censos de 1991 a 2010 [3]. Dados de municípios foram utilizados para análise, bem como as coordenadas geográficas dos centróides 89 cidades que compõem a região para aplicação da metodologia de Krigagem Ordinária (KO) e Krigagem Indicativa (KI). O índice de Desenvolvimento Humano (IDH) mede o nível de desenvolvimento humano dos países, utilizando como critérios indicadores: educação (alfabetização e taxa de matrícula), longevidade (esperança de vida ao nascer) e renda PIB per capita. Critério similar e utilizado para calcular o índice no âmbito municipal (IDH M). As Krigagens Ordinária (KO) e Indicativa (KI) são métodos geoestatísticos utilizados para prever valores e probabilidades, respectivamente, e gerar mapas. Neste trabalho utilizamos os métodos de Krigagens Ordinária e Indicativa para gerar mapas onde foi possível perceber as sub-regiões do Norte de Minas Gerais onde o índice IDH-M apresentaram inferior a média regional e também comparado com a média estadual e nacional e avaliar o nível de crescimento regional no período de 1991 a 2010.

---

## CP17 – Análise univariada dos índices Bovespa e Petrobras entre 2005-2015

Maria Eugênia Oliveira Batista<sup>1</sup>, Luciene Resende Gonçalves<sup>1</sup> e Gabriel Rodrigo Gomes Pessanha<sup>1</sup>

<sup>1</sup> Universidade Federal de Alfenas, MG, Brasil

**Email:** mag\_batista@hotmail.com

**Resumo:** O Índice Bovespa é o mais importante indicador do desempenho médio das cotações do mercado de ações brasileiro, retratando o comportamento dos principais papéis negociados na Bovespa. O Índice Petrobras diz respeito às cotações da Petrobras dentro da bolsa de valores de São Paulo (Bovespa). O objetivo deste trabalho será fazer uma análise temporal univariada das séries de retornos da Petrobras e do Ibovespa com a finalidade de obtenção de previsões de valores futuros. Os dados das séries foram retirados do site IPEADATA e ECONOMÁTICA, em seguida foram deflacionados, e, depois, foram calculados seus retornos no software R. Na modelagem de um retorno, modela-se a volatilidade por meio da classe dos modelos ARCH. Isso é feito em duas etapas. Na primeira, modelam-se os dados por meio do modelo ARIMA, e, na segunda, ajusta-se o GARCH aos resíduos ao quadrado. Para a série do índice Bovespa, foi

ajustado um ARIMA (3,0,3) e então o GARCH (1,1). Já para a série de retornos da Petrobrás, ajustou-se, primeiramente, um ARMA incompleto com os parâmetros autoregressivos de ordem 6 significativo e o de ordem 3 de médias móveis também significativo. A volatilidade, neste caso, também foi ajustada por um GARCH (1,1). As previsões dos dois retornos foram então obtidas. A explicação para os resultados deste trabalho podem estar relacionadas ao momento em que a Petrobras estava passando no início de 2015, período em que se intensificam as denúncias de corrupção contra a empresa, provocando uma maior variação na liquidez de suas ações.

---

## CP18 – Modelos lineares generalizados na análise da estrutura de comprimento da manjuba (*Opisthonema oglinum*) capturada no canal de Santa Cruz-PE

Nathielly Rego<sup>1</sup>, Renata Silva<sup>1</sup> e Humber Andrade<sup>1</sup>

<sup>1</sup> Universidade Federal Rural do Pernambuco, PE, Brasil

**Email:** nathiellylima1@gmail.com

**Resumo:** A manjuba (*Opisthonema oglinum*) é considerada um dos recursos pesqueiros mais explorados no litoral norte de Pernambuco, sendo a maior parte da produção procedente do estuário do Canal de Santa Cruz (CSC). A manjuba é classificada como uma espécie visitante, que não depende estritamente do estuário, mas o ocupa em certo período do ciclo de vida. O estudo da estrutura de comprimento do pescado capturado é fundamental para o entendimento da dinâmica populacional, resultando em informações importantes para o manejo da pescaria. Logo, o objetivo deste trabalho foi ampliar o entendimento sobre a dinâmica da manjuba no CSC a partir da aplicação de modelos lineares generalizados (MLG) para analisar a variação espacial e temporal do tamanho dos exemplares capturados com rede de emalhar. A maioria dos exemplares capturados é imatura, e em todas as áreas do CSC, a fração de maturos aumentou levemente ao longo do ano. É provável que isso seja um reflexo do processo de recrutamento para a pesca com entrada de juvenis no início e possível saída do canal no final do ano.

---

## CP19 – Descrição do consumo diário de ração para gatos usando modelos lineares e não lineares

Paula Ribeiro Santos<sup>1</sup>, Izabela Regina Cardoso de Oliveira<sup>1</sup>, Jéssica Santana dos Reis<sup>1</sup> e Flávia M. de Oliveira B. Saad<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** paullasant\_s@hotmail.com

**Resumo:** Os modelos não lineares têm sido muito utilizados para descrever inúmeros fenômenos físicos e biológicos. De forma geral, esses modelos são mais realistas e parcimoniosos que modelos lineares competidores, como os polinomiais. O objeto desse trabalho é descrever o consumo alimentar diário de rações para gatos, formuladas com cinco diferentes níveis de material inorgânico. Modelos lineares e o modelo não linear de Brody foram ajustados e comparados utilizando o Desvio Padrão Residual (DPR). Todas as análises foram feitas usando o programa R.

## CP20 – Cadeia de Markov aplicada a bioacústica

Rayssa Ribeiro<sup>1</sup>, João Marcelo Protazio<sup>1</sup> and Maria Luisa Silva<sup>1</sup>

<sup>1</sup> Universidade Federal do Pará, PA, Brasil

**Email:** rayssa02@hotmail.com

**Abstract:** Birds are an appropriate model to study environment, behaviour or eco-ethology, because, besides being mostly daytime, conspicuous and easy to observe in nature, birds have the taxonomy relatively well established and can be identified without the use of collection and handling. Birds still have the predominance of the use of sound for communication, and they are good ecological indicators because they are sensitive to environmental variations. This paper aims at using, as a model of study of vocal behaviour, a kind of bird that shows a sign of complex communication sound, such as the white-sabiá *Turdus leucomelas*, a species which is very common in cities and around the brush woods in Belém, Pará, which presents a complex melodious song. The possible influence of anthropized structure of the *T. leucomelas*'s song will be evaluated, compared to the local urban subjects and residents in barns and edges of the woods. We will record the singings of several individuals in degraded environments and better preserved with professional quality equipment.

---

## CP21 – Acidente de trânsito de motocicleta e carro em João Pessoa

Rodolfo Silva<sup>1</sup>, Alisson Brito<sup>1</sup> e Mácio Albuquerque<sup>1</sup>

<sup>1</sup> Universidade Estadual da Paraíba, PB, Brasil

**Email:** rcristian@gmail.com

**Resumo:** A Organização Mundial de Saúde estima que em 2020 os acidentes de trânsito serão a 2<sup>a</sup> causa de morte prematura no mundo. A cidade de João Pessoa, Brasil foi a região selecionada para o estudo. Os dados foram agrupados por área (bairros). A ferramenta de análise foi o agrupamento. Os resultados mostraram que as motocicletas representam 25,46% da frota cadastrada na cidade segundo dados do DENATRAN. Envolveram-se principalmente em colisão/abaloamento (35%). A média mensal de ocorrência de AT é de 254. O mês que apresentou maior contribuição foi agosto com 309 e por dia da semana foi o domingo com 558. Quanto à faixa horária, o pico ocorreu no intervalo de 18h a 19h com 209 ocorrências. As vítimas envolvidas nos acidentes de trânsito envolvendo as motocicletas eram na maioria do sexo masculino (76,83%). A faixa de idade foi de 18 a 29 anos (47,45%). Na análise espacial segundo a faixa de idade, destacou-se o bairro de Mangabeira com 292 ocorrências.

---

## CP22 – Panorama dos três corredores com maior número de acidentes na cidade de Campina Grande-PB

Sonia Santos<sup>1</sup>, Wanessa Wanderley<sup>1</sup>, Tanise Souza<sup>1</sup>, Eduarda Bizerra<sup>1</sup> e Ricardo Olinda<sup>1</sup>

<sup>1</sup> Universidade Estadual da Paraíba, PB, Brasil

**Email:** tanise\_kely@hotmail.com

**Resumo:** Hoje com o avanço da tecnologia, os veículos, antes objetos restritos a uma determinada classe social de maior poder aquisitivo na sociedade, passam a ser item integrante do patrimônio de diversas famílias. Com isso, aumentando a frota veicular e conseqüentemente os acidentes. Os dados foram fornecidos pela Superintendência de Transito e Transportes Públicos (STTP) com a parceria dos órgãos: CPTran e SAMU, da cidade de Campina Grande-PB, no ano de 2013. A proposta deste trabalho é aplicar a estatística de teste Qui-quadrado para verificar a associação entre as variáveis independentes e em seguida utilizando o teste da razão de chances (Odds Ratio). Para realização dos resultados os dados foram submetidos a uma ferramenta computacional estatística R.

---

## CP23 – Redução da mortalidade infantil no nordeste Brasileiro para os anos de 2010 e 2010 e perspectivas até 2015

Vanessa Santos<sup>1</sup>, Suy Santos<sup>1</sup> e Mateus Reis<sup>1</sup>

<sup>1</sup> Universidade Federal de Sergipe, SE, Brasil

**Email:** vks.ufs@gmail.com

**Abstract:** The United Nations includes the reduction of child mortality as one of the Millennium Development Goals. Based on the study of the issues that influence the process of access to prenatal care can implement public policies that help reduce the mortality rate being so, this study aimed to evaluate the behavior of the regions, with emphasis on the Northeast through the infant mortality rate in the years 2000 to 2010, obtained based on the percentage of reduction by 2015 calculated. In this study, we identify some external impacts that influence the studied indicator, as well as the setting it was in the state of Sergipe. Through rate were measured the health and life of the regions of interest in that, from the window it is possible to know the level of social development of the population. Finally, adjust models Dummy's were used to design rates from 2011 to 2015, to identify possible trends and compared the region / Northeast Federation Unit with greater deceleration of growth, aiming, albeit rustic, if the goal of the millennium will be achieved by the year reference – 2015. It is worth noting that the reduction in infant mortality Neonatal pact in the Northeast and in the Amazon, created in 2009 by the Federal Government to reduce regional disparities, directly influenced the decrease rate of infant mortality mainly in the Northeast, but assisting in the achievement of this goal.

---

## Sessão de Pôsters 1 / Poster Session 1

---

### Modelos de Regressão *[Regression Models]*

---

#### CP24 – Testes de Paralelismo: Aplicação para comparação de curvas analíticas de água

Adson Costanzi Filho<sup>1</sup>, Flaviane Peccin<sup>2</sup>, Thais Aline Pohren Ferreira<sup>2</sup> e Jonathan Silva<sup>2</sup>

<sup>1</sup> Universidade Federal do Rio Grande do Sul, RS, Brasil

<sup>2</sup> Souza Cruz, RS, Brasil

**Email:** adson\_filho@souzacruz.com.br

**Resumo:** A comparação de curvas analíticas pode ser de interesse de pesquisadores para verificar diferenças de fatores de preparo como operador, dia, armazenamento e etc. Neste artigo, é apresentada a aplicação de uma metodologia para avaliar a estabilidade das curvas analíticas de água, através de uma comparação entre seus coeficientes angulares. As soluções comparadas foram armazenadas em diferentes temperaturas ( $22\pm 2^\circ\text{C}$ , que foi chamada de ambiente e outra em temperatura controlada em  $5\pm 3^\circ\text{C}$ ) e os resultados analisados mostraram que após quatro dias de armazenamento há deterioração das soluções, visto que detectou-se diferença significativa entre os coeficientes angulares das curvas analíticas de água.

---

#### CP25 – Aplicações da decomposição espectral da matriz de parentesco no modelo misto poligênico

Alex Ribeiro<sup>1</sup>, Daniel F. Ferreira<sup>1</sup> e Júlia Pavan<sup>2</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> Universidade de São Paulo, SP, Brasil

**Email:** danielff@dex.ufla.br

**Resumo:** A herdabilidade representa um importante indicador para os pesquisadores na área da saúde, pois permite identificar quais caracteres possuem efeito genético pronunciado na incidência e severidade de uma determinada doença. Desta forma, o presente trabalho tem por objetivo estudar, via simulação Monte Carlo, o comportamento da estatística de teste da herdabilidade proposta por Blangero et al. (2013) e avaliar a utilização dos autovalores e autovetores da matriz de covariâncias genética para selecionar indivíduos mais informativos em delineamentos com diferentes estruturas de famílias. Nas simulações, além de estruturas fictícias, é considerada a estrutura das famílias do projeto Corações de BaependiMG que apresenta um alto grau de desbalanceamento, com famílias de três a cerca de 100 membros. A metodologia é implementada computacionalmente usando os recursos do aplicativo R. Os primeiros resultados,

obtidos por simulação com uma estrutura familiar simples (família nuclear) e constituída por 8 indivíduos ( $n=80$ ), demonstraram que a estatística de teste avaliada comportou-se muito próxima de uma qui-quadrado distribuída assintoticamente conforme relatado por Self & Liang, 1987.

---

## CP26 – Aplicação e comparação de MLG para precificação do contrato do seguro agrícola no estado de Sergipe

Amanda da Silva Lira<sup>1,2</sup>, Luelton Nunes Santos<sup>1</sup>, Giovani Oliveira Santos<sup>1</sup>, Elielma Santana de Jesus<sup>1</sup> e Kleber Regis Santoro<sup>2</sup>

<sup>1</sup> Universidade Federal de Sergipe, SE, Brasil

<sup>2</sup> PPGBEA/Universidade Federal Rural de Pernambuco, PE, Brasil

**Email:** amandalira@ufs.br

**Resumo:** O presente trabalho tem como principal objetivo propor a realização do cálculo da precificação do seguro agrícola para cultura do milho no estado de Sergipe, através de modelos lineares generalizados e realizar comparação entre distribuições paramétricas que melhor estimem o valor mais justo para precificação do seguro agrícola em Sergipe. Varias são as dificuldades das seguradoras para estabelecer uma forma correta para a precificação, pois, não existe uma teoria atuarial adequada voltada para o ramo. Nesse estudo, foram coletados dados dos anos de 1973 até 2010, nos 75 municípios do estado de Sergipe, disponível no SIDRA - Sistema IBGE de recuperação automática. Realizou o teste de Shapiro-Wilk como uma forma para verificar a normalidade dos dados. Foram aplicados para o cálculo da precificação do seguro agrícola a utilização de modelos paramétricas através de modelos lineares generalizados e ajustados as distribuições Normal, Gama e Normal Inversa para indicar qual modelo apresenta melhor estimativa para o cálculo da precificação do seguro agrícola em Sergipe. Foi observado que o modelo Normal e o modelo Gama apresentaram melhores estimativas na maioria dos municípios.

---

## CP27 – Diagnostic accuracy and inter-observer agreement measures to dependent data

Ana Carolina Cintra Nunes Mafra<sup>1</sup>, Elivane da Silva Victor<sup>1</sup> and Thaís Caldara Mussi Andrade<sup>1</sup>

<sup>1</sup> Hospital Israelita Albert Einstein, SP, Brasil

**Email:** anacarol.nunes@gmail.com

**Abstract:** Background: An important area of medical research is the study of diagnosis methods. When comparing new methods with existing gold standards are used measures such as: accuracy, sensitivity, specificity, and positive and negative predictive values. Studies that consider more than one observer still need to measure inter-observer agreement. Usually, the studies intended to infer about them through statistical analysis using tools developed under the assumption that the data are independent. Not all medical studies fulfill this assumption, for studying the various teeth of the same individual, their two knees or breasts, for example. Objective: Obtain appropriate statistical techniques to infer about accuracy and agreement to dependent data and illustrate with real data. Data: We used data from a study that compare two diagnostic imaging (multiparametric MRI exams of the prostate - with and without contrast)

with the golden standard (prostate biopsy). The suspicion level is based on the Likert classification for clinically significant prostate cancer in a sextant pattern and the MRI exams were read by two radiologists. Statistical methods: Were used generalized estimation equations to attain estimated proportions, 95% confidence intervals, inter-observer agreement coefficient (Kappa) and p-values and compare them with traditional tools. We used R 3.0.3. Results and Discussion: The largest differences were observed in estimates standard errors, generating differences in p-values and confidence intervals. Some estimates have been over or under estimated with the traditional method. Conclusion: By using statistical techniques is very important to consider the assumptions which, if not met, could result in biased results.

---

## CP28 – Uso da transformação isotônica no ajuste do acúmulo de massa seca por modelo de regressão não linear e análise da variância

Augusto Morais<sup>1</sup>, Laís Silva<sup>1</sup>, Messias J.B. Andrade<sup>1</sup>, Dâmiany Oliveira<sup>1</sup>, Elisa Morais<sup>2</sup> e Thaís Ribeiro<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> Faculdade de Jaguariúna, SP, Brasil

**Email:** elisa.alimentos@gmail.com

**Abstract:** The objective this paper was of the use of isotonic transformation as an alternative to model the growth in dry matter in common bean cultivar, under different plant densities, using the nonlinear regression model logistic. The application of isotonic transformation in the data of dry matter in function on the evaluation periods proved to be feasible, by providing reduction in the experimental error, the residual standard deviation and Akaike information criterion.

---

## CP29 – Uma comparação entre modelo fuzzy e regressão linear na relação entre indicadores de capital de giro e o risco da empresa

Beatriz Rezzieri Marchezini<sup>1</sup>, Lucas Pereira Lopes<sup>1</sup> e Gabriel Rodrigo Gomes Pessanha<sup>1</sup>

<sup>1</sup> Universidade Federal de Alfena, MG, Brasil

**Email:** bia\_marchezini@hotmail.com

**Resumo:** Analisar quais variáveis possam influenciar o risco das empresas é uma nova área dentro das finanças corporativas, que tem por essência um papel fundamental na gestão das organizações. Assim, este trabalho tem por objetivo verificar a relação entre os indicadores de capital de giro e o risco das empresas (beta) do modelo CAPM das companhias de capital aberto do segmento de siderurgia, no período de 1998 a 2015. Para tanto e em jus ao objetivo, utilizou-se dois modelos matemáticos aplicados à gestão, a regressão linear múltipla e o sistema baseado em regras fuzzy. Destaca-se que o modelo de regressão linear múltipla é uma ferramenta consagrada no cálculo de relação de variáveis, permitindo uma análise mais quantitativa, já o sistema baseado em regras fuzzy tem sido destacado na literatura, principalmente na área de gestão, como um modelo promissor para essa mesma verificação, possibilitando estabelecer comparações. O software utilizado foi o R Studio 2015. Através dos resultados obtidos, é notória a aplicabilidade dos dois modelos na descrição dos dados, no entanto, o modelo baseado em regras fuzzy

se sobressaiu sobre o de regressão linear múltipla.

---

## CP30 – Regressão logística com dados repetidos aplicada à modelagem de incontinência urinária feminina

Cecilia Candolo<sup>1</sup>, Ana Carolinne Portela Rocha<sup>2</sup> e Fernanda Vieira Rodovalho Callegari<sup>3</sup>

<sup>1</sup> DEs/Universidade Federal de São Carlos, SP, Brasil

<sup>2</sup> DFisio/Universidade Federal de São Carlos, SP, Brasil

<sup>3</sup> DMed/Universidade Federal de São Carlos, SP, Brasil

**Email:** cecilia@ufscar.br

**Resumo:** A incontinência urinária é uma condição altamente prevalente na população feminina embora permaneça, muitas vezes, subdiagnosticada e subtratada. Na Estratégia Saúde da Família (ESF) médicos e enfermeiros com formação generalista são os responsáveis pela assistência uroginecológica. Um estudo realizado no interior de São Paulo objetivou identificar as lacunas de conhecimentos, atitudes e prática de médicos e enfermeiros da ESF, a partir da aplicação de um questionário a profissionais da área. Nesse questionário, a seção Conhecimentos é composta por nove questões baseadas em uma suposta consulta que, por construção, estão agrupadas em 4 grupos. O objetivo do presente trabalho é modelar a probabilidade de acerto das questões nos grupos em função de variáveis como formação profissional, tempo de formação, tempo de atuação na ESF, tipo de vínculo empregatício, sexo e idade. A abordagem utilizada foi a de efetuar uma análise para cada grupo, considerando a resposta (acerto ou erro) dentro dos grupos como medidas repetidas, ou seja, dados correlacionados. A análise foi feita utilizando a abordagem de equações de estimação generalizadas e efetuada utilizando o genmod do SAS, na opção OnDemand. Os resultados corroboraram com a abordagem utilizada: para cada grupo, o conjunto de covariáveis significativas foi diferente. No grupo 1 destacou-se a significância (5%) de sexo, no grupo 2 foi significativa a profissão, idade, vínculo empregatício e tempo de formação. Já no grupo 3 houve apenas significância a 10% para tempo de atuação na ESF e, para o grupo 4, houve significância para idade e tempo de formação.

---

## CP31 – Estudo da degradabilidade in situ da fibra em detergente neutro utilizando modelos não lineares mistos

César Gonçalves de Lima<sup>1</sup>, Euclides Braga Malheiros<sup>2</sup>, Simone Daniela Sartorio<sup>3</sup> e José Valmir Feitosa<sup>4</sup>

<sup>1</sup> FZEA/Universidade de São Paulo, SP, Brasil

<sup>2</sup> Universidade Estadual Paulista, SP, Brasil

<sup>3</sup> CCA/Universidade Federal de São Carlos, SP, Brasil

<sup>4</sup> Universidade Federal do Ceará, CE, Brasil

**Email:** cegdlima@usp.br

**Resumo:** Dados de degradabilidade in situ são usados na avaliação do desempenho de alimentos que compõem a dieta de um ruminante. Surgem de experimentos com medidas repetidas que podem exibir alguma correlação serial e heterocedasticidade de variâncias. O presente trabalho evidenciou as vantagens

---

do uso de modelos não lineares mistos no estudo da degradabilidade da fibra em detergente neutro (FDN) de seis rações experimentais, compostas pela combinação de três diferentes proporções de volumoso: concentrado e de dois fenos de capim-Tifton 85 de diferentes qualidades. O uso do proc nlmixed do SAS facilitou a formulação dos modelos, possibilitou o estudo visando a inclusão de efeitos aleatórios e a realização dos testes sobre os efeitos dos dois fatores de tratamento sobre os parâmetros das curvas médias de resposta.

---

## CP32 – Uso de modelos não lineares para modelagem da área basal e volume de *Eucalyptus* sp.

Elliezer de Almeida Melo<sup>1</sup>, Joyce de Almeida Alves<sup>1</sup>, Ernani Lopes Possato<sup>1</sup>,  
Marcos Antônio Isaac Jr<sup>1</sup>, Ernandes Guedes Moura<sup>1</sup> e Natalino Calegario<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil  
Email: elliezymelo@hotmail.com

**Resumo:** Este trabalho teve por objetivo avaliar modelos não lineares simples e completos pela adição de covariáveis, para modelagem da área basal e do crescimento e produção em volume de clones de eucalipto. Para o estudo foram utilizados dados de 90 unidades amostrais medidas entre 2010 e 2012, distribuídas em diferentes idades e locais. Foram ajustados o modelo assintótico de dois parâmetros desenvolvido por Bates e Watts (1988) para a predição da área basal e o modelo logístico de três parâmetros para predição do volume. Para ambos ajustes os parâmetros dos modelos foram decompostos para inclusão de variáveis do povoamento que influenciam no crescimento e produção. A capacidade produtiva do local (S) e o número de árvores por hectare (N) foram as covariáveis inseridas no modelo para predição da área basal (G). Para a predição do volume além dessas variáveis utilizou-se também a G. A seleção e a qualidade de ajuste dos modelos se deram com base no erro padrão percentual Syx (%), critério de informação de Akaike (AIC), critério de informação Bayesiano (BIC) e no teste de razão da máxima verossimilhança (TRMV), além das análises gráficas dos resíduos e da altura estimada versus observada. A manipulação, ajuste e processamento dos dados foram realizados no software R. Os modelos com adição de covariáveis foram superiores em relação aos modelos simples, com redução de aproximadamente 3% e 11% no valor do erro padrão residual para área basal e volume, respectivamente.

---

## CP33 – Uso de modelos não lineares para modelagem do crescimento das árvores dominantes e codominantes de *Eucalyptus* sp.

Elliezer de Almeida Melo<sup>1</sup>, Joyce de Almeida Alves<sup>1</sup>, Ernani Lopes Possato<sup>1</sup>,  
Marcos Antônio Isaac Jr<sup>1</sup>, Gabriela Cristina Costa Silva<sup>1</sup> e Natalino Calegario<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil  
Email: elliezymelo@hotmail.com

**Resumo:** Este trabalho teve por objetivo avaliar dois modelos não lineares para a modelagem do crescimento médio em altura das árvores dominantes e codominantes (HDC) e construção de curvas de sítio em povoamentos de eucalipto. Para o estudo foram utilizados dados de 90 unidades amostrais medidas entre 2010 e 2012, distribuídas em diferentes idades (1,7 a 6,9 anos) e locais. Foram testados o modelo

de Chapman e Richards de três parâmetros e o modelo de Schumacher de dois parâmetros. A seleção e a qualidade de ajuste dos modelos se deram com base no erro padrão residual [Syx (%)], nas análises gráficas dos resíduos e da altura estimada versus observada. A manipulação, ajuste e processamento dos dados foram realizados no software R. De acordo com as estatísticas de ajuste, os modelos atenderam as pressuposições da análise de regressão. O modelo de Chapman e Richards apresentou leve superioridade em comparação ao modelo de Schumacher, com erro padrão residual de 5,78% e 5,68%, respectivamente. Sendo o melhor modelo, o modelo de Chapman e Richards foi utilizado para a estimativa da HDC e construção das curvas de índice de sítio. Foram definidas três classes de sítio, sendo a classe I (33,5 m - 37,5 m) a mais produtiva, a II (29,5 m - 33,4 m) a de média produtividade e a classe III (25,5 m - 29,4 m) a de baixa produtividade. As alturas médias das árvores dominantes e codominantes estão, em maior frequência, nos sítios I e II.

---

## CP34 – Análise do desempenho da probabilidade de cobertura para o modelo Garma com distribuição condicional de Poisson

Elza Blank<sup>1</sup>, Adriana Strieder Philippsen<sup>1</sup> e Marinho G. Andrade<sup>2</sup>

<sup>1</sup> Universidade Estadual do Paraná, PR, Brasil

<sup>2</sup> Universidade de São Paulo, SP, Brasil

**Email:** adriestrieder@yahoo.com.br

**Resumo:** Este trabalho contém os resultados parciais de um projeto de iniciação científica que ainda se encontra em andamento. A motivação do nosso trabalho é verificar o desempenho da probabilidade de cobertura para o modelo Poisson GARMA fazendo variar o tamanho da amostra e a ordem do modelo. Quando construímos um intervalo de confiança desejamos que a verdadeira probabilidade de cobertura esteja próxima do nível de confiança pré-estabelecido. Por probabilidade de cobertura entende-se que é a proporção de intervalos gerados que contém o verdadeiro valor do parâmetro, o qual é fixado ao realizar o estudo de simulação. O ideal é que esta taxa de cobertura dos intervalos de confiança esteja muito próxima do nível estipulado independente do valor do parâmetro. Para atingir essa motivação foi necessário um estudo de simulação, pois o mesmo permite investigar, analisar e descrever a performance ou comportamento de um sistema e ao utilizarmos essa técnica em estudos estatísticos, permite que seja possível realizar uma análise preliminar de dados, de modelação, estimação de parâmetros bem como da probabilidade de cobertura.

---

## CP35 – Análise do ciclo reprodutivo das lagostas vermelha (*Panulirus Argus*) e verde (*Panulirus Laevicauda*) com modelos lineares generalizados

Evelyn Oliveira<sup>1</sup> e Humber Andrade<sup>1</sup>

<sup>1</sup> Universidade Federal Rural de Pernambuco, PE, Brasil

**Email:** humber.andrade@gmail.com

**Resumo:** No Brasil, a sustentabilidade da pesca de lagosta é baseada em diversas regulamentações, sendo que uma das mais tradicionais é o defeso, que em princípio tem a finalidade de proteger o estoque no

período reprodutivo. Porém as análises que levaram às decisões de defeso são fragmentadas, e o conjunto de informações não foi explorado de maneira integrada. Neste trabalho as séries temporais foram avaliadas de forma completa, através do uso dos modelos lineares generalizados, visto que as decisões que levaram ao estabelecimento do período de defeso foram baseadas apenas em análises descritivas e inspeções visuais. Os dados utilizados para este estudo foram recompilados de uma série de artigos publicados por Paiva e Costa (1963, 1964, 1965); Paiva (1966, 1967); Paiva e Costa (1968, 1969, 1970). Inicialmente, foi efetuada uma análise exploratória dos dados e em seguida aplicado os modelos lineares generalizados na série temporal integral das informações disponíveis para elucidar os padrões nas variações das proporções de lagostas em reprodução. Para a seleção das variáveis explicativas e a escolha do modelo mais adequado utilizou-se o critério de informação de Akaike (AIC) e testes de hipóteses baseado na análise da deviança.

---

## CP36 – Modelos sigmoidais para descrever o afilamento do fuste de árvores de eucalipto

Joyce de Almeida Alves<sup>1</sup>, Ernani Lopes Possato<sup>2</sup>, Geisi Azevedo Silva<sup>3</sup>, Elliezer de Almeida Melo<sup>4</sup> e Natalino Calegario<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> Instituto Bioterra, Brasil

<sup>3</sup> Universidade Federal de Uberlândia, MG Brasil

<sup>4</sup> Instituto Federal Goiano, GO, Brasil

**Email:** joyce\_a.alves@yahoo.com.br

**Resumo:** As funções de afilamento são fundamentais para a quantificação de multiprodutos. Com elas, é possível estimar o diâmetro a qualquer altura do fuste, a altura comercial para qualquer diâmetro, o volume total do tronco e também o volume de determinadas seções do fuste com comprimentos desejados. O objetivo deste estudo foi comparar a qualidade do ajuste entre a função de afilamento sigmoidal, proposta por Biging (1984), e outra função sigmoidal, pouco utilizada na área florestal para descrever o afilamento do fuste das árvores de eucalipto. Os dados foram obtidos da cubagem rigorosa de 92 árvores de um plantio clonal localizado no sudeste de Goiás. As equações foram comparadas pelo erro padrão residual (%) e gráfico de resíduos. Para comparar a habilidade das equações em prever diâmetros ao longo do fuste calculou-se: Bias, Média das diferenças absolutas (MDA), Desvio padrão das diferenças (SD) e Soma dos quadrados dos resíduos relativos (SSRR) para dados independentes. Apesar das diferenças entre as duas equações testadas não terem sido tão acentuadas, a função de afilamento proposta por Biging (1984) foi mais adequada para representar o afilamento das árvores de eucalipto da região estudada, pois apresentou menor valor de Syx (%), melhor distribuição de resíduos e melhor habilidade de prever diâmetros para a maioria das porções do fuste das árvores. Além de ser mais parcimoniosa.

---

## CP37 – Testes de hipótese para os parâmetros do modelo de regressão logística

Juliana Maia<sup>1</sup> e Verônica Lima<sup>1</sup>

<sup>1</sup> Universidade Federal da Bahia, BA, Brasil

**Email:** vmcadena@gmail.com

**Resumo:** A estimação dos parâmetros nos modelos de regressão logística é usualmente baseada na teoria de máxima verossimilhança. As estatísticas razão de verossimilhanças, Wald e escore são usadas para testar hipóteses relativas aos parâmetros deste modelo. Os valores críticos destes testes são obtidos a partir de aproximações válidas em grandes amostras. Em geral, há uma considerável distorção do tamanho dos testes em pequenas amostras. O objetivo deste trabalho é comparar o resultado dos testes da razão de verossimilhanças, razão de verossimilhanças corrigido de Bartlett e razão de verossimilhanças bootstrap considerando o modelo de regressão logística através da aplicação a um conjunto de dados reais. Os dados consistem em um estudo epidemiológico sobre distúrbios musculoesqueléticos em trabalhadores de fábricas de calçados no estado da Bahia. Os resultados mostram que os testes da razão de verossimilhanças e da razão de verossimilhanças corrigido apresentam resultados similares. Porém, o teste da razão de verossimilhanças bootstrap produz resultados diferentes.

---

## CP38 – Utilização de modelos não lineares na descrição do acúmulo de boro no feijoeiro

Kelly Pereira de Lima<sup>1</sup>, Augusto Ramalho Morais<sup>1</sup>, Messias José Bastos de Andrade<sup>1</sup>, Sidcleide Barbosa de Sousa<sup>1</sup> e Dâmianny Pádua Oliveira<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** kelly\_limaadm@hotmail.com

**Resumo:** O estudo do crescimento do feijoeiro por meio do acúmulo do micronutriente boro ao longo do tempo é uma ferramenta informativa para o auxílio do manejo adequado e detecção de fatores que afetam o desenvolvimento da planta. Com isso, a modelagem desse tipo de crescimento pode subsidiar melhor conhecimento e aplicação eficiente das práticas agrícolas. Desta maneira, este trabalho teve como objetivo o ajustar os modelos não lineares logístico, Von Bertalanffy e exponencial usando os dados originais e transformados do acúmulo de boro, em dois sistemas de plantio e três densidades de semeadura para o feijão cultivar Jalo, o experimento foi realizado na Universidade Federal de Lavras, em um delineamento em bloco casualizados, contendo três repetições; em cada sistema de plantio usou-se o esquema fatorial 3 x 7, sendo constituído por três densidades de semeadura (75, 215 e 355 mil planta por hectares) e sete épocas de avaliação (13, 23, 33, 43, 53, 63, 73 dias após emergência). Foi analisado o acúmulo de boro nas partes das plantas hastes; haste e folhas; haste, folha e vagem e total. Os modelos não lineares logístico e von Bertalanffy são adequados para a descrição do acúmulo de boro no feijoeiro cv Jalo. A transformação de dados utilizada assegurou a presença de normalidade e homogeneidade dos erros, porém os dados alteraram a característica de acúmulo em função do tempo, necessitando usar o modelo exponencial para representá-los. Com isso, a validade da análise da ANOVA, assim, assegurada.

---

## CP39 – Modelos lineares hierárquicos normais

Kelly Pereira de Lima<sup>1</sup>, Sidcleide Barbosa de Sousa<sup>1</sup>, Ronald Targino Nojosa<sup>2</sup> e Augusto Ramalho Morais<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> Universidade Federal do Ceará, CE, Brasil

**Email:** kelly\_limaadm@hotmail.com

**Abstract:** There are situations where data have investigated a hierarchy structure or moiety or are characterized by the presence of aggregated units in other experimental larger units . This article aims to present the hierarchical linear model structure.

---

## CP40 – Comparação dos modelos de Gompertz e Verhulst no ajuste do acúmulo de nitrogênio no feijoeiro cv. Jalo

Laís Silva<sup>1</sup>, Allan Fernandes<sup>1</sup>, Thaís Ribeiro<sup>1</sup>, Augusto Morais<sup>1</sup>, Joel Muniz<sup>1</sup>,  
Messias J.B. Andrade<sup>1</sup> e Neiva Vieira<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** laismesquita@hotmail.com

**Resumo:** Com o objetivo de comparar por meio dos modelos não lineares de Verhulst e Gompertz o acúmulo do nutriente nitrogênio durante o crescimento do feijoeiro, cv. Jalo, foi conduzido um experimento na Universidade Federal de Lavras, empregando-se o delineamento em blocos casualizados, com três repetições. Os tratamentos foram dispostos em esquema fatorial  $5 \times 7$ , envolvendo cinco densidades de semeadura (75, 145, 215, 285 e 355 mil plantas por  $ha^{-1}$ ) e sete períodos (13, 23, 33, 43, 53, 63 e 73 dias após a emergência) de avaliação. Os dados de acúmulo total de nitrogênio (kg/ha) foram submetidos a análise da variância e em cada densidade de semeadura foi utilizada a análise de regressão para obter as curvas de crescimento, em acúmulo de nitrogênio, em relação aos períodos de avaliação, com a utilização dos modelos não lineares de Gompertz e Verhulst. Utilizou-se o critério de Akaike para identificar que o modelo de Verhulst foi o que apresentou melhor ajuste.

---

## CP41 – Identificação dos fatores de riscos associados à ocorrência de anomalias congênitas em recém nascidos utilizando a regressão logística

Lidia Dayse Araujo de Souza<sup>1</sup>, Isis Milane Batista de Lima<sup>1</sup>, Ianne Rafaella Santos  
Melo<sup>1</sup> e Jessica da Silva Braga<sup>1</sup>

<sup>1</sup> Universidade Federal da Paraíba, PB, Brasil

**Email:** lidiadayse@hotmail.com

**Resumo:** Os modelos de regressão logística vêm sendo aplicado intensamente em várias áreas de conhecimento e, em especial, na área da saúde. A condição de se estudar variáveis respostas binárias em função de um conjunto de fatores explicativos tem se tornado cada vez mais comum em estudos epidemiológicos. Assim, o presente estudo tem como objetivo utilizar um modelo de regressão logística para investigar os fatores de risco associados à ocorrência de malformação congênita em crianças de um hospital de João Pessoa – PB. Os resultados mostraram que a idade, a escolaridade da mãe, o uso de corticóides durante a gravidez, o tipo de parto e as medidas de APGAR de 1 e 5 minutos estavam associadas com a probabilidade de nascimento de filhos com alguma anomalia congênita. Além disso, constatou-se que o modelo ajustado conseguiu classificar corretamente mais de 93% dos casos examinados.

## CP42 – Modelo linear misto na análise de perfis de pesos de filhotes de ratos com mudança acentuada de tendência

Luzia Pedroso de Oliveira<sup>1</sup>, César Gonçalves de Lima<sup>2</sup>, Vera Lúcia Scherholz Salgado de Castro<sup>3</sup> e Aline de Holanda Nunes Maia<sup>3</sup>

<sup>1</sup> Universidade Federal de São Carlos, SP, Brasil

<sup>2</sup> FZEA/Universidade de São Paulo, SP, Brasil

<sup>3</sup> EMBRAPA Meio Ambiente, SP, Brasil

**Email:** professoraestatisticaluzia@yahoo.com.br

**Resumo:** Nos experimentos para avaliar a toxicidade de um produto ou mesmo efeitos de dietas são geralmente utilizados ratos, sendo usual observar e pesar os animais diariamente ao longo de algum período. O efeito do tratamento na variação de peso é uma característica importante na detecção de toxicidade, sendo comumente discutido nos artigos da área. Na análise de dados longitudinais o modelo linear misto mostra-se bastante apropriado, pois possibilita considerar as correlações entre as observações coletadas num mesmo indivíduo, assim como outros padrões de variação como o aumento da variabilidade ao longo do tempo. O objetivo deste trabalho é apresentar uma aplicação do modelo linear misto para testar os efeitos de dois fatores: doses de um herbicida e fases de gestação em que as doses são aplicadas nas ratas, sobre o crescimento de seus filhotes ao longo da lactação. Uma vez que os perfis médios de pesos dos filhotes, apresentaram uma mudança de tendência acentuada após a idade de abertura dos olhos, utilizou-se spline e regressão por partes, ambos quadráticos e com um único ponto de mudança fixo, para descrever as alterações dos pesos ao longo do tempo. Essas duas abordagens possibilitam comparar o ganho médio de peso nos tratamentos e a aceleração nos dois períodos de crescimento mais e menos acentuados, atendendo aos objetivos do pesquisador. Foram considerados efeitos aleatórios de rata e de filhote e diferentes estruturas para a matriz de covariâncias dos erros, sendo os modelos comparados por meio de critérios de informação e de análise de resíduos.

## CP43 – Modelo logístico para ajuste de curvas de crescimento de *Eucalyptus* sp.

Magali Teresopolis Reis Amaral<sup>1</sup> e Carlo Roberto Padovani<sup>1</sup>

<sup>1</sup> Universidade Estadual Paulista, SP, Brasil

**Email:** mteresopolis@hotmail.com

**Resumo:** O presente trabalho consiste em utilizar o modelo logístico de crescimento para relacionar funcionalmente a altura e a idade de árvores da espécie *Eucalyptus* sp. Para a aplicação do procedimento estatístico considerou-se uma amostra de 300 árvores com idade entre 12 e 70. No contexto metodológico utilizou-se o algoritmo de simulação de Monte Carlo em Cadeia de Markov (Metropolis-Hastings) com abordagem bayesiana com densidades a priori não informativa de Jeffreys. O procedimento utilizado permitiu estimativas consistentes da altura da árvore de eucalipto em função da idade.

## CP44 – Comparação dos modelos individuais e híbridos de Credit Scoring utilizando regressão logística e rede neural

Mariana Nolde Pacheco<sup>1</sup> e Lisiane Priscila Roldão Selau<sup>1</sup>

<sup>1</sup> Universidade Federal do Rio Grande do Sul, RS, Brasil

**Email:** marinolde@yahoo.com.br

**Resumo:** Nas últimas décadas, houve um crescimento progressivo na demanda e distribuição de crédito no Brasil, aumentando a necessidade de as empresas buscarem melhorar a assertividade da concessão de crédito para novos clientes. Diante disso, diferentes técnicas quantitativas vêm sendo utilizadas para a construção de modelos de previsão de risco de crédito. Essas técnicas são baseadas em informações cadastrais dos clientes, cuja combinação de diversas variáveis, predizem um comportamento padrão de risco. O objetivo deste artigo é comparar modelos individuais e híbridos de regressão logística e rede neural na assertividade da classificação de bons e maus clientes de uma empresa. Para avaliação dos modelos estatísticos propostos, os valores de percentual de acerto geral, do teste KS e da área abaixo da curva ROC foram avaliados. O modelo foi aplicado em uma amostra de clientes de uma rede de farmácias. Os resultados do estudo apontaram uma superioridade para o método de modelagem híbrida comparado a técnica individual mais usual na análise de crédito, ressaltando que a utilização dos resultados da regressão logística, como nós de entrada da rede neural, melhora as características da técnica de análise de rede neural, diminuindo o tempo de processamento e aprendizado das redes e facilitando a interpretação dos resultados obtidos.

---

## CP45 – Evolution of electricity consumption residential in geographical regions of Brazil: Comparison of linear regression coefficients

Matheus Ferreira do Amaral<sup>1</sup>, Raphael Silva<sup>1</sup>, Gabriela Maria Rocha Bolaina<sup>1</sup> and Ednaldo C. Guimarães<sup>1</sup>

<sup>1</sup> Universidade Federal de Uberlândia, MG, Brasil

**Email:** gabibolaina@gmail.com

**Abstract:** The aim of this study was to verify, by comparing the linear regression coefficient, the power consumption rate differed significantly between geographical regions of Brazil. The results indicated that the growth rate of electricity consumption did not differ statistically between the regions although it has different levels of consumption among regions.

## CP46 – Adjustment of nonlinear models applied to data from the artificial population dynamics of blowflies

Paulo Henrique Sales Guimarães<sup>1</sup>

<sup>1</sup> Universidade Estadual Paulista, SP, Brasil

**Email:** mmportsatori@ibb.unesp.br

**Abstract:** Nonlinear time series occur frequently in numerous natural, economic and social phenomena. Thus, to model these types of series nonlinear models must be considered. These, in recent decades, showed a significant development and new methods have been proposed to evaluate the nonlinearity in time series. This study aims to evaluate the fit of different nonlinear and linear models for a number of population of blowflies of an experiment. Applied a test to verify the nonlinearity of the series, and proposed two linear models and two nonlinear to assess the population dynamics of blowflies. The results indicated the superiority of the proposed models and the need to consider other models to consider, for example, Markov regime change that is able to estimate parameters of each series regime.

---

## CP47 – Data analysis with repeated measures for an experiment sugar cane

Ricardo Olinda<sup>1</sup>, Pedro Chinelato<sup>2</sup>, João Gil Luna<sup>1</sup> and Ednário Mendonça<sup>3</sup>

<sup>1</sup> Universidade Estadual da Paraíba, PB, Brasil

<sup>2</sup> Universidade de São Paulo, SP, Brasil

<sup>3</sup> Universidade Federal de Campina Grande, PB, Brasil

**Email:** pedrochinelato@gmail.com

**Abstract:** Designs with repeated measures usually involve two factors: treatments and times, and are frequently used in sugarcane tests considering different measures on the same plot over time, in general, is based on the classical model of the design in installments subdivided in time. However, this analysis does not consider the effects of correlation (dependence) between observations taken on the same field at different times. Thus, the analysis may not be valid because the errors corresponding to the experimental units can not meet the restrictions of independence and homoscedasticity. In this context, this study was to evaluate the effects of stricter application of repeated measures analysis (longitudinal data), using different structures in the matrix of variance and covariance.

---

## CP48 – Utilização da linguagem Stan na análise de dados de dose-resposta

Roseli Aparecida Leandro<sup>1</sup>, Erasnilson Vieira Camilo<sup>1</sup> e José Bruno Malaquias<sup>1</sup>

<sup>1</sup> ESALQ/Universidade de São Paulo, SP, Brasil

**Email:** roseli.ap.leandro@gmail.com

**Resumo:** Neste trabalho pretende-se motivar a utilização da linguagem Stan na inferência bayesiana. Apresenta-se os resultados obtidos na análise de dados dose-resposta em dados de compostos de bio-ativos de erva-doce *Foeniculum vulgare* Mill na largata do algodoeiro Alabama argillacea Huebner utilizando a interface do Stan (rstan) com o software R.

---

## CP49 – Additive regression model: An example without and with dynamic covariates

Rosemeire Fiaccone<sup>1</sup>

<sup>1</sup> Universidade Federal da Bahia, BA, Brasil

**Email:** rose.fiaccone@gmail.com

**Abstract:** One way to analyse recurrent event data is to use longitudinal techniques. A very general and mathematically convenient way to model these events is through their intensity functions. In addition to this, this approach is important when individuals experience dynamic or changing conditions that affect event occurrence and a previous event history affects the future events. Using data from a randomised double blind placebo controlled trial in children aged 6 to 48 months at baseline, we present an alternative approach to address several dynamic features of the longitudinal design using an additive discrete time regression model. The idea is to see how other covariates influence instantaneous change in the diarrhoea process through incidence measures and also to include dynamic covariates for reflect the previous history of events, such as, time since last event or any function of the past since effects may change over time. Their inclusion take into account the possibility of a child experiences different vulnerability for new event according to his history of diarrhoea events. Many choices can be made in relation to the construction of dynamic covariates. The choice may depend on considerable insight into the nature of the data. We defined dynamic covariates as the historical subject-specific rate of episodes, days with diarrhoea, days with fever and days with vomit. However, the inclusion of a dynamic covariate in an analysis may distort the estimation of the effects of the fixed covariates.

---

## CP50 – Análise de trilha na influência de componentes morfométricos do bagre americano (*Ictalurus punctatus*) no seu rendimento de filetagem

Tadeu Vilela de Souza<sup>1</sup>, Adriano Carvalho Costa<sup>2</sup>, João Domingos Scalon<sup>3</sup>, Rilke Tadeu Fonseca de Freitas<sup>3</sup> e Thiago Archangelo Freato<sup>3</sup>

<sup>1</sup> Instituto Federal do Sul de Minas Gerais, MG, Brasil

<sup>2</sup> Instituto Federal Goiano, GO, Brasil

<sup>3</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** tadeu\_vilela@hotmail.com

**Resumo:** O objetivo deste trabalho foi aplicar a análise de trilha para avaliar e determinar quais são os caracteres morfométricos que mais contribuem para o rendimento corporal (rendimento de filé) do bagre americano (*Ictalurus punctatus*). Setor de Piscicultura da Universidade Federal de Lavras-MG. Para obtenção das medidas morfométricas e do rendimento de filé, os peixes amostrados foram insensibilizados

através da secção da medula, sacrificados por sangria, pesados, submetidos à avaliação morfológica e dissecados para a separação e pesagem dos componentes corporais. Além do rendimento corporal, foram realizadas as seguintes medidas morfológicas: comprimento padrão (CP), compreendido entre a extremidade anterior da cabeça e o menor perímetro do pedúnculo (inserção da nadadeira caudal); comprimento de cabeça (CC), compreendido entre a extremidade anterior da cabeça e a borda caudal do opérculo e largura do corpo (LC) tomada na região do primeiro raio das nadadeiras dorsal. As correlações entre as variáveis independentes e a variável dependente foram desdobradas em efeitos diretos e indiretos, obedecendo todos os pressupostos estatísticos necessários da análise de trilha. Os resultados obtidos mostraram que o comprimento padrão (CP) foi o caractere morfológico que apresentou o maior efeito direto sobre o rendimento de filé e, portanto, o mais indicado para a seleção indireta com o objetivo de obter peixes com maior rendimento de filé.

---

## CP51 – Regressão ridge: Uma abordagem para solucionar o problema da multicolinearidade

Tatiana Kazue Yamanouchi<sup>1</sup> e Pollyane Vieira da Silva<sup>1</sup>

<sup>1</sup> Universidade de São Paulo, SP, Brasil

**Email:** vieira\_polly@hotmail.com

**Resumo:** A multicolinearidade é um problema comumente encontrado quando se trabalha com regressão linear múltipla, em que um par ou mais das variáveis regressoras apresentam altas correlações. Esse problema se deve a um mal condicionamento da matriz de delineamento, inflacionando portanto, a variância dos parâmetros dos estimadores obtidos pelo método de mínimos quadrados. Assim, a presença de multicolinearidade nos dados pode alterar significativamente as estimativas do parâmetro do modelo. Neste trabalho será apresentado um método para o diagnóstico para verificar a presença de multicolinearidade nos dados (Fator de Inflação da Variância) bem como uma alternativa para corrigir este problema, a regressão ridge. A finalidade da regressão ridge é em encontrar um valor “ótimo” para  $k$ . Aqui abordaremos duas propostas para o  $k$  ótimo e discutiremos as vantagens de se utilizar a regressão ridge sobre os estimadores de mínimos quadrados, utilizando uma aplicação em dados reais.

---

## CP52 – Ajuste do modelo de Kuhn na descrição da isoterma de adsorção aplicada a dados de guaraná em pó

Thais Ribeiro<sup>1</sup>, Laís Silva<sup>1</sup>, Joel Muniz<sup>1</sup>, Augusto R. Morais<sup>1</sup> e Gelson Ribeiro Júnior<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** armorais@dex.ufla.br

**Resumo:** A água em geral representa um dos constituintes em maior proporção nos alimentos, e dependendo da quantidade de água presente, finalidade do produto e tempo de vida de prateleira pode ocorrer uma série de mudanças organolépticas alterando a qualidade do alimento. A relação entre a umidade de equilíbrio do alimento e a sua atividade de água em uma determinada temperatura é chamada de isoterma. Estudos com isotermas de sorção são importantes para poder avaliar o tempo de armazenamento dos ali-

mentos em condições específicas. Este trabalho teve como objetivo descrever as isotermas de guaraná em pó nas temperaturas de 15 e 35°C pelo modelo não linear de Kuhn. A estimação dos parâmetros foi feita pelo método dos quadrados mínimos, com base no algoritmo de Gauss-Newton, utilizando-se o software SAS. Foram usados como avaliadores da qualidade do ajuste o Coeficiente de Determinação Ajustado ( $R^2_{aj}$ ), Desvio Padrão Residual (DPR), e o teste de Durbin-Watson (DW) para testar a independência dos resíduos. Encontrou-se autocorrelação no resíduos, modelando por processo autoregressivo de ordem 1 ao longo da atividade de água na temperatura de 15°C. Os valores de  $R^2_{aj}$  foram em geral altos a cima de 0,96 e DPR baixo indicando boa qualidade nos ajustes aos dados em ambas temperaturas.

## CP53 – Uso dos modelos exponencial simples com três parâmetros e Henderson para descrever a secagem de polpa de jabuticaba

Thais Ribeiro<sup>1</sup>, Michele Barbosa<sup>1</sup>, Joel Muniz<sup>1</sup> e Marcio Balestre<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil  
Email: t.destefani.ribeiro@gmail.com

**Resumo:** A jabuticabeira (*Myrciaria spp.* Berg) é uma árvore frutífera nativa do sul do Brasil, sendo a cultivar ‘Sabará’ a mais produzida no país, encontrada desde o Pará até o Rio Grande do Sul. O fruto é rico em vitaminas, no entanto a jabuticaba é muito perecível e necessita de técnicas adequadas no pós colheita, a secagem da polpa de jabuticaba surge como uma saída para se reduzir estas perdas. Este trabalho teve como objetivo descrever a cinética de secagem de polpa de jabuticaba na temperatura de 60°C pelos modelos: Exponencial simples com três parâmetros e Henderson, e compará-los utilizando o software SAS (versão 9.0).os modelos foram comparados utilizando o Coeficiente de Determinação Ajustado ( $R^2_{aj}$ ) e o Desvio Padrão Residual (DPR). Avaliou-se também a possível existência de correlação entre os resíduos bem como as medidas de curvatura de Bates e Wattes. Os resultados do teste de DW indicaram a presença do erro autoregressivo de ordem 1 ao longo do tempo para ambos os modelos. Os valores de  $R^2_{aj}$  foram em geral altos a cima de 0,99 e DPR baixos indicando boa qualidade nos ajustes dos modelos aos dados. O modelo de Henderson o mais indicado para descrever os dados de acordo com as medidas de curvatura de Bates e Wattes.

## CP54 – Qualidade do ajuste do modelo Poisson-Weibull de longa duração na presença de covariáveis

Valdemiro Vigas<sup>1</sup>, Giovana Oliveira<sup>2</sup> e Francisco Louzada<sup>3</sup>

<sup>1</sup> Universidade Federal do Mato Grosso do Sul, MS, Brasil  
<sup>2</sup> Universidade Federal da Bahia, BA, Brasil  
<sup>3</sup> ICMC/Universidade de São Paulo  
Email: valdemiro.vigas@ufms.br

**Resumo:** Neste trabalho é proposto o modelo de mistura de longa duração usando a distribuição Poisson-Weibull. Este modelo provém da distribuição Poisson-Weibull (PW), que generaliza as distribuições exponencial-Poisson (EP) e Weibull, e é deduzida na situação em que não é possível observar o valor exato da variável tempo de sobrevivência, mas apenas o valor mínimo dos tempos é possível observar.

Os parâmetros do modelo são estimados pelo método de máxima verossimilhança. Além disso, medidas de influência e análise de resíduos são utilizados para verificar a adequabilidade deste modelo. Por fim, a utilidade do modelo proposto foi ilustrada por meio de sua aplicação à conjunto de dados.

---

## Sessão de Pôsters 1 / Poster Session 1

---

### Bioestatística e Modelos Biométricos 1

[*Biostatistics and Biometrical Models 1*]

---

#### CP55 – Parameterization of Davis dynamic mathematical model to individually estimate the performance of cross beef cattle

Adrielle Giaretta Biase<sup>1</sup>, Carlos Tadeu dos Santos Dias<sup>1</sup>, Luís Gustavo Barioni<sup>2</sup>, Tiago Zanett Albertini<sup>3</sup>, Lucieta Guerreiro Martorano<sup>4</sup>, James W. Oltjen<sup>5</sup>, Dante Pazzanese Lanna<sup>1</sup>, Patrícia Perondi Anchão Oliveira<sup>6</sup>, Sérgio Raposo de Medeiros<sup>7</sup> and Roberto Augusto de Almeida Torres Júnior<sup>7</sup>

<sup>1</sup> ESALQ/Universidade de São Paulo, SP, Brasil

<sup>2</sup> Embrapa Informática Agropecuária, Brasil

<sup>3</sup> ArrobaTech, SC, Brasil

<sup>4</sup> Embrapa Amazonia Oriental, Brasil

<sup>5</sup> University California, EUA

<sup>6</sup> Embrapa Pecuária Sudeste, Brasil

<sup>7</sup> Embrapa Gado de Corte, Brasil

**Email:** ctsdias@usp.br

**Abstract:** Dynamic models to generate growth curves for individual growth curves for bovines are commercially unavailable in Brazil. This scenario is about to as individual performance monitoring systems became available. The prediction of chemical composition by Davis Growth Model (DGM) can help to determine the optimal point of slaughter of cattle allowing the producers to meet stringent specifications of the international market and to increase profitability. The aim of this study was to present an approach for prediction of individually determined parameter values of the DGM to estimate the dynamics of growth and body composition for each animal. Data for the parameter estimation was produced in a trial run at the Embrapa Beef Cattle, located in Campo Grande, Mato Grosso do Sul (MS) state, Brazil, in 2007. The data was comprised of 27 crossbred cattle 1/2 Red Angus  $\times$  1/4 Caracu  $\times$  1/4 Nellore [9 female and 9 young bulls] and Red Angus  $\times$  Nellore (4 female and 5 bulls). Before weaning, the animals were raised in two systems: pasture system (PS) and confined system (CS). Multivariate covariance analysis were performed based on the individual parameter estimates by DGM. Additionally, Mahalanobis distances between the scenarios were calculated to verify the dissimilarity between groups. Covariance multivariate analyzes of the individual parameter vectors and the Mahalanobis distances revealed pronounced effects of production systems, genres and significant interaction between those factors. The DGM expressed potential application to estimate the growth performance in production systems in animal in Brazil.

## CP56 – Dimensões de unidades experimentais básicas e de parcelas em tremoço branco

Alberto Cargnelutti Filho<sup>1</sup>, Bruna Mendonça Alves<sup>1</sup>, Marcos Toebe<sup>2</sup> e Giovani Facco<sup>2</sup>

<sup>1</sup> Universidade Federal de Santa Maria, RS, Brasil

<sup>2</sup> Fundação Universidade Federal do Pampa, RS, Brasil

**Email:** alberto.cargnelutti.filho@gmail.com

**Resumo:** O objetivo deste trabalho foi verificar a influência do tamanho da unidade experimental básica (UEB) na estimativa do tamanho ótimo de parcela, para a avaliação da massa de matéria verde de tremoço branco (*Lupinus albus* L.). Avaliou-se a massa de matéria verde em 432UEB de  $1m \times 1m$  ( $1m^2$ ) e formaram-se 16 planos de UEB com tamanhos entre 1 e  $16m^2$ . Para cada plano de UEB foi determinado o tamanho ótimo de parcela, pelo método da curvatura máxima do modelo do coeficiente de variação. Foi calculada a diferença mínima significativa pelo teste de Tukey, em 448 cenários formados pelas combinações de 16 tamanhos ótimos de parcela, i tratamentos ( $i = 5, 10, 15e20$ ) e r repetições ( $r = 3, 4, 5, 6, 10, 15e20$ ), para o delineamento blocos ao acaso. A estimativa do tamanho ótimo de parcela depende do tamanho da unidade experimental básica. É indicado avaliar a massa de matéria verde em UEB de menor tamanho possível, para serem usadas na estimação do tamanho ótimo de parcela.

---

## CP57 – Practical application of the sign test on data from trial and meta-analysis

Alfredo José Barreto Luiz<sup>1</sup>, Adriane da Rocha Costa<sup>2</sup>, Eunice Reis Batista<sup>1</sup> and Rosa Toyoko Shiraishi Frighetto<sup>1</sup>

<sup>1</sup> Embrapa Meio Ambiente, SP, Brasil

<sup>2</sup> Universidade Federal Rural da Amazônia, PA, Brasil

**Email:** alfredo.luiz@embrapa.br

**Abstract:** Many practical situations faced in agricultural research does not allow assuming the normality of the data to be analyzed. In some of these situations, the researcher, not statistical, want to test hypotheses with minimal statistical security, but finds himself in trouble by not dominate approaches that require deeper knowledge in that discipline. The sign test has a very strong appeal for interpretability of results and is often sufficient for decision-making on the further direction of research or analysis. In this paper, we present two examples of application and interpretation of the test in the form available at Univariate routine SAS, one for data obtained for phosphorus content of the meta-analysis of available soil in areas under different forms of management, from dozens of published articles. The second example is data from testing in an area with strong physical condition of location of treatment and repetitions because the plants are perennials (coffee) and the treatments comprise the application of CO<sub>2</sub> through a fixed system, expensive and rare (FACE - Free Air CO<sub>2</sub> Enrichment). The results allowed us to evaluate the relevance of the hypotheses tested in a satisfactory manner considered by the researchers responsible for data.

## CP58 – Martingales: Teoria, simulações e aplicações

Allan Robert Silva<sup>1</sup> e Cláudio T. Cristino<sup>2</sup>

<sup>1</sup> Universidade Federal de Sergipe, SE, Brasil

<sup>2</sup> Universidade Federal Rural de Pernambuco, PE, Brasil

**Email:** ctcrisino@gmail.com

**Resumo:** Originalmente, martingale que se refere a uma classe de apostas estratégias popular na França no século 18. O conceito de martingale na Teoria das Probabilidades foi introduzido por Paul Lévy em 1934, embora ele não o tenha nomeado desta forma. O termo martingale foi introduzido posteriormente por Ville (1939), que por uma motivação inicial queria mostrar a impossibilidade de estratégias de apostas bem-sucedidas. Atualmente, existem diversas aplicações em diferentes áreas como finanças, ciências atuariais, análise de sobrevivência, etc. Neste trabalho descrevemos sobre a Teoria de Martingale, com uso de teoremas e de alguns exemplos. Exploraremos algumas aplicações a partir de simulações computacionais, onde a primeira tem um objetivo didático e a segunda tem o objetivo de comparar uma estratégia de investimento de compra de dólares com base na ideia de martingale com investimento de poupança no mesmo período.

---

## CP59 – Avaliação da associação entre o desenvolvimento humano da população de Sergipe em relação à pecuária leiteira

Amanda da Silva Lira<sup>1,2</sup>, Gabrielle Monteiro Souza<sup>1</sup>, Vanessa Kelly dos Santos<sup>1,2</sup>,  
Luydson Jamyson do Nascimento Souza<sup>1</sup> e Kleber Régis Santoro<sup>2</sup>

<sup>1</sup> Universidade Federal de Sergipe, SE, Brasil

<sup>2</sup> PPGBEA, Universidade Federal de Pernambuco, PB, Brasil

**Email:** amandalira@ufs.br

**Resumo:** A bovinocultura e a pecuária leiteira são atividades que vem ganhando cada vez mais espaço na economia sergipana, ao mesmo tempo percebe-se que o IDH nos municípios também vem apresentando crescimento em seus índices. Diante disso, o presente trabalho objetiva avaliar possíveis influências que o crescimento na pecuária leiteira possa estar refletindo sobre os índices do IDHM das últimas três décadas. Com os dados disponibilizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e Programa das Nações Unidas para o Desenvolvimento (PNUD), foi possível organizar o banco de dados com as variáveis de efetivo bovino, produção leiteira, valor da produção, populações urbana e rural e o Índice de Desenvolvimento Humano Municipal (Longevidade, Renda e Educação) nos anos 1991, 2000 e 2010. Com os dados executaram-se técnicas como correlação e testes que avaliam a adequabilidade da aplicação da Análise Fatorial, uma técnica multivariada que permite identificar fatores para explicar a relação entre as variáveis em estudo. Os resultados mostraram que não há relação entre o crescimento do IDHM e a pecuária leiteira no estado de Sergipe, de forma que o crescimento de um não depende do crescimento do outro.

## CP60 – Use of factor analysis to evaluate everyday themes students of the 6th grade of elementary school

Amanda Liz Pacífico Manfrim Peticarrari<sup>1</sup>, Clara Augusta Marques Natalin<sup>1</sup>,  
Andreia da Silva Meyer<sup>1</sup> and Aline Maria Pacífico Manfrim<sup>2</sup>

<sup>1</sup> Universidade Estadual Paulista, SP, Brasil

<sup>2</sup> Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, SP, Brasil

**Email:** amanda@fcav.unesp.br

**Abstract:** This paper presents a survey conducted under the educational Statistics and Mathematics to identify everyday themes of students Elementary school and, secondly, through the information obtained from these students, describe and investigate the similarity between these topics (variables) using a multivariate analysis of factors to subsequently develop a didactic material that can be used to support the learning statistical in mathematics discipline for Elementary Education. Were interviewed students the 6th year of elementary school and explored some topics to identify those that are part of daily life and/or interest of these students. As a result, it was possible to divide the study into four dimensions expressed in the factor loadings of four factors, which allowed to act independently in these four dimensions.

---

## CP61 – Longitudinal analysis of inter-rater agreement and associated factors in medical records audit

Ana Carolina Cintra Nunes Mafra<sup>1</sup>, Renata Rafaella Santos Tadeucci Medina<sup>1</sup>,  
Ederson Haroldo Pereira Almeida<sup>1</sup> and Mario Maia Bracco<sup>1</sup>

<sup>1</sup> Hospital Israelita Albert Einstein, SP, Brasil

**Email:** mario.bracco@einstein.br

**Abstract:** Background: Patient's medical records are widely used in observational studies. However, the reliability of the information extracted from them is a matter of concern in audit processes, in order to ensure inter-rater agreement. Objective: To develop a statistical model to measure the agreement among observers over time and identify possible associated factors with the audit process at the Hospital Municipal Dr. Moysés Deutsch. Data: In 2015, monthly medical records audits were conducted and concordance analysis occurred every three months. 307 records were audited by 21 auditors who have evaluated each record by, at least, two auditors. Each auditor evaluated 15 items that could be rated as "not applicable", "appropriate" or "inappropriate". Statistical Methods: It was obtained the Gwet's agreement coefficient among auditors for each item and audits. For temporal trends analyses, it was considered the percent agreement, which was estimated by beta models. Possible variations were verified among items and the presence of the auditors at monthly meetings. Results and Discussion: Although the percent agreement overestimates the agreement coefficient, it was observed that both have a similar temporal variation, allowing inferences about its results. Over the time there was a general increase in the percent agreement of 1.33% (p-value = 0.037). If any auditor have not attended the meeting, the agreement decreased by 1.05%, but was not significant (p-value = 0.526). Conclusion: Tracking the progress of agreement allows setting up goals and identifying associated factors to improve medical record audit processes.

## CP62 – Latent class analysis with covariates for evaluation of asthma symptoms

Beatriz Santos<sup>1</sup>, Leila D.A.F. Amorim<sup>1</sup> and Ana Clara Campos<sup>2</sup>

<sup>1</sup> IM e ISC/Universidade Federal da Bahia, BA, Brasil

**Email:** beatrizschindler@hotmail.com

**Abstract:** Latent class analysis (LCA) is a statistical methodology commonly used to identify subgroups (latent classes) based on observed response patterns from a set of categorical variables. LCA aims to display large arrays of data in a parsimonious format while revealing important features of the data. LCA with covariates is a methodological extension of LCA with increasing use for answering scientific questions about the effects of factors on latent responses. In this paper we illustrate these methodologies analyzing data about asthma symptoms in children participating in SCAALA-Salvador Study (Barreto et al, 2006). The results indicate that the chance of belonging to the latent class called “Symptomatic” in relation to latent class “Healthy” is approximately 11.2% lower for boys than for girls (95%CI = 0.66; 1.20); 88% greater for atopic children; and decreases with age, being 47% e 56% lower for 6 – 7 and = 8 years-old compared to < 6 years. The disclosure of this type of methodology among Brazilian researchers can answer questions that have not been properly analyzed by other methods more commonly known by data analysts.

---

## CP63 – Relações lineares entre caracteres em genótipos transgênicos de milho

Bruna Mendonça Alves<sup>1</sup> e Alberto Cargnelutti Filho<sup>1</sup>

<sup>1</sup> Universidade Federal de Santa Maria, RS, Brasil

**Email:** brunamalves11@gmail.com

**Resumo:** O objetivo deste trabalho foi avaliar as relações lineares entre caracteres fenológicos, morfológicos, produtivos e nutricionais em genótipos transgênicos de milho. O experimento foi realizado no ano agrícola 2009/2010, no delineamento blocos ao acaso com três repetições. Foram avaliados 18 genótipos transgênicos de milho e mensurados os caracteres fenológicos, morfológicos, produtivos e nutricionais. Inicialmente, foi realizada a verificação dos pressupostos de normalidade dos erros e de homogeneidade de variâncias residuais, pelos testes de Kolmogorov Smirnov e Bartlett, respectivamente. Após, para cada um dos 13 caracteres mensurados foi realizada a análise de variância e o teste F. As médias foram comparadas por meio do teste de Scott Knott a 5% de probabilidade. Foi determinada a matriz de coeficientes de correlação fenotípica entre os caracteres e realizado o diagnóstico de multicolinearidade, por meio do número de condição. Foram realizadas duas análises de trilha, considerando os caracteres nutricionais (proteína bruta e extrato etéreo) como principais e os demais como explicativos e, realizada a análise de regressão stepwise. O número de dias da semeadura até o florescimento masculino tem relação linear positiva com a proteína bruta e com o extrato etéreo. O número de dias da semeadura até o florescimento feminino tem relação linear negativa com a proteína bruta em grãos de milho. Os caracteres número de dias da semeadura até o florescimento masculino e número de dias da semeadura até o florescimento feminino podem ser utilizados para seleção indireta em programas de melhoramento genético de milho.

## CP64 – Inclusão de variável multicategórica na análise de trilha da produtividade de alho

Brunna Rithielly Rezende<sup>1</sup>, Divino de Sousa Costa<sup>1</sup>, Hygor Amaral Santana<sup>1</sup>,  
Wilhan Valasco dos Santos<sup>1</sup> e Anderson Rodrigo da Silva<sup>1</sup>

<sup>1</sup> Instituto Federal Goiano - Câmpus Ururtaí, GO, Brasil

**Email:** brunnarithielly@hotmail.com

**Resumo:** Apesar de ser uma cultura de importância indiscutível, poucos trabalhos envolvendo análise de trilha da produtividade de alho foram realizados. Não foi encontrado nenhum trabalho do tipo utilizando informações de variáveis multicategóricas. O objetivo deste trabalho foi realizar análise de trilha da produtividade de alho em função de variáveis morfológicas relacionadas ao bulbo, sendo uma destas uma variável multicategórica. Foram utilizados dados de um experimento envolvendo 89 acessos de alho, sendo analisadas as seguintes variáveis: produtividade (variável básica), diâmetro e comprimento do bulbo, número de bulbilhos por bulbo e a variável multicategórica cor da túnica do bulbo. Entre as variáveis quantitativas o coeficiente de correlação de Pearson foi utilizado para medir o grau de associação linear. Correlações entre as variáveis quantitativas e a cor foram calculadas via correlação poliserial. Diâmetro e comprimento do bulbo foram fortemente ( $r > 0.8$ ) correlacionados com a produtividade de alho. Entretanto, apenas o diâmetro apresentou efeito direto relevante sobre a produtividade.

## CP65 – Tamanho de amostra para estimar o coeficiente de correlação de Pearson em experimentos com tomate cereja

Bruno Giacomini Sari<sup>1</sup>, Alessandro Dal'col Lúcio<sup>1</sup>, Cintya Souza Santana<sup>1</sup>, André Luis Tischer<sup>1</sup>, Andrielle Magrini Rodrigues<sup>1</sup>, Letícia Gall Dreifke<sup>1</sup> e Lucas Drebes<sup>1</sup>

<sup>1</sup> Universidade Federal de Santa Maria, RS, Brasil

**Email:** brunosari@hotmail.com

**Resumo:** O presente trabalho teve como objetivo determinar o tamanho de amostra necessário para estimar o coeficiente de correlação de Pearson entre caracteres produtivos do tomate cereja. Para isso, um ensaio de uniformidade foi instalado em ambiente protegido na estação primavera/verão de 2014. De cada planta foi mensurada as variáveis comprimento médio de fruto, largura média de fruto, peso médio de fruto, número de cachos, número de frutos por cacho, número de frutos e peso total de frutos. A partir delas, foi calculada a matriz de correlação de Pearson. Posteriormente, foram planejados 48 tamanhos de amostra, sendo o tamanho inicial composto por 10 plantas e os demais obtidos acrescentando-se 5 plantas. Para cada tamanho de amostra planejado foram obtidos 3000 estimativas do coeficiente de correlação de Pearson através de reamostragens “bootstrap” com reposição. A partir das 3000 estimativas foi calculado o percentil 2, 5%, a média e o percentil 97, 5% dos coeficientes de correlação de Pearson. O tamanho de amostra de cada coeficiente de correlação de Pearson foi determinado quando o valor da amplitude do intervalo de confiança foi menor ou igual a 0,4. A obtenção das estimativas do coeficiente de correlação de Pearson com elevada precisão é particularmente difícil para caracteres com relação linear fraca. Como consequência, observou-se que o tamanho de amostra e a magnitude do coeficiente de correlação de Pearson mostraram-se inversamente proporcionais. Para estimar o coeficiente de correlação de Pearson,

com intervalo de confiança de 0,4, são necessárias 200 plantas de tomate.

---

## CP66 – Relações entre caracteres produtivos de tomate tipo cereja

Cintya Souza Santana<sup>1</sup>, Alessandro Dal'col Lúcio<sup>1</sup>, Bruno Giacomini Sari<sup>1</sup>, Lucas Drebes<sup>1</sup>, Andrielle Magrini Rodrigues<sup>1</sup>, André Luis Tischler<sup>1</sup> e Letícia Gall Dreifke<sup>1</sup>

<sup>1</sup> Universidade Federal de Santa Maria, RS, Brasil

**Email:** css.agro@live.com

**Resumo:** O presente trabalho tem como objetivo avaliar a relação linear entre os componentes da produtividade do tomate cereja. Um ensaio de uniformidade foi instalado com o híbrido Lilli, em ambiente protegido na estação primavera/verão de 2014. As plantas foram conduzidas com duas hastes e tutoradas com fitilho. Para cada planta foi mensuradas as variáveis comprimento médio de fruto (cm), largura média de fruto (cm), peso médio de fruto (gramas), número de cachos, número de frutos por cacho, número de frutos e peso total de frutos (gramas). Os dados obtidos foram analisados por meio da matriz de correlação de Pearson e análise de trilha. A análise de trilha foi realizada considerando-se o peso total de frutos como variável principal e as demais como variáveis explicativas. Devido à presença de multicolinearidade, optou-se pela eliminação da variável número de frutos por cacho. O coeficiente de correlação de Pearson entre as variáveis explicativas e a variável principal foi significativo, indicando relação linear entre elas. Verificou-se também a inexistência de relação de causa e efeito entre a maioria das variáveis explicativas com a principal. O número de frutos por planta foi a que apresentou maior relação de causa e efeito com a produtividade. Portanto, a produtividade do tomate cereja está diretamente relacionada com o número de frutos por planta.

---

## CP67 – Simulação morfológica na caracterização espacial da perda de solo por erosão em cultivo de cana-de-açúcar

Daniel de Bortoli Teixeira<sup>1</sup>, João Fernandes da Silva Júnior<sup>2</sup>, Marcílio Vieira Martins Filho<sup>1</sup> e Gener Tadeu Pereira<sup>1</sup>

<sup>1</sup> FCAV/Univesidade Estadual de São Paulo, SP, Brasil

<sup>2</sup> Instituto Federal do Pará, PA, Brasil

**Email:** daniel.dbt@hotmail.com

**Resumo:** O conhecimento do padrão espacial da perda de solo por erosão é vital para o planejamento conservacionista. Neste estudo objetivou-se avaliar a aplicação dos procedimentos de modelagem estocástica na caracterização espacial das unidades de mapeamento de perdas de solo por erosão em dois cenários de manejo de cana-de-açúcar, cana crua e cana queimada. O experimento foi conduzido em uma área de 200 ha. Uma malha de amostragem regular foi estabelecida com intervalos equidistantes de 50 m, totalizando 626 pontos. Em cada ponto foi estimada a perda de solo por erosão. Estes resultados foram analisados por meio da simulação sequencial indicatriz com variáveis categóricas (simulação morfológica). Posteriormente, foi realizada a nova classificação morfológica e avaliada a incerteza morfológica por meio da entropia de Shannon. As maiores perdas de solo em ambos os cenários foram localizadas nas regiões com maior declividade, as quais apresentaram maior incerteza na sua classificação. Os mapas morfológicos

evidenciam os locais em que as práticas de manejo conservacionistas devem ser intensificadas, ou seja, locais em que o limite de tolerância à perda de solo é ultrapassado. A técnica de simulação morfológica fornece uma noção de precisão das informações dos mapas por meio da entropia de Shannon. A simulação morfológica deve ser recomendada para a caracterização da distribuição espacial das perdas de solo por erosão em ambos os cenários, sendo esta um instrumento imprescindível para o planejamento conservacionista de precisão.

---

## CP68 – Identificação do período de prevenção contra o vírus sincicial respiratório e ocorrência de bronquiolite viral aguda

Edilenia Queiroz Pereira<sup>1</sup>, Márcia Lorena Alves dos Santos<sup>1</sup> e Eniuce Menezes de Souza<sup>1</sup>

<sup>1</sup> Universidade Estadual de Maringá, PR, Brasil  
**Email:** edileniaqueiroz@yahoo.com.br

**Resumo:** A bronquiolite viral aguda (BVA) é ocasionada pela infecção dos bronquíolos, ramificações menores dos brônquios. O principal vírus causador do diagnóstico de bronquiolite em internações de crianças com idade inferior à 12 meses é o vírus sincicial respiratório (VSR), sendo este a principal causa de mortalidade em crianças dessa faixa etária no Brasil. Além das medidas preventivas para evitar o contágio, dado a inexistência de vacinas para este vírus, existem medicamentos para a prevenção da BVA, entre eles o anticorpo sintético Palivizumabe. Geralmente, esta medicação é administrada em cinco doses mensais, mas devido ao alto custo, sua disponibilização é restrita à casos que apresentam grande risco. Utilizando dados do DATASUS, foram calculadas as taxas de internações por bronquiolite a cada 10000 crianças com idade abaixo dos 12 meses segundo as regionais de saúde do estado do Rio Grande do Sul, com o objetivo de identificar o mês em que se deve iniciar a prevenção com o Palivizumabe. Foram utilizados modelos SARIMA para a modelagem das séries temporais das taxas de internações por bronquiolite entre 2000 e 2015, em que cada modelo foi identificado, estimado e validado, possibilitando a identificação do comportamento sazonal da BVA para cada regional de saúde de cada estado. Para as séries em que os pressupostos do modelo SARIMA não foram atendidos, optou-se pelo ajuste utilizando modelos *wavelets*. De modo geral, o efeito sazonal da BVA coincide com os meses que apresentam temperaturas amenas.

---

## CP69 – Application of marginal structural models to irregular longitudinal data

Edmore Chamapiwa<sup>1</sup>, David Reeves<sup>1</sup>, Evangelos Kontopantelis<sup>1</sup> and Darren Ashcroft<sup>1</sup>

<sup>1</sup> University of Manchester, UK  
**Email:** edmore.chamapiwa@postgrad.manchester.ac.uk

**Abstract:** Background: Marginal Structural Models (MSMs), a class of structural causal models, are being increasingly used in the analysis of complex longitudinal health data because of their ability to give unbiased effect estimates of a time-varying treatment in the presence of time-varying confound-

ing/mediating covariates. However, MSMs have shown good performance to settings where observations occur at regularly separated time points for all patients, whereas in “real-life” health record data, different patients are commonly seen and measured at different and irregular time points. In addition, the frequency with which a patient is seen may well be related to their health outcomes. The impact of irregular, but more realistic, data on the performance of MSMs is unknown. Objective: To evaluate the performance in effect estimation of inverse-probability-weighted MSMs in unbalanced longitudinal data Methods: A simulation study was conducted to compare treatment effect estimates from inverse-probability-weighted MSM, adjusted and unadjusted generalised estimating equation (GEE) models. Irregular longitudinal data was generated by sampling time between consecutive visits for an individual from an inverse Gaussian distribution. Treatment at each observation time was sampled from a Bernoulli distribution with likelihood of getting treated dependent on the confounder level, and confounder values were sampled from an inverse Gaussian distribution. Binary outcome values were simulated from a Bernoulli distribution. Data simulation and analysis were conducted in R. Results: This simulation study showed that inverse-probability-weighted MSMs outperform stratification based estimation methods when longitudinal data is unbalanced and when confounders and treatments are time-varying.

---

## CP70 – Modelling the rate of carbon dioxide absorption of the grassy species *Echinochloa crus-galli* by mixing finite distributions

Ednário Mendonça<sup>1</sup>, Ricardo Olinda<sup>1</sup>, Tiago Oliveira<sup>2</sup> and Thuenne Barros de Oliveira<sup>2</sup>

<sup>1</sup> Universidade Federal de Campina Grande, PB, Brasil

<sup>2</sup> Universidade Estadual da Paraíba, PB, Brasil

**Email:** tiagoestatistico@gmail.com

**Abstract:** In statistics, a finite mixture of distributions consists of a probabilistic model used to represent a finite number of subpopulations within the general population, without requiring such a model to identify which population belongs each observation. Here, we addressed the main concepts inherent in theory mixtures distributions and an application is made on a set of data related to the carbon dioxide absorption rate (in  $\mu\text{mol}=\text{m}^2 \text{sec.}$ ) by marram *Echinochloa crus-galli* where assessed to cold tolerance of these species. modeling were made in various models of finite mixtures in order to compare them and, through a good selection criteria, select which the best model. For the estimation of the parameters of the mixture resulting distributions used the EM algorithm. All the computational part of this work was done in the software R with the aid package “mixsmsm”.

---

## CP71 – Calculation of probabilities for daily maximum rainfall in the city of Jatai, Goiás

Elias Silva de Medeiros<sup>1</sup>, Marina Rodrigues Maestre<sup>2</sup> and Pablo Lourenço Ribeiro de Almeida<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> Universidade Estadual de Mato Grosso do Sul, MS, Brasil

**Email:** eliasestatistica@gmail.com

**Abstract:** It is known that the presence or absence of rains, in general, impacts on many sectors of society. These extreme events can cause droughts and high levels can cause significant environmental impacts. In statistics, the theory of extreme values performs a fundamental role in modeling of events associated with the very small probabilities or rare events. Thus, this study aims to estimate the probabilities at different levels of daily rainfall for the highest expected daily value in the month. The dataset used in this study is related to pluviometric precipitation of the City of Jatai (Goias), comprising a period 37 years - from January 1979 to December 2015. The first analysis was to study the behavior of the parameter shape of the GEV distribution, and computing confidence intervals for this parameter verified that Gumbel distribution is adequate to study the behavior of the largest daily value of rainfall recorded in the month in this municipality. Then, to verify the quality of the Gumbel distribution adjustment, were drawn graphs of empirical and theoretical cumulative distribution for each of the months in question, it was observed that this model fits well to the data set. From this validation was possible to determine the probability of occurrence for a minimum threshold of daily rainfall, which observed that in March there is 65% probability of rain at a minimum 50 mm in one day.

---

## CP72 – Análise quantitativa dos traçados eletrocardiográficos de pacientes do projeto Teleminas Saúde

Franciele Brito<sup>1</sup>, Aurélia A. de A. Rodrigues<sup>1</sup>, Rodolfo Ribeiro<sup>1</sup>, Elmiro Resende<sup>1</sup> e João Batista Destro Filho<sup>1</sup>

<sup>1</sup> Universidade Federal de Uberlândia, MG, Brasil

**Email:** aurelia@famat.ufu.br

**Resumo:** O eletrocardiograma (ECG) é uma ferramenta muito utilizada para rastrear e detectar doenças cardíacas, que são uma das principais causas de morte no Brasil. Alguns países, inclusive o Brasil, vêm utilizando os recursos da telemedicina para possibilitar cuidados à saúde nas situações em que a distância é um fator crítico, até mesmo no caso de doenças cardiovasculares. O Estado de Minas Gerais, por exemplo, conta desde 2006 com um sistema de telecardiologia, o projeto Teleminas Saúde, que já abrange 70% da população mineira. O projeto Teleminas Saúde consiste em seis polos, os quais são de responsabilidade das seguintes universidades: Universidade Estadual de Montes Claros, Universidade Federal de Juiz de Fora, Universidade Federal de Minas Gerais, Universidade Federal de São João Del-Rei, Universidade Federal de Uberlândia e Universidade Federal do Triângulo Mineiro. O presente trabalho tem por objetivo quantificar as informações eletrocardiográficas dos pacientes normais, sem patologias, do Projeto Teleminas Saúde. Foi realizado um estudo retrospectivo e exploratório dos eletrocardiogramas do banco de dados de 2011 do projeto Teleminas Saúde. A amostra composta por 4.000 exames foi aleatória e proporcional à população vinculada a cada polo. Os exames que foram classificados como normais no laudo médico foram agrupados e analisados quantitativamente. Observou-se que a etnia, o gênero e a idade podem influenciar o intervalo de referência de pacientes normais. Dessa forma, propõe-se, para trabalhos futuros, um estudo abrangente para a obtenção de valores de referência do eletrocardiograma normal, considerando adultos brasileiros, por gênero e idade.

## CP73 – Modelos assimétricos para dados de consumo de vitaminas em idosos

Giovana Fumes<sup>1</sup> and José Eduardo Corrente<sup>2</sup>

<sup>1</sup> ESALQ/Universidade de São Paulo, SP, Brasil

<sup>2</sup> Universidade Estadual Paulista, SP, Brasil,

**Email:** gifumesbtu@gmail.com

**Abstract:** One of the main interest in the nutrition field is to estimate the distribution of usual nutrient intake. Data from vitamin intake generally present high asymmetry mainly to the presence of outliers. This can occur because the variability of the diet and, in this case, robust estimation to get the distribution of the data can be required. Then, the aim of paper is to propose an alternative approach for estimating usual intake through asymmetric distributions with random effects applied to a data set 10 vitamins obtained from a dietetic survey for older people in Botucatu city, São Paulo, Brazil. In general, these asymmetric distributions include parameters related to mean, median, dispersion measures and such parameters provided good estimate for the intake distribution. In order to make some comparisons, a model fitted by National Cancer Institute (NCI) method with only for amount of nutrient intake were established using Akaike information criteria (AIC). NCI method is based on a Box-Cox transformation coupled with the normal distribution but in case of asymmetric data, this transformation can be not useful. It was observed that, in the presence of outliers, the asymmetric models provided a better fit than the NCI method in the major of the cases. Then, these models can be an alternative method to estimate the distribution of nutrient intake mainly because a transformation for the data are no necessary and all the information can be obtained directly from the parameters. assumptions about the logical structures of transmission and effects of interventions, having the main objective of estimating the causal effect (impact) of the intervention in a response variable. Propensity Score Matching (PSM) is one of the methods traditionally used to find an ideal counterfactual group in order to eliminate / reduce the confounding effects (selection bias). This paper summarizes the idea of the PSM approach and illustrates its use through an application with data regarding the school enrollment rate of girls in rural Burkina Faso. The results conclude that the BRIGHT program had a positive impact on school enrollment, indicating an increase of 10% to 16% in the number of girls enrolled in school depending on the adopted criteria and the causal effect estimator. PSM is a useful tool when it's assured that only observed factors are responsible for allocation to treatment. Potential bias problems can occur when assumptions are not followed or when the choice of observed variables to be included in the model is not made cautiously by the researcher.

---

## CP74 – Rede de correlação fenotípica de variáveis de alho

Hygor Amaral Santana<sup>1</sup>, Divino de Sousa Costa<sup>1</sup>, Wilhan Valasco dos Santos<sup>1</sup>,  
Brunna Rithielly Rezende<sup>1</sup> e Anderson Rodrigo da Silva<sup>1</sup>

<sup>1</sup> Instituto Federal Goiano - Campus Uruaí, GO, Brasil

**Email:** hygor\_amaral@hotmail.com

**Resumo:** A análise de redes ponderadas de correlação foi aplicada com o objetivo de descobrir estruturas de correlação e padrões de ligação entre dezesseis variáveis de alho relacionadas à folha, bulbo e outras variáveis vegetativas. Por meio do algoritmo Fruchterman-Reingold, grupos de correlação e outras

estruturas puderam ser facilmente identificados. Um padrão geral de ligação entre grupos de variáveis de folha e de bulbo foi detectado. O índice de colheita foi negativamente correlacionado com variáveis vegetativas, como esperado. Além disso, a taxa de crescimento do bulbo foi positivamente associada com o índice de área foliar, taxa de crescimento radicular e taxa de assimilação líquida da planta.

---

## CP75 – Avaliação da qualidade dos registros neonatais através de linkage e imputação de bancos de dados

Ianne Rafaella Santos Melo<sup>1</sup>, Jessica da Silva Braga<sup>1</sup>, João Batista Carvalho<sup>1</sup> e Neir Antunes Paes<sup>1</sup>

<sup>1</sup> Universidade Federal da Paraíba, PB, Brasil

**Email:** jessicasbbraga@gmail.com

**Resumo:** O objetivo consistiu em aplicar métodos probabilísticos de linkage para avaliar a qualidade da cobertura e de imputação de dados em relação aos registros de óbitos neonatais de residentes no município de João Pessoa-PB para o período de 2000 a 2013. As bases de dados foram construídas a partir dos sistemas SIM e do SINASC – Ministério da Saúde. As variáveis-chaves para a realização do linkage foram: data de nascimento, sexo do nascido vivo, endereço, idade da mãe e bairro de residência. Três situações foram investigadas para o linkage: a) dados registrados; b) dados com a exclusão das declarações com variáveis ignoradas; c) dados imputados. A população em estudo foi composta por 1.726 declarações de óbitos neonatais e 158.966 declarações de nascidos vivos em todo o período de análise. 76% foi a taxa de vinculação geral nos sistemas. Apesar de uma tendência de decréscimo do sub-registro de óbitos neonatais no período, e mesmo com a implantação do código unificado na Declaração de Nascido Vivo e na Declaração de Óbito a partir de 2010 a taxa de vinculação não ultrapassou 93%. Este achado indica a fragilidades dos sistemas que necessitam de urgentes melhorias.

---

## CP76 – Equações de volume individual para *Toona ciliata*

Joyce de Almeida Alves<sup>1</sup>, Geisi Azevedo Silva<sup>2</sup>, Ernani Lopes Possato<sup>3</sup>, Marcos Antônio Isaac Jr<sup>1</sup> e Natalino Calegario<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> Instituto Bioterra, MG, Brasil

<sup>3</sup> Universidade Federal de Uberlândia, MG, Brasil

**Email:** joyce\_a.alves@yahoo.com.br

**Resumo:** As equações de volume são utilizadas para prever o volume individual das árvores em função do diâmetro a altura do peito (DAP), ou em função do DAP e a altura total (H). Essas equações são conhecidas por equações de simples e dupla entrada, respectivamente. A pressuposição de homocedasticidade é frequentemente violada no ajuste das equações de volume, tornando necessária a aplicação de métodos para remediar esse problema. Os objetivos deste estudo foram comparar equações de volume de simples e dupla entrada para árvores de *Toona ciliata* (cedro australiano) e testar uma função de variância para contabilizar a heterocedasticidade. As equações avaliadas foram: Polinômio de grau 1, Polinômio de grau 2, Potência, Spurr e Schumacher-Hall. As três primeiras são equações de simples entrada e as demais de dupla entrada. Para a avaliação da qualidade dos ajustes utilizou-se o erro padrão residual e o gráfico de

resíduos. A qualidade do ajuste das equações de volume simples e dupla entrada foram semelhantes para a área estudo. Os valores de erro padrão residual para as equações de dupla entrada foram ligeiramente menores. Entretanto, por questões práticas, a equação de simples entrada Potência foi selecionada e então reajustada pelo método de mínimos quadrados não lineares generalizados com a inclusão da função de variância VarPower da biblioteca nlme. A função de variância representou adequadamente a heterocedasticidade para a equação selecionada.

---

## CP77 – O método simplex aplicado a dados de mistura isotópica em suco de frutas

Maria Márcia Pereira Sartori<sup>1</sup>, Luciana Balter Francisco<sup>1</sup>, Beatriz Vedovato<sup>1</sup> e Waldemar Gastoni Venturini Filho<sup>1</sup>

<sup>1</sup> Universidade Estadual Paulista, SP, Brasil  
Email: mmportsartori@ibb.unesp.br

**Resumo:** A análise isotópica tem sido usada para avaliar a legalidade em alimentos e bebidas, sendo esta metodologia reconhecida pelo MAPA. Para determinar a mistura isotópica (polpa + açúcar) utiliza-se de valores isotópicos médios, de carbono-13 do açúcar e do suco da fruta produzidos em laboratório, sendo assim a variabilidade no sinal isotópico tanto da fruta quanto do açúcar não é considerado. Com intuito de avaliar o intervalo de resultados possíveis da mistura isotópica de sucos de frutas, possibilitando assim obter maior confiabilidade na avaliação da legalidade de sucos, utilizou-se o método simplex para minimizar e maximizar a função mistura com as restrições intrínsecas do problema. Esse método possibilitou determinar os possíveis valores isotópicos para cada porcentagem de polpa de fruta adicionada ao suco.

---

## CP78 – Modelos com platô para estimativa do nível de zinco inorgânico na dieta de frangos de corte de 1 a 21 dias de idade

Maria Márcia Pereira Sartori<sup>1</sup>, Juliana Cristina Ramos Rezende<sup>1</sup>, Antônio Celso Pezzato<sup>1</sup> e José Roberto Sartori<sup>1</sup>

<sup>1</sup> Universidade Estadual Paulista, SP, Brasil  
Email: mmportsartori@ibb.unesp.br

**Resumo:** Foi realizado um experimento para avaliar o efeito dos níveis de zinco na dieta de frangos de corte aos 21 dias de idade sobre o desempenho, ganho de peso e consumo de ração. O ensaio foi realizado em aviário experimental com delineamento inteiramente casualizado com cinco níveis de zinco inorgânico, totalizando cinco tratamentos com sete repetições. As aves foram divididas em boxes, sendo cada box uma unidade experimental com 33 aves cada. Foi avaliado os parâmetros de desempenho zootécnico, ganho de peso corporal e consumo de ração de 1 a 21 dias de idade. Dois métodos de determinação do ponto ótimo foram avaliados para os dados com diferentes níveis de zinco inorgânico, aplicando análise de regressão linear e o método de broken line, com auxílio do programa estatístico SAS (2002). O nível ótimo de zinco foi superestimado quando aplicado métodos de regressão linear usual.

## CP79 – Modelagem linear mista aplicada a dados de expectativa de vida da população

Marina Rodrigues Maestre<sup>1</sup> e Elias Silva de Medeiros<sup>2</sup>

<sup>1</sup> Universidade Estadual de Mato Grosso do Sul, MS, Brasil

<sup>2</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** marina.maestre.estadistica@gmail.com

**Resumo:** Em diversas áreas de pesquisa são comuns as situações em que são tomadas medidas repetidas de uma ou mais variáveis respostas em ocasiões sucessivas na mesma unidade experimental, ao longo de um certo intervalo de tempo (fator longitudinal). É comum admitir-se que as observações feitas repetidamente em tempos distintos sejam correlacionadas e tenham variâncias heterogêneas. Para analisar esse tipo de dados podem-se utilizar modelos (lineares ou não lineares) mistos, que tem o atrativo de explicar de forma conveniente a variabilidade existente entre unidades experimentais e as correlações dentro dessas unidades. O objetivo deste trabalho foi utilizar a classe de modelos lineares mistos para ajustar um modelo para um conjunto de dados de expectativa de anos de vida da população (EAV) ao longo do tempo nas capitais das Regiões Nordeste (NE) e Sudeste (SE). Foram utilizados testes como o da razão de verossimilhanças, os critérios de informação AIC e BIC e gráficos de diagnóstico dos resíduos condicionais padronizados para alcançar o melhor modelo. Para a parte fixa, o modelo resultante ficou com duas curvas quadráticas, sendo uma para cada região. Estas curvas diferiram somente pelo intercepto (paralelas). Quanto a inclusão de efeito aleatório, foi significativa somente no Intercepto da curva NE. Por meio dos testes utilizados, houve heterocedasticidade para os níveis do fator Ano.

---

## CP80 – Evaluation of allelopathic effect of *Senecio* sp. in the germination of pearl millet seeds (*Pennisetum glaucum*) through a log-gamma model

Matheus Henrique Dal Molin Ribeiro<sup>1</sup>, Vanderlei Aparecido de Lima<sup>1</sup>, Patrícia Carla Giloni de Lima<sup>2</sup> and Flávia Boruch<sup>2</sup>

<sup>1</sup> Universidade Tecnológica Federal do Paraná, PR, Brasil

<sup>2</sup> Universidade Estadual do Centro-Oeste, PR, Brasil

**Email:** pgiloni2012@gmail.com

**Abstract:** Today the plant allelochemicals extraction has been of great interest to replace the use of pesticides. The bioassays germination evaluating the allelopathic effect of these substances in seeds. Thus, the aim of this study was to apply the methodology of generalized linear models (GLM) to evaluate the effect of hexanolic extract *Senecio* sp. on the germination of pearl millet (*Pennisetum glaucum* L. R. Brown) under laboratory conditions. Three attributes were evaluated: variable rate, index and germination rate. In this sense, glm gamma, with logarithmic link function was used in data modeling and contrast for comparison between groups. Residual and diagnostics analysis were satisfactory, since the statistical results corroborated with biological evidence. From this methodology, it was found that the time and the speed of germination and germination index were different when compared to controls millet germination bioassays ( $p < 0.05$ ).

## CP81 – Seleção de modelos lineares misto na estimativa de quantidade de beta-caseína do leite UAT e comparação entre os softwares SAS e R

Michele Barbosa<sup>1</sup>, Márcio Balestre<sup>1</sup> e César Gonçalves de Lima<sup>2</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> Universidade de São Paulo, SP, Brasil

**Email:** michele.barbosa@unifal-mg.edu.br

**Resumo:** O enfoque do trabalho foi o desenvolvimento de um modelo linear misto para a estimativa de quantidade de proteína beta-caseína do leite UAT, através da escolha de diferentes estruturas de covariâncias e de modelos de regressão para as médias dos grupos. Realizando comparações entre o proc Mixed do SAS e os argumentos da função lme() do software R observando as vantagens e restrições destes softwares. Após a modelagem através dos modelos propostos, verificou-se uma significativa melhora dos parâmetros das informações estatísticas (AIC e BIC).

---

## CP82 – Time-to-event analysis of subcellular exocytotic events of insulin release

Morten Gram Pedersen<sup>1</sup>, Giuliana Cortese<sup>1</sup>, Nikhil Gandasi<sup>2</sup> and Sebastian Barg<sup>2</sup>

<sup>1</sup> University of Padova, Italy

<sup>2</sup> Uppsala University, Sweden

**Email:** pedersen@dei.unipd.it

**Abstract:** Whereas time-to-event analysis is well established for clinical studies, it has so far not been applied to cell biological data. Here we investigate the subcellular events underlying insulin release from secretory granules in healthy and diabetic pancreatic beta-cells. Reduced insulin secretion is the major cause of diabetes, one of the most frequent chronic diseases in the world. Insulin is secreted in exocytotic events that can be followed with state-of-the-art imaging techniques. Exocytosis is triggered by calcium, entering the cell via voltage-sensitive calcium channels (“CaV proteins”), which can be followed over time. Since the absolute calcium concentration cannot be reliably estimated, we treated the rise-time of calcium at each granule as covariate in a Cox regression model with exocytosis as the events of interest. In another dataset, the amount of CaV proteins was considered as time-dependent covariate. Granules within a cell were treated as clustered data and marginal Cox models were used to estimate robust standard errors. Since the recordings of calcium and CaV proteins were performed only for a randomly selected subset of the granules that failed to undergo exocytosis during the experiments, a weighted Cox regression model was used, following the lines of a case-cohort analysis. Our results confirm biological intuition, but moreover allow quantification of the effects and the uncertainty over time of the experiment. We believe this novel application of time-to-event analysis should be of general interest to the analysis of cellular events. Moreover, this context presents statistical challenges, which have so far remained unaddressed.

## CP83 – Efeito da interação entre fontes e doses de biomassa carbonizada na produção de mudas de uma espécie florestal na região Amazônica

Paulo Souza<sup>1</sup>, Gustavo Silva<sup>1</sup>, Carlos Felipe Link<sup>1</sup>, Roni Balduino<sup>1</sup>, Ernandes Moura<sup>2</sup>, Jodean Silva<sup>2</sup> e Fabrício Andrade<sup>1</sup>

<sup>1</sup> Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso, MT, Brasil

<sup>2</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** ernandesfederal@hotmail.com

**Resumo:** O *Enterolobium contortisiliquum* (Vell.) Morong, conhecido popularmente como orelha de macaco apresenta crescimento inicial rápido. Objetivou-se com o presente trabalho avaliar o desempenho de fontes e doses de biomassa carbonizada, na produção de mudas de tamboril. O experimento foi conduzido em estufa com telado sombreado a 50% no viveiro de pesquisas de plantas lenhosas do IFMT, campus de Juína – MT. O delineamento experimental foi o inteiramente casualizado, em esquema fatorial 2 x 4, sendo os fatores constituídos por duas fontes de biomassa carbonizada (maravalha carbonizada e moinha de carvão vegetal) e quatro concentrações destes resíduos (0; 80; 160 e 240 g dm<sup>-3</sup>) com quatro repetições. Avaliou-se a percentagem de germinação e mortalidade, altura de plantas, diâmetro do coleto e teor relativo de clorofila. Com exceção da percentagem de plantas mortas, todas as demais variáveis morfofisiológicas foram significativamente influenciadas pelos tratamentos. Os dados de altura e diâmetro do coleto de plantas ajustaram-se ao modelo quadrático de regressão com os valores máximos para esses parâmetros a partir da aplicação acima de 144 g dm<sup>-3</sup> de biomassa carbonizada no substrato. Recomenda-se o enriquecimento dos substratos formados a partir de solo e esterco com biomassa carbonizada.

## CP84 – Análise de modelos probabilísticos para os dados mensais de chuva no município de Toledo - PR

Rúbia Wilhelms França<sup>1</sup> e Rosângela Aparecida Botinha Assumpção<sup>1</sup>

<sup>1</sup> Universidade Tecnológica Federal do Paraná, PR, Brasil

**Email:** rosangelaa@utfpr.edu.br

**Resumo:** A utilização de estimativas da necessidade da água da chuva através de modelos probabilísticos é de extrema importância para a obtenção de resultados confiáveis. No presente trabalho utilizou-se uma série histórica de precipitações pluviométricas mensais no período de janeiro de 2003 a dezembro de 2014, totalizando 12 anos. O objetivo foi avaliar a variabilidade mensal por meio da distribuição temporal e analisar o potencial das distribuições de probabilidade Weibull, Gama, Log-Normal e Exponencial na determinação de totais mensais de precipitação na cidade de Toledo, Paraná. A aderência das distribuições decorrente dos dados pluviométricos mensais foi verificada via teste de Kolmogorov-Smirnov. As funções Weibull e Gama apresentaram aderência, porém os resultados obtidos indicaram que o modelo Gama ajustou-se mais adequadamente a precipitação mensal da cidade de Toledo.

Part X

**Contributed Posters 2**



---

## Sessão de Pôsters 2 / Poster Session 2

---

### Bioestatística e Modelos Biométricos 2

[*Biostatistics and Biometrical Models 2*]

---

#### CP85 – Previsão do número de pousos no aeroporto Santa Maria em Aracaju-SE

Carlos Silva<sup>1</sup>, Juliana Silva<sup>1</sup> e Suzana Russo<sup>1</sup>

<sup>1</sup> Universidade Federal de Sergipe, SE, Brasil

**Email:** chsestatistica@gmail.com

**Abstract:** The Santa Maria Airport, located in Aracaju / SE, besides serving the local population, it also serves the executives and tourists seeking the state of Sergipe for business and leisure. Daily, six thousand people circulate through the airport terminal, totaling nineteen daily flights and a monthly average of sixty thousand passengers. According to Infraero airport has an airport site with an area of 3,874,742.13m<sup>2</sup>, with patio 22,356 aircraft m<sup>2</sup>, four positions for parking aircraft and passenger terminal with an area of 9321 m<sup>2</sup>. The work intends to use the models of Holt-Winters to find a predictive model that fits well to the landing data in Aracaju.

---

#### CP86 – Um método de amostragem georreferenciado para fiscalização, controle e erradicação da doença grenning (Huanglongbing/HLB) da cultura do citros com base na distribuição Beta-Binomial

Edilenia Queiroz Pereira<sup>1</sup>, Marcos Jardel Henriques<sup>1</sup>, Oilson Alberto Gonzatto Junior<sup>1</sup>, Terezinha Aparecida Guedes<sup>1</sup>, Vanderly Janeiro<sup>1</sup> e Ana Beatriz Tozzo Martins<sup>1</sup>

<sup>1</sup> Universidade Estadual de Maringá, PR, Brasil

**Email:** edileniaqueiroz@yahoo.com.br

**Resumo:** Atualmente a fruta mais consumida e produzida no mundo é a laranja (*Citrus sinensis* L. Osbeck) sendo o Brasil, o maior produtor. Assim como as demais culturas agrícolas, a da laranja necessita de manejos e cuidados especiais contra pragas, plantas daninhas e, principalmente, doenças para a garantia de produção. Atualmente merece destaque a doença de huanglongbing que, praticamente, dizimou plantações inteiras de laranja em Estados norte-americanos em curto prazo de tempo, pelo simples fato de profissionais e agricultores não terem se dado conta da gravidade da doença. Como ainda não há controle da infecção a única maneira, até então, de tentar fazer com que esta não aumente o número de plantas infectadas é realizar inspeções regularmente, eliminar plantas doentes e controlar o inseto vetor

(psilídeo) com pulverizações periódicas. Dependendo da porcentagem de plantas infectadas no pomar, por lei, o agricultor é obrigado a arrancar todas as plantas da área, pois, sem os devidos cuidados, em pouco mais de cinco anos uma área inteira pode se tornar improdutiva para a cultura da laranja. Desta forma, torna-se imprescindível o uso de técnicas de amostragem para a coleta de dados para se detectar o percentual de plantas infectadas em pomares. Em virtude disso, a Agência de Defesa Agropecuária do Paraná (ADAPAR), que por meio de uma parceria com o Curso de Graduação em Estatística do Departamento de Estatística da Universidade Estadual de Maringá – PR, conseguiu elaborar planos amostrais utilizando-se a distribuição Beta-binomial.

---

## CP87 – Otimização na seleção entre famílias de cana-de-açúcar via classificação por K-nearest neighbor

Édimo Moreira<sup>1</sup>, Luiz Peternelli<sup>2</sup> e Kelly Rosa<sup>1</sup>

<sup>1</sup> Instituto Federal do Triângulo Mineiro, MG, Brasil

<sup>2</sup> Universidade Federal de Viçosa, MG, Brasil

**Email:** edimo.fernando@gmail.com

**Resumo:** Um dos grandes desafios nos diversos programas de melhoramento é a seleção de genótipos promissores nas fases iniciais do programa. O objetivo deste trabalho é comparar a classificação via K-Nearest Neighbor (KNN), num contexto de aprendizado estatístico, e a seleção via tonelada de cana por hectare estimada (TCHe), como alternativas para seleção entre famílias de cana-de-açúcar. Os dados utilizados neste trabalho foram provenientes de 5 experimentos, com 22 famílias cada, no delineamento em blocos casualizados com 5 repetições. Nestes experimentos foram coletados os componentes de produção, número de colmos (NC), diâmetro de colmos (DC) e a altura de colmos (AC), bem como foi obtida a produção real, expressa em tonelada de cana por hectare (TCHr). Para o treinamento dos métodos de seleção foram utilizados como variáveis de entrada os caracteres de produção NC, DC e AC. A variável de saída utilizada foi a TCHr. Previamente a análise, os valores de NC, DC e AC foram padronizados e, foram simulados valores de NC, DC, AC e TCHr para 1,000 famílias com base na estrutura de médias e covariância de cada i-ésimo experimento. As análises foram processadas em cinco diferentes cenários de acordo com o experimento utilizado para simulação. A seleção via KNN apresenta alta concordância com o método de seleção ideal, via TCHr, podendo ser utilizada com sucesso na seleção entre famílias de cana-de-açúcar. A seleção via KNN apresenta desempenho.

---

## CP88 – Modelagem da variabilidade em experimentos de mistura com variáveis de ruído

Edmilson Rodrigues Pinto<sup>1</sup>, Leandro Alves Pereira<sup>1</sup> e Aurélia A. de A. Rodrigues<sup>1</sup>

<sup>1</sup> Universidade Federal de Uberlândia, MG, Brasil

**Email:** edmilson@famat.ufu.br

**Resumo:** Em experimentos de mistura com variáveis de ruído ou variáveis de processo que não podem ser controladas, investigar e tentar controlar a variabilidade da variável resposta é muito importante, especialmente para a melhoria da qualidade em processos industriais. Desta forma, modelar a variabil-

idade em experimentos de mistura com variáveis de ruído vem a ser de grande importância e tem sido considerado na literatura com abordagens que requerem a escolha de uma função de perda quadrática ou pelo uso do método delta. Neste trabalho, é utilizado o método delta e também é proposta uma abordagem alternativa, a qual é baseada na modelagem conjunta da média e dispersão (MCMD). As duas técnicas são utilizadas em um experimento de mistura com variáveis de ruído a fim de obter modelos conjuntos para média e variância. Seguindo as ideias de Taguchi sobre planejamento robusto é construído e resolvido um problema de otimização para minimizar a variância enquanto a média fica próxima a um valor alvo. No final, é fornecido uma discussão sobre as duas metodologias consideradas.

---

## CP89 – Análise de trilha em caracteres de genótipos de milho de ciclo precoce

Jéssica Andiará Kleinpaul<sup>1</sup>, Alberto Cargnelutti Filho<sup>1</sup>, Bruna Mendonça Alves<sup>1</sup>,  
Giovani Facco<sup>1</sup>, Rosana Marzari Thomasi<sup>1</sup>, Daniela Barbieri Uliana<sup>1</sup>, Cleiton  
Antonio Wartha<sup>1</sup>, Rafael Vieira Pezzini<sup>1</sup> e Daniela Lixinski Silveira<sup>1</sup>

<sup>1</sup> Universidade Federal de Santa Maria, RS, Brasil

**Email:** alberto.cargnelutti.filho@gmail.com

**Resumo:** O objetivo deste trabalho foi avaliar as relações lineares entre caracteres fenológicos, morfológicos e produtivos de genótipos de milho de ciclo precoce. No ano agrícola 2014/2015 foi conduzido um experimento com 16 genótipos de milho de ciclo precoce, no delineamento experimental blocos ao acaso, com três repetições. Em cada unidade experimental foram mensurados os caracteres número de dias da sementeira até a emergência, número de dias da sementeira até 50% do florescimento masculino, número de dias da sementeira até 50% do florescimento feminino, altura de plantas na colheita, altura de espigas na colheita, posição relativa da espiga, número de plantas, número de espigas, índice de espigas e produtividade de grãos. Inicialmente, foram verificados os pressupostos de normalidade dos erros e de homogeneidade de variâncias residuais, pelos testes de Kolmogorov Smirnov e Bartlett, respectivamente. Os dados foram submetidos à análise de variância e teste F. As médias foram comparadas por meio do teste de Scott Knott a 5% de probabilidade. Foi determinada a matriz de coeficientes de correlação fenotípica entre os caracteres e realizado o diagnóstico de multicolinearidade pelo número de condição. Após, realizou-se a análise de trilha considerando a produtividade de grãos como o caractere principal e os demais caracteres como explicativos e, realizada a análise de regressão stepwise. O número de espigas apresenta relação linear positiva com a produtividade de grãos e pode ser utilizado para a seleção indireta.

---

## CP90 – Análise de agrupamento em genótipos de milho de ciclos precoce e superprecoce

Lucas Silva Stefanelo<sup>1</sup>, Alberto Cargnelutti Filho<sup>1</sup>, André Lavezo<sup>1</sup>, Bruna  
Mendonça Alves<sup>1</sup>, Gabriela Görden Chaves<sup>1</sup>, Jéssica Andiará Kleinpaul<sup>1</sup>, Cláudia  
Marques de Bem<sup>1</sup>, Cleiton Antonio Wartha<sup>1</sup> e Daniela Lixinski Silveira<sup>1</sup>

<sup>1</sup> Universidade Federal de Santa Maria, RS, Brasil

**Email:** lstefanelo@gmail.com

**Resumo:** Este trabalho teve como objetivo avaliar a divergência genética, por meio de análise de agrupamento de 16 genótipos de milho de ciclo precoce e dez genótipos de milho de ciclo superprecoce. Foram avaliados os caracteres: número de dias da sementeira até a emergência; número de dias da sementeira até 50% do florescimento masculino; número de dias da sementeira até 50% do florescimento feminino; altura de planta na colheita; altura de espigas na colheita; posição relativa da espiga; estande final; número de espigas; índice de espigas e produtividade de grãos. Para cada caractere, foi realizada a análise de variância e as médias foram comparadas por meio do teste de Scott Knott, a 5% de probabilidade. Determinou-se a matriz de coeficientes de correlação fenotípica e realizou-se o diagnóstico de multicolinearidade. Determinou-se a contribuição relativa de cada caractere para divergência genética. Para a análise de agrupamento, foi utilizada a distância generalizada de Mahalanobis como medida de dissimilaridade e os grupos foram formados pelo método vizinho mais próximo. Foram construídos dendrogramas, calculados os coeficientes de correlação cofenética, distorção e estresse e realizado o corte, usando como critério de significância entre grupos 35% e 45% de dissimilaridade nos genótipos precoce e superprecoce, respectivamente. Há variabilidade genética para genótipos de milho ciclos precoce e superprecoce. Por meio do dendrograma há formação de onze grupos para os genótipos de milho de ciclo precoce e sete grupos para os genótipos de milho de ciclo superprecoce.

## CP91 – Análise de trilha em caracteres de genótipos de milho de ciclo superprecoce

Rosana Marzari Thomasi<sup>1</sup>, Alberto Cargnelutti Filho<sup>1</sup>, Bruna Mendonça Alves<sup>1</sup>,  
Denison Esequiel Schabarum<sup>1</sup>, Jéssica Andiará Kleinpaul<sup>1</sup>, Cleiton Antonio  
Wartha<sup>1</sup>, Daniela Barbieri Uliana<sup>1</sup>, Fernanda Martins Simões<sup>1</sup> e Rafael Vieira  
Pezzini<sup>1</sup>

<sup>1</sup> Universidade Federal de Santa Maria, RS, Brasil

**Email:** rvpezzini@hotmail.com

**Resumo:** O objetivo deste trabalho foi avaliar as relações lineares entre a produtividade de grãos e os caracteres agrônômicos de milho. O experimento foi realizado no ano agrícola 2014/2015, no delineamento blocos ao acaso com três repetições e foram avaliados dez genótipos de milho de ciclo superprecoce. Em cada parcela foram mensurados os caracteres fenológicos: número de dias da sementeira até a emergência, número de dias da sementeira até 50% do florescimento masculino e número de dias da sementeira até 50% do florescimento feminino; morfológicos: altura de plantas na colheita, altura de espigas na colheita e posição relativa da espiga; produtivos: número de plantas, número de espigas, índice de espigas e produtividade de grãos. As análises estatísticas foram iniciadas com a verificação dos pressupostos de normalidade dos erros e de homogeneidade de variâncias residuais, pelos testes de Kolmogorov Smirnov e Bartlett, respectivamente. Após, foi realizada a análise de variância e o teste F. As médias foram comparadas por meio do teste de Scott Knott a 5% de probabilidade. Foi determinada a matriz de coeficientes de correlação fenotípica entre os caracteres e realizado o diagnóstico de multicolinearidade pelo número de condição. Após, realizou-se a análise de trilha considerando a produtividade de grãos como o caractere principal e os demais caracteres como explicativos e, realizada a análise de regressão stepwise. A posição relativa da espiga e o número de espigas apresenta relação linear positiva com a produtividade de grãos. A altura de planta tem relação linear negativa com a produtividade de grãos.

## CP92 – Improvement of impact assessment through combination of alternative strategies: Empirical results of a training program for rice production in Indonesia

Samila Sena<sup>1</sup> and Leila D.A.F. Amorim<sup>1</sup>

<sup>1</sup> Universidade Federal da Bahia, BA, Brasil

**Email:** mylasenna@hotmail.com

**Abstract:** In observational studies, it is not possible to make random allocation of units to groups, leading to analysis that require special methods to adjust the effects for possible differences in the covariates between the groups. There are several techniques to take this into account, including Propensity Score Matching (PSM), Regression Discontinuity Design (RDD), Instrumental Variables (IV) and Difference-in-Difference (DD). In general, these methods are used to reduce bias and to correct the differences in covariates between groups, making the groups comparable. This work considers PSM, DD and its combination for impact assessment. Information about a training program among farmers in Indonesia is used for illustration. The program was created in the late 80's in order to disseminate concepts and practices for rice-based production in East Asia, leading to changes in the farmer's behavior with respect to pest management, development of their analytical skills, critical thinking and creativity. The use of improved methods for impact assessment can contribute for a better evaluation of effectiveness of intervention programs in the communities.

---

## CP93 – Aceitabilidade de pão tipo bisnaguinha adquirido na cidade de Araras/SP por meio de mapas de preferência

Simone Daniela Sartorio<sup>1</sup>, Marta Regina Verruma-Bernardi<sup>1</sup>, Carolina Ibelli Bianco<sup>1</sup>, Jéssica Adriana de Jesus<sup>1</sup> e Ionele do Socorro Ferreira Almeida<sup>1</sup>

<sup>1</sup> Universidade Federal de São Carlos, SP, Brasil

**Email:** sisartorio@cca.ufscar.br

**Resumo:** A maioria dos resultados de testes com consumidores vem sendo avaliados por ANOVA e testes de comparações múltiplas de médias, sendo uma análise global univariada, em que considera conjuntamente as validações de todos os consumidores. O Mapa de Preferência é de grande utilidade, sendo um procedimento multivariado que considera a individualidade dos consumidores e não somente a média do grupo. O objetivo do trabalho foi verificar a aceitabilidade e a preferência entre 8 diferentes marcas comerciais do pão tipo bisnaguinha presentes na cidade de Araras/SP, por meio da técnica de mapas de preferência. Cinquenta consumidores do pão realizaram o teste de aceitabilidade. Os mapas mostram que houve diferença entre as marcas de pão e permitiu visualizar que dentro do grupo das amostras com menor aceitação global estavam as duas marcas restritas ao comércio da cidade de Araras e região. Já as marcas de maior aceitação foram aquelas que apresentaram bisnaguinhas com os seguintes atributos: coloração caramelo da casca mais clara; casca com aparência lisa e homogênea; maciez e elasticidade tátil medianas; considerável aderência ao palato quando consumida; presença de aroma fermentado, porém não tão intenso quanto ao aroma de massa de pão crua contendo fermento biológico; aroma adocicado, mas não tão doce quanto ao de pão doce; um leve aroma de queimado da casca, além de sabor não tão doce e com certo residual de fermentado. Pode-se inferir que as marcas mais disseminadas são as responsáveis

por estabelecerem os atributos e as respectivas intensidades que caracterizam um pão “tipo bisnaginha”.

---

## Sessão de Pôsters 2 / Poster Session 2

---

### Análise de Sobrevivência [*Survival Analysis*]

---

#### CP94 – The Kumaraswamy transmuted Fréchet distribution

Angela Lima da Silva<sup>1</sup>, Giovana Oliveira Silva<sup>1</sup>, Jailson de Araujo Rodrigues<sup>2</sup> and Marcio Claudio Mercês Brito<sup>3</sup>

<sup>1</sup> Universidade Federal da Bahia, BA, Brasil

<sup>2</sup> Instituto Federal de Educação, Ciência e Tecnologia da Bahia, BA, Brasil

<sup>3</sup> Instituto Federal de Educação, Ciência e Tecnologia Baiano, BA, Brasil

**Email:** angelalimadasilva@yahoo.com.br

**Abstract:** In this work is introduced a new distribution called the Kumaraswamy transmuted Fréchet. This distribution includes as special cases the inverse exponential, Fréchet, inverse Rayleigh and exponentiated Fréchet distributions. The study examines various properties of the new model. The flexibility of this distribution is illustrated in an application to a real data set.

---

#### CP95 – Modelo de regressão para eventos recorrentes utilizando a distribuição Weibull modificada em análise de sobrevivência

Jackson Santos da Conceição<sup>1</sup> e Giovana Oliveira Silva<sup>1</sup>

<sup>1</sup> Universidade Federal da Bahia, BA, Brasil

**Email:** giovana@ufba.br

**Resumo:** Em análise de sobrevivência, o evento de interesse pode ocorrer várias vezes no decorrer do acompanhamento de uma dada unidade de investigação. Esses eventos são chamados de eventos recorrentes e surgem em várias áreas, por exemplo, estudos biomédicos, criminologia, demografia e confiabilidade. Contudo a estratégia mais usada em análise de sobrevivência é o modelo de Cox que sob a ótica de dados com eventos recorrentes não é adequado devido a sua suposição de independência. Alguns autores propuseram como alternativa o uso dos modelos marginais e condicionais que são extensões do modelo de Cox para análises de dados com eventos recorrentes. Outra proposta para analisar dados com eventos recorrentes surge do uso de modelos multiplicativos que sob o ponto de vista paramétrico utiliza uma distribuição de probabilidade para a função de risco basal. Em geral, as distribuições de probabilidades utilizadas nesta abordagem são a distribuição exponencial e a distribuição Weibull. Dessa forma, será proposto neste trabalho o uso da distribuição Weibull Modificada, que diferentemente das distribuições exponencial e Weibull apresenta função de taxa de falha em forma de U, para função de risco basal do modelo de regressão multiplicativo avaliando-se a adequacidade do modelo e as estimativas dos parâmetros dos modelos multiplicativos paramétricos através de estudos de simulação e um exemplo de aplicação

desta metodologia.

---

## CP96 – The exponentiated generalized Lindley distribution: Properties and application

Jailson A. Rodrigues<sup>1</sup>, Ana Paula Silva<sup>2</sup> and Angela Silva<sup>3</sup>

<sup>1</sup> Instituto Federal da Bahia, BA, Brasil

<sup>2</sup> Universidade Federal São João del-Rei, MG, Brasil

<sup>3</sup> Universidade Federal da Bahia, BA, Brasil

**Email:** jailsondearaujo@yahoo.com.br

**Abstract:** A new distribution called the exponentiated generalized Lindley is proposed and studied. This distribution includes as special cases the Lindley and exponentiated Lindley distributions. Some structural properties of the proposed distribution are studied, such as moments, moment generating function, characteristic function, distribution of the order statistics. The flexibility of this distribution is illustrated in an application to a real data set.

---

## CP97 – Análise de sobrevivência aplicada a durabilidade de pneus 11.00R22

Pablo Lourenço<sup>1</sup>, Elias Silva<sup>1</sup>, Tiago Almeida<sup>2</sup> e Isaac Pereira<sup>2</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> Universidade Estadual da Paraíba, PB, Brasil

**Email:** pablo\_lourenco@hotmail.com

**Resumo:** O estudo da durabilidade (quilometragem) dos pneus em caminhões que transportam produtos agrícolas se faz importante, pois, por meio deste é possível antecipar a necessidade de troca ou recauchutagem antes que o pneu venha a falhar. Caso essa análise não seja executada, o transporte corre o grande risco de ficar por horas, ou até dias, parado em uma borracharia, impactando na produção. Assim, a realização desse estudo é de fundamental importância no escoamento da safra, no que diz respeito à logística no processo de transporte. Uma ferramenta estatística importante para esse tipo de estudo é a Análise de Sobrevivência, também conhecida como teoria da Confiabilidade. Neste estudo utilizamos o estimador de Kaplan-Meier para obter as estimativas de sobrevivência, em que se observou um tempo médio de falha no instante que o pneu completava 31201,45 Km. Foi utilizado o teste Log-rank com o interesse de verificar estatisticamente se existe diferença entre os três tipos de vida dos pneus, no qual o teste foi significativo entre os três grupos, indicando que houve diferença no tempo de vida desses grupos. Para se verificar a adequabilidade dos dados a uma determinada distribuição, foi utilizado o teste de Anderson – Darling, em que se observou que a distribuição Weibull melhor se adequou aos dados de durabilidade dos pneus.

---

## Sessão de Pôsters 2 / Poster Session 2

---

### Estatística Multivariada [Multivariate Statistics]

---

#### CP98 – Analysis with distal outcomes: latent effects on ENADE performance

Ana Clara Campos<sup>1</sup>, Sheila Regina Pereira<sup>1</sup>, Samila Sena<sup>1</sup>, Lilia Carolina Costa<sup>1</sup>  
and Leila D.A.F. Amorim<sup>1</sup>

<sup>1</sup> ISC, IM e FACED/Universidade Federal da Bahia, BA, Brasil  
**Email:** leila.d.amorim1@gmail.com

**Abstract:** New statistical methods have been proposed to deal with complex research questions and data coming from non-trivial sample designs. One of the simplest methods involving mixture models is the Latent Class Analysis (LCA), which has been widely used in a variety of problems to identify subgroups of individuals based on observed response patterns. A recent approach includes LCA with distal outcomes, which aims to assess the impact of latent variables on observed (distal) outcomes. Some alternatives procedures to estimate such effects have been discussed in the literature, including pseudo class method (Wang et al., 2005; Clark & Muthén, 2009), a method based on Bayes Theorem (Lanza tan, and Bray, 2013) and a method that incorporates measurement error, proposed by Vermunt (2010) and later extended by Muthén for distal outcomes (Asparouhov & Muthén, 2014). In this work these methods are presented and compared to estimate effects related to a continuous distal outcome. The illustration is done through the National Student Performance Examination Data (ENADE) from 2010.

---

#### CP99 – A new approach for testing the genetic covariance

Anderson Silva<sup>1</sup>, Carlos Tadeu dos Santos Dias<sup>2</sup>, Claudio Souza Junior<sup>2</sup>, Trevor Bailey<sup>3</sup> and Wojtek Krzanowski<sup>3</sup>

<sup>1</sup> Instituto Federal Goiano, GO, Brasil

<sup>2</sup> ESALQ/Universidade de São Paulo, SP, Brasil

<sup>3</sup> University of Exeter, UK

**Email:** anderson.silva@ifgoiano.edu.br

**Abstract:** When a genetic factor is being studied for more than one response variable, estimates of the genetic covariances are essential, especially for genetic breeding programs. In a multivariate analysis of variance, mean cross-products (covariances) due to genetic and residual (often called environmental) effects are computed. Thus, an important question would be: is the genetic covariance significant in comparison with the residual one? To answer this question we need to quantify the magnitude of the

joint variation of two response variables due to the genetic effect with respect to the variation due to the residual effect. We proposed a test based on the empirical distribution of the simple ratio of genetic and residual mean cross-products under the null hypothesis (null genetic covariance). The empirical distribution was obtained by using a Monte Carlo process to perform simulations of Wishart matrices based on estimates of covariances. We have noticed distribution densities similar to those of nonstandard Cauchy.

---

## CP100 – Correlação canônica entre caracteres fenológicos, morfológicos e produtivos de cultivares de milho

André Lavezo<sup>1</sup>, Alberto Cargnelutti Filho<sup>1</sup>, Lucas da Silva Stefanelo<sup>1</sup>, Bruna Mendonça Alves<sup>1</sup>, Jéssica Andiar Kleinpaul<sup>1</sup>, Diego Nicolau Follmann<sup>1</sup>, Gabriela Görgen Chaves<sup>1</sup>, Cleiton Antonio Wartha<sup>1</sup> e Fernanda Martins Simões<sup>1</sup>

<sup>1</sup> Universidade Federal de Santa Maria, RS, Brasil

**Email:** cleiton.ufsm@gmail.com

**Resumo:** O objetivo do trabalho foi avaliar as relações lineares entre os grupos de caracteres fenológicos e produtivos e os grupos de caracteres morfológicos e produtivos, de cultivares de milho de ciclo precoce e superprecoce. No ano agrícola 2014-2015 foi conduzido um experimento com 16 cultivares de milho de ciclo precoce e outro com dez cultivares de milho de ciclo superprecoce, em blocos ao acaso, com três repetições. Em cada unidade experimental mensurou-se os caracteres fenológicos: número de dias da sementeira até a emergência, número de dias da sementeira até 50% do florescimento masculino e número de dias da sementeira até 50% do florescimento feminino; morfológicos: altura de planta, altura de espiga e posição relativa da espiga; e produtivos: número de plantas, número de espigas, índice de espigas e produtividade de grãos. Os dados foram submetidos à análise de variância, e realizados os testes da normalidade dos erros e da homogeneidade de variâncias residuais. Foi determinada a matriz de correlação fenotípica e realizado o diagnóstico de multicolinearidade. A seguir, foi realizada a análise de correlação canônica entre os grupos de caracteres: fenológicos e produtivos, morfológicos e produtivos. Nas cultivares de milho de ciclos precoce e superprecoce, há associação linear entre os caracteres fenológicos e produtivos e os caracteres morfológicos e produtivos. Maior associação ocorre entre o número de dias da sementeira até 50% do florescimento masculino e o número de plantas, e entre a altura de espiga e a produtividade de grãos, para cultivares de ciclos precoce e superprecoce, respectivamente.

---

## CP101 – LISA cluster maps of incidence rates of occupational accidents in Brazil

Andre Luis Santiago Maia<sup>1</sup>, Isabelle Galdino de Almeida<sup>1</sup> and Gecynalda Soares da Silva Gomes<sup>2</sup>

<sup>1</sup> MTPS/Fundação Jorge Duprat e Figueiredo, BA, Brasil

<sup>2</sup> Universidade Federal da Bahia, BA, Brasil

**Email:** andresantiagomaia@gmail.com

**Abstract:** The intensive process of economic growth and job creation that Brazil experienced in the last years often is associated with an important dimension where this process is far from to be satisfactory: the high incidence rates of occupational accidents. Hence, it is interesting to analyze the spatial distribution of this rate in order to identify critical regions of incidence of occupational accidents in Brazil and thus to detect possible areas with an excess of the issue under study. Results for the years from 2002 to 2012 are presented and show, for these years, significant spatial dependence between the incidences rates registered in Brazilian municipalities and some critical regions in the country.

---

## CP102 – Seleção de descritores agronômicos em acessos de tabaco por meio de Técnicas Multivariadas

Antonio Leandro da Silva Conceição<sup>1</sup>, Carlos Alberto da Silva Ledo<sup>2</sup>, Mauricio dos Santos da Silva<sup>1</sup>, Clailto Carvalho dos Santos<sup>3</sup> e Sandra Domingos João Afonso<sup>1,4</sup>

<sup>1</sup> Universidade Federal do Recôncavo da Bahia, BA, Brasil

<sup>2</sup> Embrapa Mandioca e Fruticultura, BA, Brasil

<sup>3</sup> Ermor Tabarama Tabacos do Brasil Ltda., BA, Brasil

<sup>4</sup> Instituto Superior Politécnico do Kwanza Sul, Angola

**Email:** leandrosilvaufbr@hotmail.com

**Resumo:** O objetivo deste trabalho foi selecionar descritores agronômicos e determinar sua importância relativa na caracterização, bem como propor um número mínimo capaz de quantificar a divergência entre acessos de tabaco no intuito de fornecer maior conhecimento da variabilidade genética para conservação e melhoramento genético da espécie. Para a caracterização de 15 acessos de tabaco tipo Sumatra foram usados 17 descritores quantitativos. Foi realizada a identificação dos descritores quantitativos redundantes, por dois procedimentos: 1) seleção direta, proposta por Jolliffe e 2) Seleção baseada no coeficiente de Singh. Para auxiliar na decisão de descarte, foram estimados os coeficientes de correlação de Spearman entre todos os descritores. Foram selecionados 10 descritores quantitativos (altura da planta, número de folhas, diâmetro médio do caule, Comprimento da 3ª folha, comprimento da 10ª folha, largura da base da 10ª folha, ângulo de inserção da 10ª folha, comprimento dos internódios, engrossamento do tubo da flor e comprimento da corola), os quais são importantes na caracterização do germoplasma de tabaco em estudo. O descarte de 41,18% dos descritores não provocou perda de informação considerável, uma vez que os descritores redundantes estão correlacionados aos descritores remanescentes, em que esse descarte também possibilitará a redução de custos e dará maior dinâmica ao manejo e caracterização da cultura.

---

## CP103 – Técnicas multivariadas aplicadas a biomarcadores metabólicos e o método Z-scan

Aubin Elisete<sup>1</sup>, Viviana Giampaoli<sup>1</sup>, Nágila Damaceno<sup>1</sup>, Antonio Figueiredo Neto<sup>1</sup>, Maria Camila Freitas<sup>1</sup> e Milena Barbosa<sup>1</sup>

<sup>1</sup> IME, IF e FSP/Universidade de São Paulo, SP, Brasil

**Email:** vivig@ime.usp.br

**Resumo:** A doença cardiovascular (DCV) inclui mecanismos complexos, sendo a aterosclerose a base fisiopatológica de seus eventos clínicos. Estudos recentes indicam que a lipoproteína de baixa densidade oxidada (oxLDL) é um fator de risco independente para DCV. Para mensurar o nível de oxidação dessas partículas pode ser utilizada a técnica de varredura Z ou Z-scan. Outras variáveis foram avaliadas como características bioquímicas medidas do plasma dos indivíduos, lipoproteína de baixa densidade oxidada (oxLDL), lipoproteína de baixa densidade eletronegativa [LDL(-)], glicose, apolipoproteína A-I (APOAI), apolipoproteína B (APOB), assim como o tamanho e densidade das partículas de LDL. Como o objetivo de estabelecer associações foi realizado um estudo transversal que incluiu 63 indivíduos de ambos os sexos. A técnica de Análise de Componentes Principais (PCA) foi utilizada em duas situações, a primeira considerando quinze biomarcadores cardiometabólicos e a segunda analisando os parâmetros do perfil aterogênico do LDL, com o intuito de tentar reduzir a dimensão do conjunto de variáveis relacionadas e, ao mesmo tempo, tentar encontrar uma interpretação da “função biológica” destas componentes. Após a interpretação, que estabeleceu seu padrão aterogênico ou anti-aterogênico de cada componente, por meio de um estudo de correlação vislumbrou-se a associação de cada uma delas com a medida denominada parâmetro  $\theta$ , mesurada com a técnica de Varredura-Z. Análises estatísticas futuras são necessárias para reforçar a conclusão deste trabalho, de que a técnica Z-scan parece ser uma ferramenta auxiliar e complementar, pois o parâmetro  $\theta$  parece indicar aspectos oxidativos que não são avaliados pelas outras variáveis.

---

## CP104 – Análise de componentes principais na identificação dos principais fatores relacionados à satisfação dos alunos: uma análise em uma IES Pública

Beatriz Rezzieri Marchezini<sup>1</sup>, Lucas Pereira Lopes<sup>1</sup> e Gabriel Rodrigo Gomes Pessanha<sup>1</sup>

<sup>1</sup> Universidade Federal de Alfena, MG, Brasil

**Email:** bia\_marchezini@hotmail.com

**Resumo:** O presente trabalho tem por objetivo através da Análise de Componentes Principais (ACP) analisar a satisfação dos alunos de uma Instituição de Ensino Superior (IES), buscando validar dimensões inerentes a este construto. Para isso, utilizou-se como instrumento de mensuração uma escala do tipo Likert de 10 pontos que continha 44 variáveis. A análise dos dados ocorreu com o auxílio do programa R. Pelo Alfa de Cronbach a escala utilizada apresentou um resultado de 0,91. A relevância deste artigo está na contribuição que ele proporcionará para a instituição estudada, bem como nos insights teóricos para novas incursões científicas. O modelo pode ser explicada pela extração de sete componentes principais, que juntos, explicam 61,39% da variação total dos dados, com uma prevalência dos seguintes temas: os professores demonstram interesse pelo aprendizado do estudante, relacionam a teoria com a prática, incentivam o uso dos livros disponíveis, atendem os alunos, a biblioteca possui um bom acervo de livros e periódicos e a infraestrutura do campus fornece suporte para um bom convívio. Com este resultado, a universidade conseguirá trabalhar melhor os seus recursos para o aprimoramento da qualidade de seus serviços.

## CP105 – A dependência espacial e análise discriminante canônica para a bovinocultura no estado de Mato Grosso

Camila Ozelame<sup>1</sup>, Kuang Hongyu<sup>1</sup> e Anderson Oliveira<sup>1</sup>

<sup>1</sup> Universidade Federal de Mato Grosso, MT, Brasil

**Email:** csgarioniozelame@gmail.com

**Resumo:** O trabalho procura distinguir o comportamento dos municípios matogrossenses quanto a atividades de bovinocultura utilizando métodos multivariados e espaciais, variáveis métricas e classificatórias. De acordo com a análise discriminante canônica, pode-se perceber a clara diferença entre os grupos da região sul e os grupos da região norte do estado. Com a análise de dependência espacial, constata-se que existe uma influência dos vizinhos mais próximos para a formação do comportamento bovinocultor dos municípios.

---

## CP106 – Análise do aquecimento de motores elétricos utilizando a análise de componentes principais Biplot

Cláudio Cruz da Silva<sup>1</sup>, Aline Xavier Porto Alves<sup>1</sup>, Maria Joseane Cruz da Silva<sup>1</sup> e Sônia Eliane Gonçalves dos Santos<sup>1</sup>

<sup>1</sup> Universidade Estadual de Paraíba, PB, Brasil

**Email:** clacscruz@gmail.com

**Resumo:** Na engenharia elétrica, um problema que vem preocupando os pesquisadores é o aquecimento de motores elétricos. Este problema pode causar fadiga nos materiais, interrupção na produção e consequentemente um alto custo na manutenção e processo produtivo. Para avaliar este aquecimento foi observado a temperatura de 7 motores elétricos do processo produtivo de 69 máquinas têxtil em 3 meses de avaliação (Janeiro, Fevereiro, Abril e Maio). A temperatura é considerada alta quando passa de uma determinada classe de isolamento conhecida como classe F, isto significa que a temperatura em seu bobinado pode atingir até 155°C, sem comprometer o seu desempenho e vida útil. Neste trabalho, foi utilizado a análise de componentes principais para a análise do aquecimento de diferentes motores elétricos. Com base nos resultados, em relação as máquinas, observou-se que algumas delas precisam de uma intervenção programada (Manutenção Preventiva) afim de se analisar as causas. Em relação aos motores, os resultados indicaram que eles podem estar apresentando perdas de potência em seu funcionamento. Para uma conclusão mais precisa é interessante analisar as temperaturas dos motores em relação as máquinas em cada mês de avaliação.

## CP107 – Redução de dimensionalidade e visualização: comparação entre os modelos AMMI Ponderado e AMMI Robusto

Cristiane N. Pamplona Souza<sup>1</sup> e Paulo Canas Rodrigues<sup>2</sup>

<sup>1</sup> Universidade Federal da Bahia, BA, Brasil

**Email:** crissouza.ufpa@gmail.com

**Resumo:** Com o aumento na quantidade de dados recolhidos diariamente, o desenvolvimento de novas estratégias para redução de dimensionalidade e visualização são essenciais. Uma das áreas de aplicação onde estas técnicas são mais utilizadas é a estatística genética quantitativa, onde a complexidade da interação entre genótipos e ambientes dificulta o trabalho dos melhoristas em selecionar melhores genótipos. Neste trabalho estamos interessados no estudo dos efeitos da redução de dimensionalidade utilizando o modelo AMMI - Additive Main Effects and Multiplicative Interaction Model, um dos modelos mais empregados para analisar e estruturar a interação entre genótipo e ambiente no contexto de melhoramento de plantas. Contudo, nos casos em que algumas variáveis e/ou indivíduos têm uma maior importância para a modelagem dos dados deve-lhes ser atribuído uma maior ponderação. Desta forma, é realizada uma comparação entre o modelo AMMI, uma versão ponderada do modelo definida por W-AMMI e um modelo robusto denominado R-AMMI tendo como principal objetivo diminuir a influência de outliers no modelo como também a redução do ruído na análise. Para tanto, realizou-se um estudo de simulação considerando diferentes cenários de contaminação. Como principal resultado constatou-se a discrepância no comportamento dos genótipos e ambientes quando se utiliza o modelo AMMI padrão e outras versões do modelo, isto é, em cenários com baixo percentual de contaminação o modelo AMMI foi o mais adequado, contudo, quando analisamos cenários com contaminação o modelo ponderado e robusto W-AMMI e R-AMMI permitiram reduzir a influência de observações menos precisas.

---

## CP108 – Análise de agrupamento dos municípios do estado de Sergipe em relação à precipitação e região

Eucymara Nunes<sup>1</sup> e Inajá Francisco Sousa<sup>1</sup>

<sup>1</sup> Universidade Federal de Sergipe, SE, Brasil

**Email:** eucymara@gmail.com

**Resumo:** A análise multivariada de agrupamentos foi utilizada com o objetivo de agrupar municípios do Estado de Sergipe, relacionadas a variáveis meteorológicas e coordenadas geográficas. A medida de distância Euclidiana foi utilizada juntamente ao método de agrupamento hierárquico de ligação completa. Os dados foram coletados em uma série de 30 anos, para 35 municípios sergipanos, referente às variáveis de precipitação e temperatura média, longitude, latitude, altitude e área demográfica. Mostrou-se uma boa técnica para agrupar os municípios, evidenciando os municípios mais similares por área, litoral, agreste e sertão, nos períodos chuvoso e seco.

## CP109 – Análise de fatores no agrupamento de caracteres de acessos de pimentas (*Capsicum chinense*)

Giselda M. Pereira<sup>1</sup>, Maicon Nardino<sup>1</sup>, Elisia R. Corrêa<sup>1</sup>, Willian Barros<sup>1</sup>, Ivo R. Silva<sup>2</sup>, Adilson R. Schuelter<sup>3</sup> and Fernando Finger<sup>2</sup>

<sup>1</sup> Universidade Federal de Pelotas, RS, Brasil

<sup>2</sup> Universidade Federal de Viçosa, MG, Brasil

<sup>3</sup> Universidade Estadual do Oeste do Paraná, PR, Brasil

**Email:** [ffinger@ufv.br](mailto:ffinger@ufv.br)

**Abstract:** The objective of this work was to reduce the dimensionality of original variables through factor analysis aiming to cluster traits of *Capsicum chinense* to specific factors for applying in the *Capsicum* breeding program. Were evaluated 11 variables in 41 genotypes, being 11 genitors and 30 F1 hybrids by partial diallel crosses between spicy and sweet genitors. Data were subjected to a factor analysis. Four factors had been retained explaining 81.72% of the total variation. The communality was above 70% for ten of the eleven variables. The factor one (F1) clustered variables associated with production and fruit quality (MTF, MTS, % MTS, SST, COMP and LARG). The variables vitamin C (VIT C) and number of seeds per fruit were retained in factor 2 (F2) and 3 (F3) respectively. The fourth factor (F4) has been related with capsaicinoids (CAP), adding as well the variables total fruit mass (MTF), dry total fruit mass (MTS) and flesh thickness (ESP). Factor analysis allowed the clustering of variables according to their biologic and functional concepts, furthermore acknowledged one factor associated with pungency, the most attractive trait for the peppers consumers. Factor analyses will be efficient for the Breeding Program of *C. chinense*.

## CP110 – Classificação no desfecho da doença meningocócica invasiva usando análise discriminante

Izabela R. C. Oliveira<sup>1</sup>, Daniel F. Ferreira<sup>1</sup>, Eliana T. Masuda<sup>2</sup>, Eliseu A. Waldman<sup>2</sup> e Francisco Chiaravalloti Neto<sup>2</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> FSP/Universidade de São Paulo, SP, Brasil

**Email:** [izabela.rco@gmail.com](mailto:izabela.rco@gmail.com)

**Resumo:** Frequentemente pesquisadores deparam-se com o problema de classificar um novo indivíduo em uma das  $k \geq 2$  populações, com base em um conjunto de variáveis observadas. A análise discriminante é uma ferramenta útil nessas situações e possui grande aplicabilidade em diversas áreas do conhecimento. Na medicina essa técnica pode ser aplicada, por exemplo, para uma doença com diagnóstico difícil ou com diferentes níveis de severidade. Problemas dessa natureza surgem na área de epidemiologia, em doenças graves como meningite, por exemplo, cujo diagnóstico correto e precoce aumenta as chances de recuperação do paciente. Nesse trabalho, a técnica multivariada de análise discriminante é aplicada a dados de ocorrência de meningite no município de São Paulo, no período de 1998 a 2004. Modelos de classificação para o desfecho em “óbito” e “não óbito” foram obtidos para três diferentes faixas etárias (0 a 17 anos, 18 a 39 anos e acima de 40 anos). Além da variável “severidade”, foram usadas variáveis preditoras clínicas e laboratoriais. Dentro de cada grupo etário, todas as variáveis foram consideradas em um modelo geral e, em seguida, foram obtidos modelos de classificação para variáveis clínicas e

laboratoriais, separadamente. As análises foram feitas utilizando o procedimento DISCRIM do programa SAS.

---

## CP111 – Componentes independentes, análise de agrupamento e medidas de similaridades no monitoramento da potência elétrica

Jair Rocha do Prado<sup>1</sup>, Thelma Sáfyadi<sup>2</sup> and Joaquim Paulo da Silva<sup>2</sup>

<sup>1</sup> Universidade Federal de Uberlândia, MG, Brasil

<sup>2</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** jairmat1@yahoo.com.br

**Abstract:** The power monitoring is a very important task in the various sectors of society, are companies, universities, industries, etc. Generally the analyzes in energy variables underuse or even neglect statistical methods for troubleshooting. In this work the aim is to show that statistics can be used in research in the field of electricity, and that this can contribute effectively in this business. For this study will be used electricity data from the Federal University of Lavras (UFLA). From an initial analysis of the data, there was the idea of using a new statistical approach, which considers a number of methods. The statistical methodologies were: independent component analysis (ICA), cluster analysis (AA) and indexes Moran and Geary. Comparisons were made considering the independent components (IC's), obtained from the power series of departments / sectors. It can be concluded from this work that ICA methods, AA and indexes Moran and Geary, can together be an efficient way to monitor electricity at a university or a company any.

---

## CP112 – Um estudo de diferentes técnicas de seleção do modelo AMMI

João Vitor Ribeiro Silva<sup>1</sup>, Lúcio B. de Araújo<sup>1</sup>, Mírian F. C. Araújo<sup>1</sup> and Carlos Tadeu dos Santos Dias<sup>2</sup>

<sup>1</sup> Universidade Federal de Uberlândia, MG, Brasil

<sup>2</sup> ESALQ/Universidade de São Paulo, SP, Brasil

**Email:** joao\_vitor0001@hotmail.com

**Abstract:** In this paper will be applied the additive main effects and multiplicative interaction model (AMMI) in order to analyze the interaction between genotypes and environments in multi-environments experiments. In this analysis are presented the use of two different techniques: singular value decomposition (multivariate technique based on the use of eigenvalues and eigenvectors from the interaction matrix genotype x environment) and cross-validation (technique based on the omission of an element  $x_{ij}$  of X matrix, where  $X = UDV'$ , to obtain a prediction value  $\hat{x}_{ij}$  for each choice of components and measure the discrepancy between the measured value and the original value), in order to obtain a comparison result.

## CP113 – Utilização dos modelos AMMI e GGE Biplot para os dados de interação genótipo por ambiente

Kuang Hongyu<sup>1</sup>, Fabiane Silva<sup>2</sup>, Lúcio Araújo<sup>3</sup> and Carlos Tadeu dos Santos Dias<sup>4</sup>

<sup>1</sup> Universidade Federal de Mato Grosso, MT, Brasil

<sup>2</sup> Universidade Federal do Recôncavo da Bahia, BA, Brasil

<sup>3</sup> Universidade Federal de Uberlândia, MG, Brasil

<sup>4</sup> Universidade de São Paulo, SP, Brasil

**Email:** kuang\_hongyu@hotmail.com

**Abstract:** Genotype by environment is an extremely important issue in plant breeding and production. The selection and recommendation of superior genotypes are hampered due to the constant interaction occurrence, it represents a major challenge for researchers. Thus, the multi-environment trials are essential for the selection of effective strain and recommendation of cultivars. There are several models in the literature for the analysis of multi-environmental data, models AMMI (Additive Main effects and Multiplicative Interaction) and GGE (Genotype main effects + Genotype environment interaction) biplot are the most used. The main objectives of this study with the use of AMMI and GGE biplot models were mega-environment research to the understanding of the target environment, evaluation of genotypes and environments within each mega-environment and comparison of AMMI models GGE biplot.

## CP114 – Avaliação dos registros de nascimentos das microrregiões do Semiárido Brasileiro utilizando análise fatorial

Lígia Maia<sup>1</sup>, Clarissa Cavalcanti<sup>1</sup>, Diego Silva<sup>1</sup> e Neir Paes<sup>1</sup>

<sup>1</sup> Universidade Federal da Paraíba, PB, Brasil

**Email:** clarissaocavalcanti@gmail.com

**Resumo:** Os registros dos nascimentos fazem parte de importantes indicadores demográficos epidemiológicos, mas que quando não declarados ou declarados erroneamente podem conduzir a estimação de indicadores inadequados. Esta situação é encontrada em várias regiões do Brasil, particularmente na região semiárida. Teve-se como objetivo avaliar a qualidade dos registros de nascimentos relacionados com indicadores de desenvolvimento das microrregiões do semiárido brasileiro em 2010. Os dados foram extraídos do SINASC e dos microdados do Censo Demográfico 2010. Recorreu-se à Análise Fatorial considerando os microdados de variáveis relacionadas às condições de vida dos registros de nascimentos até 10 anos de idade. Os resultados mostraram uma boa captura para a maior parte das variáveis na Declaração dos Nascimentos e a identificação de dois fatores. O primeiro foi formado por indivíduos com condições de vida mais favoráveis ao registro do nascimento e o segundo com menos condições favoráveis ao registro de nascimento. As estatísticas apontam para uma convergência e a universalização dos registros de nascimento no semiárido como um todo.

## CP115 – Aplicação do modelo multivariado de resposta ao item em um exame vestibular da UFLA

Manoel Vitor de Souza Veloso<sup>1</sup>, Caio Azevedo<sup>2</sup> e Júlio Bueno<sup>3</sup>

<sup>1</sup> Universidade Federal de Alfenas, MG, Brasil

<sup>2</sup> Universidade de Campinas, SP, Brasil

<sup>3</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** manoel.souza@unifal-mg.edu.br

**Resumo:** Esse estudo teve como principal motivação descobrir a estrutura de dependência entre habilidades nas provas de um exame de seleção, por meio do ajuste de um Modelo Multivariado de Resposta ao Item. O ajuste do modelo multivariado foi também comparado ao de modelos univariados da TRI em uma aplicação com dados do Vestibular 2006-2 da UFLA. A análise bayesiana foi implementada no programa R por meio de algoritmos do tipo Monte Carlo via Cadeias de Markov (MCMC), usando-se o amostrador de Gibbs com dados aumentados. As habilidades em cada prova foram bem estimadas com o modelo proposto, resultando em estimativas de maior correlação entre provas de mesma área (a exemplo, Física e Matemática), na capacidade de discriminar grupos de habilidades altamente associadas e gerar estimativas mais precisas.

## CP116 – Relações de interdependência da emissão de CO<sub>2</sub> e atributos de solo em áreas de cana-de-açúcar, analisadas por meio de estatística multivariada

Mara Regina Moitinho<sup>1</sup>, Daniel de Bortoli Teixeira<sup>1</sup>, Alan Rodrigo Panosso<sup>1</sup> e Newton La Scala Jr<sup>1</sup>

<sup>1</sup> Universidade Estadual Paulista, SP, Brasil

**Email:** maramoitinho@gmail.com

**Resumo:** Dentre as estratégias consideradas mais eficientes para redução das emissões de CO<sub>2</sub> do solo (FCO<sub>2</sub>) em áreas agrícolas, encontram-se a adoção de práticas de manejo mais adequadas do solo visando a maximização do sequestro de carbono. Neste estudo objetivou-se caracterizar a FCO<sub>2</sub> e sua relação com os atributos do solo em áreas sob sistemas de colheita de cana crua e queimada, e em área de reforma de canavial, utilizando-se estatística multivariada. Foram utilizadas áreas adjacentes: dois talhões apresentando os sistemas de manejo da cana-de-açúcar: cana crua (CC) e cana queimada (CQ), uma área em reforma do canavial e uma mata nativa. Ao longo de 10 dias foram avaliadas nestes sistemas, a FCO<sub>2</sub>, a temperatura e a umidade do solo e, ao final destas, foram realizadas coletas de solo para determinação dos atributos físicos e químicos do solo. As análises multivariadas de agrupamento e de componentes principais, possibilitaram a visualização das interações entre os atributos e os sistemas em estudo. Valores elevados da FCO<sub>2</sub> foram observados na área de reforma do canavial e no sistema de cana queimada, e relacionados à maior aeração (macroporos e poros livres de água) e a temperatura mais elevada do solo encontradas nestes sistemas. O sistema de colheita da cana crua quando comparado ao sistema com queima da cana propicia maior estoque de carbono no solo.

## CP117 – Determinantes da mortalidade na infância no Brasil: Uma análise por componentes principais

Marcela de Marillac Carvalho<sup>1</sup> and Manoel Vitor de Souza Veloso<sup>1</sup>

<sup>1</sup> Universidade Federal de Alfenas, MG, Brasil

**Email:** manoel.souza@unifal-mg.edu.br

**Abstract:** The paper investigates, as the main socioeconomic determinants affect the mortality rate in children in Brazil through a Principal Component Analysis. It was thus possible to affirm the importance of the performance indicators together, highlighting the expansion of public investments that enable access to important determinants, such as basic sanitation and education in the process of reducing mortality in childhood.

## CP118 – Identificação da similaridade entre tratamentos em locais distintos com a utilização da análise de múltiplos fatores

Paulo Cesar Ossani<sup>1</sup>, Tatiane Carvalho Alvarenga<sup>1</sup> e Renato Ribeiro de Lima<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** ossanipc@hotmail.com

**Resumo:** A cultura do feijão é uma das atividades agrícolas do estado de Minas Gerais, sendo que o Brasil é o maior produtor mundial. Neste trabalho, foi considerado o experimento com a cultivar BRSMG Madrepérola (do grupo carioca), dispostas em três locais: Lavras, Machado e Lambari, sendo que em cada local o delineamento experimental utilizado foi em blocos casualizados, com três repetições, sendo os tratamentos dispostos em esquema fatorial 2x3x4, constituídos pelos fatores: inoculação (sem, com), adubação de base (0, 20, 40) e adubação de cobertura (0, 30, 60 e 90 quilograma (kg) por hectare (ha) de nitrogênio usando como fonte a uréia). Na colheita foram determinados os números de plantas (NP) ou estande foral, pela contagem das plantas na área útil; o rendimento de grãos, o qual foi transformado para ; o número de vagens por planta (NV); o número de grãos por vagem (NGV) e o peso de 100 grãos (P100) em gramas. Utilizou-se a análise de múltiplos fatores para avaliação do experimento, na qual, conclui-se que houve similaridades entre os tratamentos em cada local, mostrando com isso a viabilidade da técnica análise de múltiplos fatores na identificação dos tratamentos.

## CP119 – Obtenção de limite de acurácia seletiva e coeficiente de variação para classificação de ensaios de cana de açúcar

Sandra Elisa Fonrozo<sup>1</sup>, Rubens Leite do Canto Braga Júnior<sup>2</sup> e Guilherme Augusto Peres Silva<sup>2</sup>

<sup>1</sup> ICMC/Universidade de São Paulo, SP, Brasil

<sup>2</sup> Centro de tecnologia Canavieira, SP, Brasil

**Email:** sandra.fonrozo@gmail.com

**Resumo:** Historicamente a precisão de ensaios de cana de açúcar é medida através do coeficiente de variação (CV). Contudo, o mesmo não leva em consideração algumas fontes de variação que causam o aumento da variância residual. Ficou provado em diversas culturas que ensaios conduzidos em ambientes restritivos tendem a ter coeficientes de variação mais elevados. Além disso, ensaios de avaliação de cultivares devem ser abordados do ponto de vista genético e estatístico. Assim, um dos parâmetros mais relevantes para avaliação da qualidade de um experimento é a acurácia seletiva (AS). Tem-se que os valores de CV e AS dos experimentos variam de acordo com a cultura e variável resposta em estudo. Assim, o objetivo desse projeto foi definir limites de acurácia seletiva e coeficiente de variação para classificação de experimentos de cana de açúcar para as variáveis FIBRA (porcentagem de fibra na cana), POL (porcentagem de sacarose na cana) e TCH (toneladas de cana por hectare). Os limites foram definidos a partir dos dados de 401 experimentos de cana planta e soca em blocos completos plantados nas principais regiões produtoras de cana de açúcar do Brasil e conduzidos sob a coordenação do Centro de Tecnologia Canavieira (CTC), com sede em Piracicaba, SP. Passando a classificar os ensaios através do Coeficiente de Variação e da Acurácia Seletiva, simultaneamente, ao invés de utilizar apenas a primeira estatística obteve-se uma melhor avaliação da precisão dos ensaios.

---

## CP120 – Análise de correlação canônica da emissão de CO<sub>2</sub> e atributos do solo em área de soja na região de cerrado

Vivian A. Brancaglioni<sup>1</sup>, Jacqueline Domingues<sup>1</sup>, Maria F. F. Almeida<sup>1</sup>, Alan R. Panosso<sup>1</sup>, Newton La Scala Junior<sup>1</sup> e José S. Govone<sup>1</sup>

<sup>1</sup> Universidade Estadual Paulista, SP, Brasil

**Email:** lascala@fcav.unesp.br

**Resumo:** O processo de perda de carbono do solo para a atmosfera é resultante da atividade microbiana (oxidação química) e respiração das raízes, considerada, atrás apenas dos oceanos, a segunda maior fonte de CO<sub>2</sub> para atmosfera. O objetivo do trabalho foi aplicar a técnica de análise de correlação canônica para analisar a associação entre o conjunto de variáveis ligadas às dinâmicas do carbono e o conjunto de atributos físico-químicos do solo, em área de cultivo de soja no sistema de plantio direto, no cerrado, realizado no município de Selvíria-MS. Foram consideradas nove variáveis para o estudo, sendo elas: FCO<sub>2</sub> (emissão de CO<sub>2</sub>), k (taxa de emissão de CO<sub>2</sub> por estoque) e EstC (estoque de carbono). Já o segundo conjunto é composto por: Ds (densidade do solo), Macro (macroporosidade), PT (porosidade total), P (Fósforo), SB (soma de bases) e CTC (capacidade de troca de Cátions). De acordo com os resultados obtidos, pontos com maior estoque de carbono estão associados a valores razoavelmente altos de Ds, SB e CTC, e razoavelmente baixos os valores de Macro e PT.

---

## Sessão de Pôsters 2 / Poster Session 2

---

### Big Data, Redes Neurais e Estatística Espacial

[*Big Data, Neural Networks and Spacial Statistics*]

---

#### CP121 – Aplicação de máquinas de suporte vetorial no reconhecimento de neuropatias periféricas

Albaro Paiva<sup>1</sup>, Carenne Ludeña<sup>2</sup> e Tiago Espínola<sup>1</sup>

<sup>1</sup> Universidade Federal Rural de Pernambuco, PE, Brasil

<sup>2</sup> Universidad Central de Venezuela, Venezuela

**Email:** albaropaiva@gmail.com

**Resumo:** Implementou-se uma ferramenta para a classificação de biópsias com neuropatias periféricas, do nervo sural, em imagens obtidas por óptica microscópica. Para isto foi desenvolvido algoritmos para o processamento de imagens; “thresholding”, segmentação, e extração de bordas. Definiram-se os vetores característicos dos componentes geométricos de área, perímetro e circunferência, número de Euler. Para a classificação de quatro patologias se utilizou a máquina de suporte vetorial, SVM, obtendo-se resultados aceitáveis para cada uma das patologias, destacando a entrada do número de Euler para a classificação das mesmas.

---

#### CP122 – Accuracy Assessment in Probabilistic Linkage Procedures using Big Data

André Alves<sup>1</sup>, Samila Sena<sup>1</sup>, Rosemeire Fiaccone<sup>1</sup>, Leila D. A. F. Amorim<sup>1</sup>, George Barbosa<sup>1</sup>, Marcos Barreto<sup>1</sup> and Maurício Lima Barreto<sup>1</sup>

<sup>1</sup> IM e ISC/Universidade Federal da Bahia, BA, Brasil

**Email:** rose.fiaccone@gmail.com

**Abstract:** The debate about evaluation of linkage procedures has been updated due to the increasing availability of large administrative databases. When there is a unique identifier in the multiple databases, the linkage is deterministic and trivial. In the absence of this identifier, probabilistic linkage is a widely used strategy. However, one of the biggest challenges of those approaches is about assessing the quality of the linkage process because a gold standard is usually unavailable. Thus, this study aims to evaluate the accuracy of a probabilistic linkage procedure using big data from administrative systems of the Brazilian federal government through definition of appropriate sensitivity and specificity measures, and positive predictive values. Manual and automatic search were performed to establish the gold standard and to define true matches and non matches. The quality of the probabilistic linkage procedures using big data

has to be carefully addressed because it has important consequences to data analysis.

---

## CP123 – Estudo exploratório da configuração espacial pontual dos alvéolos do pão francês

Carolina Cristina Bicalho<sup>1</sup>, Raul Antônio Viana Madeira<sup>1</sup> e João Domingos Scalon<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** carolinabicalho@gmail.com

**Resumo:** A textura do pão depende, entre outras variáveis, da quantidade, distribuição e tamanho dos alvéolos (bolhas de ar) na massa. Assim, essas variáveis devem ser analisadas durante o processo de fabricação para que o pão atinja uma qualidade adequada para o consumo. Métodos estatísticos convencionais (não espaciais) começam a ser utilizados para a análise dessas variáveis. Observa-se que métodos da estatística espacial ainda não foram utilizados com este propósito. Assim, o objetivo deste trabalho foi analisar a distribuição espacial dos alvéolos localizados na massa congelada de pão francês através da aplicação de métodos exploratórios de processos pontuais univariados não marcados no plano. As coordenadas da posição dos alvéolos foram obtidos por processamento de imagem de uma fatia transversal de pão francês com uma área de, aproximadamente, 16,32 cm<sup>2</sup> que apresentou 855 alvéolos. Utilizou-se métodos para explorar as propriedades de primeira ordem (estimação do kernel) e segunda ordem (funções F, G, J e K). Os resultados mostraram que o número de alvéolos por unidade de área de massa é, aproximadamente, constante e que os alvéolos tendem a se localizarem mais afastados uns dos outros do que seria esperado em uma distribuição de alvéolos completamente aleatória. Os métodos utilizados possibilitaram caracterizar a distribuição espacial dos alvéolos no pão francês e, conseqüentemente, apontar que o pão analisado apresenta qualidade adequada para o consumo.

---

## CP124 – Análise Espacial dos Resultados do ENEM 2014 das Escolas do Município de Fortaleza/CE

Débora Ferreira de Assis<sup>1</sup> e Júlio Francisco Barros Neto<sup>2</sup>

<sup>1</sup> PET Estatística/Universidade Federal do Ceará, CE, Brasil

<sup>2</sup> DEMA/ Universidade Federal do Ceará, CE, Brasil

**Email:** jfbarros@dema.ufc.br

**Resumo:** O Exame Nacional do Ensino Médio (ENEM) é um prova para avaliar os alunos do ensino médio, a partir do qual o estudante tem a oportunidade de ingressar no ensino superior em instituições públicas, também pode obter bolsa de estudo em instituições particulares ou ainda para obter certificação de ensino médio. Esse trabalho tem por objetivo buscar características que influenciem nas notas dos alunos que participaram desse exame, avaliando diferentes indicadores e identificando a existência ou não de dependência espacial das notas. A análise foi realizada a partir de dados do INEP/MEC, sendo avaliados os resultados do município de Fortaleza/CE onde foram registradas 276 escolas de ensino médio com alunos que realizaram o ENEM em 2014 e foram utilizados os softwares R e ArcGIS. Dentre os resultados, pôde-se concluir que as escolas particulares apresentaram notas superiores às escolas públicas, bem como, nos bairros que possuíam os melhores indicadores de Índice de Desenvolvimento Humano -

Educação as melhores notas foram observadas. Também, verificou-se que as escolas que estavam caracterizadas como pertencentes entre as melhores do país estas possuíam baixo fator de permanência dos alunos nas escolas.

---

## CP125 – Delineamento de unidades de mapeamento com base na incerteza espacial da suscetibilidade magnética e teor de argila

Gener Tadeu Pereira<sup>1</sup>, Daniel de Bortoli Teixeira<sup>1</sup>, Diego Silva Siqueira<sup>1</sup>, Vinícius Vasconcellos<sup>1</sup> e José Marques Júnior<sup>2</sup>

<sup>1</sup> FCAV/Univesidade Estadual de São Paulo, SP, Brasil

<sup>2</sup> DG/Universidade de Brasília, DF, Brasil

**Email:** genertp@fcav.unesp.br

**Resumo:** O objetivo deste estudo foi utilizar as incertezas espaciais da suscetibilidade magnética (SM) e teor de argila para o aprimoramento de unidades de mapeamento do solo. Na área experimental, abrangendo 870 ha, foi instalada uma malha de amostragem regular contendo 371 pontos, onde se avaliou a SM e teor de argila na profundidade de 0-0,25 m. Com o auxílio do modelo digital de elevação e observações de campo, foi identificada uma transeção na área em estudo de 4440 m a partir do topo da vertente até o sopé, no sentido do decaimento mais suave do declive. As incertezas nas estimativas espaciais foram contabilizadas pelos mapas de desvio padrão estimados com base em 200 realizações da simulação sequencial gaussiana. Para a definição das zonas de transição na transeção, foram selecionadas as isolinhas das incertezas representativas da transição entre as unidades de mapeamento do solo. Ambos os atributos apresentaram valores de pico das incertezas próximos à mudança de unidades de mapeamento previamente conhecidas. As zonas de incerteza, determinadas na transeção previamente delineada, variaram de 45 a 210 m e de 60 a 170 m, para a SM e teor de argila, respectivamente. Entretanto, após a extrapolação das incertezas para a lateral da transeção, verificaram-se mudanças na amplitude das zonas de incerteza. O delineamento de unidades de mapeamento incorporando as incertezas da SM e teor de argila foi semelhante. Entretanto, devido ao menor custo e rapidez, a SM passa a ser a alternativa mais viável.

---

## CP126 – Redes neurais aplicadas à classificação de imagens radiográficas de sementes

Isabel Leite<sup>1</sup>, Felipe Gonzalez<sup>1</sup>, Thelma Sáfyadi<sup>2</sup> e Maria Laene Carvalho<sup>2</sup>

<sup>1</sup> Instituto Federal de Educação, Ciência e Tecnologia da Bahia, BA, Brasil

<sup>2</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** iccleite@gmail.com

**Resumo:** As Redes Neurais Artificiais (RNA) são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura do cérebro humano para tentar simular sua forma inteligente de processar informações e adquirir conhecimento através da experiência. São amplamente utilizadas no reconhecimento de padrões. O presente estudo teve como objetivo implementar uma RNA para a classificação de imagens radiográficas de sementes de girassol (*Helianthus annuus* L.) segundo a sua morfologia interna, separando-as em três classes: sementes cheias, sementes deformadas ou com danos graves e se-

mentes com danos leves. O método foi aplicado com o algoritmo de aprendizado Backpropagation a partir de um conjunto de 445 imagens previamente processadas pelas técnicas estatísticas análise de componentes principais (PCA - Principal Component Analysis) e análise de componentes independentes (ICA - Independent Component Analysis), usadas para redução da dimensão do conjunto original de dados. Buscou-se qual entre os conjuntos (RNA com ICA e RNA com PCA) apresentava melhores resultados na identificação das referidas classes de sementes. O trabalho foi desenvolvido no ambiente computacional do software Matlab com a utilização de toolbox específica para implementação de RNA e, dentre as diversas arquiteturas de redes testadas, obteve-se como melhor resultado a estrutura de RNA com uma camada oculta com 35 neurônios e dados provenientes da decomposição pela técnica ICA. O desempenho atingido nesta configuração apresentou globalmente 87% de acertos na classificação das imagens, demonstrando que a aplicação do método pode contribuir para uma avaliação rápida e menos subjetiva de imagens radiográficas de sementes.

---

## CP127 – Análise da autocorrelação espacial dos setores censitários de Rio Claro - SP para os casos de dengue, através do índice de Moran

Jacqueline Domingues<sup>1</sup>, Maria F. F. Almeida<sup>1</sup>, Vivian Ap. Brancaglioni<sup>1</sup> e José S. Govone<sup>1</sup>

<sup>1</sup> Universidade Estadual Paulista, SP, Brasil  
**Email:** vivian.brancaglioni@gmail.com

**Resumo:** No presente trabalho buscou-se identificar através da Análise Exploratória de Dados Espaciais, clusters (aglomerações) dos casos notificados de dengue no município de Rio Claro - São Paulo por setor censitário. Desta forma, através de mapas e análise de autocorrelação espacial, calculada pelo índice global e local de Moran, verificou-se existência de autocorrelação espacial para a variável de interesse entre os setores censitários.

---

## CP128 – The process of search networks applied to the MDM

Lilia Carolina Costa<sup>1</sup>

<sup>1</sup> Universidade Federal da Bahia, BA, Brasil  
**Email:** liliacarol@gmail.com

**Abstract:** A Multiregression Dynamic Model (MDM) is a class of multivariate time series that represents multiple dynamic causal processes in a graphical way. One of the advantages of this class is that, in contrast to many other Dynamic Bayesian Networks, the hypothesised relationships accommodate conditional conjugate inference. We demonstrate for the first time how it is straightforward to search over all possible connectivity networks with dynamically changing intensity of transmission to find the Maximum a Posteriori Probability (MAP) model within this class. This search method is made feasible by using a novel application of the integer programming algorithm. The search over all possible directed (acyclic or cyclic) graphical structures can be made especially fast by utilising the fact that, within this class of models, the joint likelihood factorizes. We then demonstrate that the MDM-IPA is not only useful method for detecting the existence of edges, but also for estimating its direction. We apply this method

to fMRI datasets for estimating brain connectivities.

---

## CP129 – Distribuição espacial dos focos de queimadas no estado de São Paulo no período de 2008 a 2015

Mayara P.L.S. Sicchieri<sup>1</sup> e Estela C. Carneseca<sup>1</sup>

<sup>1</sup> ProEstat - Consultoria Estatística e Pesquisa de Mercado, SP, Brasil

**Email:** mayapiani@hotmail.com

**Resumo:** É comum o uso do fogo nas áreas urbanas de todo o Brasil, no entanto nas áreas rurais é onde as queimadas acontecem com maior frequência. Em particular, a queima da palha da cana-de-açúcar é uma prática comum em grande parte do Estado de São Paulo. Diante das várias pesquisas envolvendo as queimadas como fator de risco à saúde, algumas medidas para eliminar a queima foram propostas pelo governo do Estado. A mais recente, firmada no protocolo agroambiental de 2007, propõe como o prazo final para a erradicação total das queimadas o ano de 2017. Com o prazo final se aproximando, existe a necessidade de monitorar os focos de queimadas bem como entender sua evolução ao longo dos anos. Este trabalho mostra a distribuição espacial dos focos de queimadas entre os anos de 2008 e 2015 e demonstra como esta distribuição evoluiu após o acordo firmado em 2007. O arquivo contendo a malha digital do estado de São Paulo foi obtido no site do IPEAGeo e o programa QGIS foi utilizado para elaborar os mapas. Foram observadas grandes melhoras na região sudoeste do Estado, enquanto que a principal macrorregião canavieira, Ribeirão Preto, mostrou que mais da metade dos seus municípios se mantiveram com os piores índices de queimadas ao longo dos três períodos.

---

## CP130 – Funções L homogênea e não homogênea em testes de hipótese de completa aleatoriedade espacial

Sidcleide Barbosa de Sousa<sup>1</sup>, João Domingos Scalon<sup>1</sup> e Kelly Pereira de Lima<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** sbarbosas1987@yahoo.com.br

**Resumo:** O objetivo principal deste trabalho é investigar o poder de testes contra a hipótese de completa aleatoriedade espacial em configurações pontuais simuladas que apresentam não homogeneidade e dependência. Os testes utilizam as funções L homogênea e não homogênea na estatística de Kolmogorov-Smirnov. Os resultados mostram que a função L homogênea não deveria ser usada em configurações suspeitas de serem não homogêneas e que a função L não homogênea não deveria ser usada em configurações homogêneas.

## CP131 – Análise da distribuição espacial da resistência do solo à penetração em área experimental

Wilhan Valasco dos Santos<sup>1</sup>, Divino de Sousa Costa<sup>1</sup>, Brunna Rithielly Rezende<sup>1</sup>, Hygor Amaral Santana<sup>1</sup> e Anderson Rodrigo da Silva<sup>1</sup>

<sup>1</sup> Instituto Federal Goiano-Câmpus Ururutaí, GO, Brasil

**Email:** wilhanvalasco@hotmail.com

**Resumo:** A compactação do solo limita o crescimento radicular das plantas, podendo causar perdas na produção agrícola. A resistência à penetração é uma medida frequentemente empregada para avaliar o grau de compactação do solo. O objetivo deste trabalho foi avaliar a variabilidade espacial da resistência mecânica a penetração de uma área experimental onde havia cultivo de pastagem, empregando metodologias distintas, quais sejam: teste e correlograma de Mantel e o ajuste de semivariogramas geoestatísticos. Foram coletados 181 pontos num grid amostral de dimensão 50 x 70 m, de forma semi-aleatória, isto é, com distâncias entre pontos vizinhos variando entre 3 m, 5 m e 7 m, nas camadas de 0-20 cm e 20-40 cm de profundidade. Observou-se que a área é considerada compactada para a maioria das culturas de interesse econômico. O correlograma de Mantel indicou resultado discordante do semivariograma ajustado para a camada superficial, devido a variabilidade em termos de resistência a penetração da mesma.

---

## Sessão de Pôsters 2 / Poster Session 2

---

### Geoestatística e estatística não paramétrica [*Geostatistics and non-parametric statistics*]

---

#### CP132 – Applied Nonparametric Analysis of covariance: One study of the effect of the bandwidth choosing

Fabio da Silva<sup>1</sup>

<sup>1</sup> Centro Federal Tecnológica de Minas Gerais, MG, Brasil

**Email:** fabiorochadasilva@yahoo.com.br

**Abstract:** In this work, we compare the relationship between homicide rate and the size of the population through the time in several places. For each place (group) we estimate that relationship using polynomial local regression. The optimal bandwidths are estimated using Fan and Gijbel's method and to test the equality of the regression functions we used Dette and Neumeyer's method. We also compare the performance of the bandwidths estimated by Fan and Gijbels's method with those used by Dette and Neumeyer.

---

#### CP133 – Influence of symmetry of the data in a geostatistical analysis

Lidiane Rosa<sup>1</sup>, Gerson Santos<sup>1</sup>, Adriana Santos<sup>1</sup> and Italo Ferreira<sup>1</sup>

<sup>1</sup> Universidade Federal de Viçosa, MG, Brasil

**Email:** prof.gersonrodrigues@gmail.com

**Abstract:** Geostatistics is used to study the continuity of a phenomenon. It is characterized by using all the available information, including the geographical position of the sample data, to characterize the variable under study. The aim of this study was to assess the commitment of a geostatistical analysis between asymmetric data and symmetric data adopting three different data transformations, starting from the natural asymmetry of data arriving in a statistically acceptable symmetry. It was found that the benefits achieved in geostatistical analysis were not significant in the face of obtained disadvantages, among these, there is the significant increase in the variance of kriging, which was 0.01772606 to the original data and came to 51.8138 for the best transformation.

## CP134 – Estudo da variabilidade da emissão de CO<sub>2</sub> do solo em área de reforma do canavial, utilizando-se análises multivariada e geoestatística

Mara Regina Moitinho<sup>1</sup>, Antonio Sergio Ferraudo<sup>1</sup>, Daniel de Bortoli Teixeira<sup>1</sup>, Fabiana De Souza Cannavan<sup>2</sup>, Siu Mui Tsai<sup>2</sup>, Elton da Silva Bicalho<sup>1</sup> e Newton La Scala Junior<sup>1</sup>

<sup>1</sup> FCAV/Universidade Estadual Paulista, SP, Brasil

<sup>2</sup> CENA/Universidade Estadual Paulista, SP, Brasil

**Email:** maramoitinho@gmail.com

**Resumo:** Devida a grande variabilidade da emissão de CO<sub>2</sub> do solo (FCO<sub>2</sub>), o emprego de ferramentas como as análises multivariada e a geoestatística, fornece informações relevantes, auxiliando na compreensão da dinâmica do CO<sub>2</sub> entre o solo e a atmosfera. Assim, objetivou-se caracterizar a FCO<sub>2</sub> associada aos atributos físicos, químicos e biológicos do solo em área de reforma do canavial, por meio das estatísticas multivariada e geoestatística. O estudo foi conduzido em área agrícola localizada no Estado de São Paulo, Brasil. Foram conduzidas 10 avaliações da FCO<sub>2</sub>, temperatura e umidade do solo, num gradeado regular de 90 x 90 m contendo 100 pontos. Posteriormente, em cada ponto foram coletadas amostras de solo na profundidade de 0-0,20 m para determinação dos atributos do solo. Foram realizadas análises multivariadas para seleção das variáveis relacionadas a FCO<sub>2</sub> e, geoestatísticas para a caracterização da variação espacial e mapeamento dos atributos. Na análise de fatores foram identificados dois processos (fatores), sendo o primeiro fator o que reteve os atributos responsáveis pelos processos de produção e transporte de CO<sub>2</sub> no solo. No mapeamento da FCO<sub>2</sub>, foram identificadas duas regiões na área em estudo com diferentes potenciais de emissão para a quantificação do número de cópias dos genes 16S rRNA e pmoA, por meio das análises de PCR quantitativo em tempo real (qPCR). Os padrões espaciais da FCO<sub>2</sub> foram similares aos atributos: macroporos, porosidade livre de água e matéria orgânica. Os resultados da qPCR para os genes 16S e pmoA foram relacionados aos resultados observados para a FCO<sub>2</sub>.

## CP135 – Comparação entre modelo paramétrico e não paramétrico para dados de contagem

Paulo Henrique Sales Guimarães<sup>1</sup> e Sue Éllen Ester Queiroz<sup>2</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas, MG, Brasil

**Email:** phsg13@yahoo.com.br

**Resumo:** A distribuição de Poisson é comumente utilizada para o ajuste de dados de contagens. Assim, para dados que apresentam esta característica, modelos de regressão linear não devem ser aplicados a estes, sendo necessário o uso de modelos lineares generalizados (MLG). A aproximação não paramétrica é mais flexível, pois determina uma função  $f$  a partir de uma família de funções, sendo que um caso especial importante do modelo geral é regressão não paramétrica simples, onde existe apenas um preditor:  $x_i$ . Neste trabalho buscou-se fazer ajustes paramétricos e não paramétricos ao conjunto de dados trabalhados. Os resultados mostraram uma relativa superioridade do GML em detrimento do método de kernel.

Entretanto, o último revelou ser bastante flexível e coeso para ajustar os dados do experimento em estudo.

---

## Sessão de Pôsters 2 / Poster Session 2

---

### Séries Temporais, Cópulas e Econometria

[*Time series, Copulas and Econometrics*]

---

#### CP136 – Comparação de métodos de estimação do risco relativo

Agda Galletti<sup>1</sup>, Mateus Ananias<sup>1</sup>, Rodrigo Lassance<sup>1</sup>, Bernardo Andrade<sup>1</sup> e Joanlise Marco de Leon Andrade<sup>1</sup>

<sup>1</sup> Universidade de Brasília, DF, Brasil

**Email:** joanlise@unb.br

**Resumo:** O risco relativo (RR) é uma medida de associação bastante utilizada em estudos epidemiológicos com delineamentos transversal e de coorte. Vários métodos de estimação do RR tem sido propostos. Entre eles cita-se modelos de regressão log binomial (RLB) e as aproximações por modelos logístico, de Poisson e de Cox. O modelo RLB, estima diretamente o RR, porém requer um algoritmo apropriado para o cálculo numérico do estimador de máxima verossimilhança (EMV), o que não é o padrão nas rotinas utilizadas em software estatísticos e pode acarreta em falhas de convergência em determinadas situações. Neste trabalho avaliou-se dois métodos de estimação do RR por meio de simulações em diferentes processos de geração de dados. Comparou-se um método indireto de estimação do RR via regressão logística (proposto por Santos et. al., 2008) com um método de estimação direta que utiliza RLB impondo-se restrições ao espaço paramétrico para se contornar problemas de convergência (proposto em Andrade e Carabin, 2006). Observou-se que o método direto forneceu melhores resultados com viés muito menor e maior precisão e cobertura em simulações com amostras de tamanho 500.

---

#### CP137 – Análise de previsibilidade da velocidade do vento no Nordeste Brasileiro

Antonio Silva<sup>1</sup>, Leda Santana<sup>1</sup>, Luiz Dore<sup>1</sup> e Tatijana Stosic<sup>1</sup>

<sup>1</sup> Universidade Federal de Pernambuco, PE, Brasil

**Email:** valeria\_araua@hotmail.com

**Resumo:** Foram analisadas séries diárias de velocidade do vento em 72 pontos de grade distribuídos sobre a região Nordeste (NE) do Brasil disponíveis pelo European Centre for Medium-Range Weather Forecasts (ECMWF). Em cada ponto de grade foi obtido a média e o desvio padrão, em seguida calculada a entropia de permutação normalizada (PE Norm). Os resultados para cada ponto de grade foram apresentado de forma espacialmente explícita através da técnica do Inverse distance Weighing (IDW) com peso  $a = 2$ . Percebeu-se que as sub-regiões Agreste, Zona da Mata e parte do Sertão do NE apresentam as maiores médias, já o desvio padrão é mais acentuado na região norte do NE e menos expressivo no Maranhão.

A distribuição da dinâmica de entropia no nordeste manteve um padrão bem definido, independente do tamanho do vetor ( $D=3,4,5$  e  $6$ ), contudo é visto que a entropia diminui a medida que aumenta-se o valor de  $D$ . Os menores valores de entropia são observados na Bahia, Sertão de Pernambuco e parte do Ceará e Piauí, enquanto a região Oeste tem comportamento oposto, altas entropia. Ainda é importante mencionar a formação de contornos concêntricos, ressaltando altos/baixos valores de entropia, característica peculiar do IDW. O resultados de entropia obtido para a região NE podem ser justificados pelas características do relevo da região, da presença dos ventos alísios de Leste e as brisas marítimas que influência na velocidade do vento.

---

## CP138 – Análise da produção de biodiesel no Brasil usando a metodologia de Box e Jenkins

Bárbara Alves Lopes<sup>1</sup> e Leticia Lima Milani<sup>1</sup>

<sup>1</sup> Universidade Federal de Alfenas, MG, Brasil

**Email:** camilam.ribeiro@yahoo.com.br

**Resumo:** O biodiesel é um dos tipos existentes de biocombustível, utilizado para substituir o uso de diesel, sua matéria prima deriva-se de diversos tipos de óleo animal ou vegetal, este tipo de mistura vem sendo utilizada em uma escala crescente, isso se deve ao fato de sua composição reduzir de forma considerável o risco de poluição, causada pelo petróleo que é um combustível não renovável (Santos e Pinto, 2009). No Brasil este tipo de biocombustível foi legalmente aceita em 2008 (Kohlhepp, 2010). Este artigo tem como objetivo fazer uma análise temporal sobre as observações feitas da produção do biodiesel no Brasil, com as informações colhidas mensalmente entre janeiro de 2008 até junho de 2015, segundo a Agência Nacional de Petróleo, fazendo a previsão e comparando os dados da série original e a prevista no período de julho de 2015 até dezembro de 2015. Para isso inicialmente foram feitos os testes de Cox-Stuart que analisa tendência (Morettin e Toloi, 2004) e o teste de Fisher que verifica sazonalidade (Priestley, 1989), inicialmente foi analisado a presença de tendência e sazonalidade na série que ficou estacionária após a diferenciação, para testar foi feita a autocorrelação. Em seguida foi realizado o teste de resíduo e observando o Critério de Akaike que considera seu menor valor como sendo o melhor (Alves, Costa Leite e Machado, 2008), foi considerado o melhor modelo o ARIMA (1,1,1), seguindo a metodologia Box-Jenkins. a previsão do modelo foi superestima 50% dos valores e subestima os outros 50%.

---

## CP139 – Análise econométrica da estabilidade do mercado acionário do setor bancário Brasileiro

Beatriz de Souza Bernadinho<sup>1</sup>, Gabriel Rodrigo Gomes Pessanha<sup>1</sup>, Caio Tonolli Dias Gentile<sup>1</sup> e Lucas Pereira Lopes<sup>1</sup>

<sup>1</sup> Universidade Federal de Alfenas, MG, Brasil

**Email:** caiodiasgentile@gmail.com

**Resumo:** Este trabalho tem como objetivo identificar os determinantes da estabilidade do mercado do setor financeiro brasileiro. Particularmente, pretende-se detectar os efeitos das fusões e aquisições, dos marcos regulatórios e de variáveis macroeconômicas (PIB, Taxa de juros, Taxa de câmbio, Inflação) na

estabilidade do mercado acionário. Os dados necessários para realização deste trabalho foram obtidos a partir de fontes secundárias, no período compreendido entre o primeiro trimestre de 1995 e o segundo trimestre de 2015. Foram formuladas 7 hipóteses a serem testadas pela metodologia empírica adotada no desenvolvimento da pesquisa. A análise das variáveis utilizadas no trabalho foi realizada por meio da técnica de regressão linear múltipla (RLM). Sendo assim, conclui-se que as variáveis de mercado (sistema regulatório, estratégias de F&A e concentração de mercado) geram, em média, elevações na variabilidade dos retornos afetando negativamente a estabilidade do mercado acionário, em outras palavras, causam um incremento no risco do investidor. Semelhante comportamento observou-se com relação às variáveis referentes aos aspectos econômicos (crescimento econômico – PIB, inflação e câmbio). Assim, destaca-se a importância das variáveis macroeconômicas no comportamento do investidor e na formação do risco sistemático. Também constatou-se que as fusões e aquisições exercem aumento na instabilidade do mercado acionário, ressaltando a importância do estudo das operações. Portanto, os resultados encontrados respondem os objetivos traçados, já que constatou-se que a variável dependente em estudo (estabilidade de mercado) está estreitamente relacionada com as variáveis independentes incluídas no modelo, reforçando a importância de se avaliar e estudar as variáveis relativas ao mercado.

---

## CP140 – Moving block quantile residual Bootstrap in GARMA models: An application for the time series of dengue case count

Breno Silveira de Andrade<sup>1</sup>, Jacek Léskow<sup>1</sup> and Marinho Gomes de Andrade Filho<sup>2</sup>

<sup>1</sup> ICMC/Universidade de São Paulo, SP, Brasil

<sup>2</sup> Institute of Mathematics, Cracow Technological University, Poland

**Email:** marinho@icmc.usp.br

**Abstract:** Generalized autoregressive moving average (GARMA) models were developed to extend the univariate Gaussian ARMA time series. In particular, such models can be applied to describe discrete time series. Our paper introduces a resampling method called Moving Block Quantile Residual Bootstrap (MBQRB) to improve inference for the parameters of the GARMA model. We provide consistency theorem for MBQRB as well as simulated and real data example related to hospitalizations caused by Dengue disease in the state of Paraíba in Brazil.

---

## CP141 – Determinantes da concentração de mercado: Uma análise econométrica do setor bancário Brasileiro no período de 1995 a 2015

Caio Tonolli Dias Gentile<sup>1</sup>, Gabriel Rodrigo Gomes Pessanha<sup>1</sup>, Lucas Pereira Lopes<sup>1</sup> e Beatriz de Souza Bernadinho<sup>1</sup>

<sup>1</sup> Universidade Federal de Alfenas, MG, Brasil

**Email:** caiodiasgentile@gmail.com

**Resumo:** O objetivo deste trabalho é analisar o impacto das fusões e aquisições, dos marcos regulatórios e de variáveis macroeconômicas (PIB, Taxa de juros, Taxa de câmbio, Inflação) na concentração de mercado do setor bancário brasileiro no período compreendido entre o primeiro trimestre de 1995 e o segundo trimestre de 2015. A análise das variáveis utilizadas no trabalho foi realizada por meio da técnica

de regressão linear múltipla (RLM) com o auxílio do software Gretl 1.9.90. A primeira etapa do trabalho se refere à verificação da aderência dos dados e modelos ajustados aos pressupostos da RLM. Uma vez testados os pressupostos, procedeu-se a análise dos dados obtidos por meio do ajuste dos modelos. De maneira geral, pode-se afirmar que a concentração de mercado é afetada, embora com valores pouco expressivos para o coeficiente de determinação ( $R^2$ ), pelas demais variáveis incluídas nos modelos. É possível inferir que as variáveis de mercado (sistema regulatório e estratégias de F&A) geraram, em média, maior poder de mercado para os bancos à medida que forneceram uma relação direta com a concentração de mercado. Ademais, pode-se concluir que situações macroeconômicas que demonstram instabilidade, ou seja, com alta taxa de juros, inflação e câmbio e baixo crescimento econômico podem enriquecer a concentração e o poder de mercado das instituições. Nestas circunstâncias de instabilidade econômica, as oportunidades de F&A podem ser constatadas e aproveitadas pelas instituições adquirentes.

---

## CP142 – Ajuste de modelo de séries temporais para o número de óbitos por causas externas no Brasil

Camila Moreira Ribeiro<sup>1</sup> e Leticia Lima Milani<sup>1</sup>

<sup>1</sup> Universidade Federal de Alfenas, MG, Brasil

**Email:** rodrigues.milani.l@gmail.com

**Resumo:** Entende-se por óbito por causa externa aquilo que decorre de lesão provocada por violência (homicídio, suicídio, acidente ou morte suspeita), independente do tempo entre o evento lesivo e a morte. O índice de mortalidade por causas externas no Brasil vem apresentando significativo crescimento ao longo das décadas. Constata-se que o crescimento dos óbitos está ocorrendo, principalmente entre os adultos jovens, especialmente do sexo masculino. Pela frequência com que ocorrem e por serem estes grupos mais atingidos, os acidentes de trânsito, os homicídios e os suicídios respondem por cerca de dois terços dos óbitos no país. O estudo da mortalidade pode trazer contribuições importantes para o país que permitam orientar as discussões quanto a medidas preventivas para esta epidemia característica da modernidade. Desta forma, este trabalho tem o objetivo de apresentar uma análise temporal da série de óbitos por causas externas no Brasil, no período de janeiro de 2000 a dezembro de 2013, a partir do modelo ajustado por meio da metodologia SARIMA.

---

## CP143 – Uso de séries temporais na análise do consumo de energia elétrica da Universidade Federal de São João Del-Rei

Caroline Oliveira Santos<sup>1</sup>, Letícia Lima Milani<sup>2</sup>, Andréa Cristiane Delfino<sup>3</sup> e Ana Lúcia Mateus<sup>4</sup>

<sup>1</sup> Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, RJ, Brasil

<sup>2</sup> Universidade Federal de Alfenas, MG, Brasil

<sup>3</sup> Universidade Federal de São João del-Rei, MG, Brasil

<sup>4</sup> Universidade Federal de Santa Maria, RS, Brasil

**Email:** carolineoliviera@yahoo.com.br

**Resumo:** Diversas áreas do conhecimento utilizam a modelagem via séries temporais para estudar o comportamento de um conjunto de dados observado ao longo do tempo. Neste trabalho pretende-se utilizar a metodologia de Box e Jenkins (1976), para modelar a série do consumo de energia elétrica (kWh) do Campus Santo Antônio da Universidade Federal de São João del-Rei (UFSJ), no período de Janeiro de 2007 a Dezembro de 2014. O critério de informação de Akaike foi utilizado para a seleção do melhor modelo dentre os modelos possíveis. A análise inicial da série deu-se pela elaboração do gráfico temporal da série e do gráfico da função de autocorrelação. Esta análise preliminar apontou a presença de tendência crescente, mas não foi possível confirmar a presença de sazonalidade na série. Ainda para averiguação dos componentes de tendência e sazonalidade, aplicaram-se os testes de Cox-Stuart e Fisher, respectivamente. Estes testes confirmaram a existência das duas componentes na série original. Eliminou-se a tendência e a sazonalidade com uma diferença simples e uma diferença sazonal. Desta forma, a estacionariedade da série pôde ser confirmada pela verificação do gráfico da função de autocorrelação. Para a série sem tendência e dessazonalizada o modelo selecionado de acordo com o critério de Akaike foi o SARIMA(1,1,1)(0,1,1). De acordo com os resultados deste estudo, pode-se concluir que o modelo selecionado ofereceu valores previstos dentro dos limites aceitáveis, sendo considerados como boas previsões.

---

## CP144 – Análise do volume do reservatório da usina hidrelétrica de Três Marias usando modelos de séries temporais

Charles Shalimar Felipe da Silva<sup>1</sup>, Paulo César Moraes Ribeiro<sup>1</sup>, Sidcleide Barbosa de Sousa<sup>1</sup>, Ricardo Vitor Ribeiro dos Santos<sup>1</sup> e Marcelo Silva de Oliveira<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** charlesfsilva@yahoo.com.br

**Resumo:** A aplicação da metodologia de Box e Jenkins de séries temporais na análise do volume útil do reservatório da usina hidrelétrica de Três Marias conseguiu descrever adequadamente o comportamento da série em estudo, por meio de um modelo  $SARIMA(1,0,1)(3,1,2)_{12}$  visto que este reproduziu, de modo eficiente, a sazonalidade presente nas observações. O modelo prevê que, para os próximos meses, o volume continuará oscilando em torno de valores percentuais abaixo de 25 por cento do total.

---

## CP145 – Um teste para qualidade de ajuste de Cópulas bivariadas via marginal obtida na diagonal

Danielle Prado<sup>1</sup>, Lucas Chaves<sup>2</sup> e Devanil Souza<sup>2</sup>

<sup>1</sup> Universidade Tecnológica Federal do Paraná, PR, Brasil

<sup>2</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** danigoprado@hotmail.com

**Resumo:** É notório o grande número de estudos recentes relacionados à qualidade de ajuste para cópulas e inegável a sua relevância em diversas áreas, principalmente na economia. Eleger uma família de cópulas para ajustar um determinado conjunto de dados é uma tarefa importante e bem complexa. Porém, ainda não existe um método dito mais adequado para tal. Nos últimos anos, vários métodos têm sido propostos observando-se os diferentes tipos de características dos dados. O objetivo principal deste trabalho é propor

um teste para verificar a qualidade de ajuste de cópulas para dados bivariados via marginais nas diagonais principal e secundária baseado no teste qui quadrado para distribuições de frequências. A fim de verificar sua adequabilidade, foram utilizadas as seguintes famílias de cópulas: Clayton, Gumbel, Normal e Frank. Os cálculos utilizados nos dois testes foram feitos com o auxílio do software livre R. A conclusão foi feita através da verificação do controle das taxas de erros tipo I e tipo II.

---

## CP146 – Determinantes do retorno das ações: Um estudo econométrico na Eletrobrás no período de 1994 a 2014

Ericksen Souza<sup>1</sup>, Gabriel Rodrigo Gomes Pessanha<sup>1</sup> e Luciene Resende Gonçalves<sup>1</sup>

<sup>1</sup> Universidade Federal de Alfenas, MG, Brasil

**Email:** ericksen\_wcs@hotmail.com

**Resumo:** Este trabalho, a partir de uma análise quantitativa, tem como objetivo principal evidenciar a relação entre o retorno das ações da estatal brasileira Eletrobras e variáveis macroeconômicas, bem como quantificar o impacto que essas variáveis têm sobre o retorno das ações, no período de agosto de 1994 e vai até julho de 2014. Para a realização desta pesquisa foi usada regressão linear múltipla, com auxílio do teste de correlação e o teste de Variance Inflation Factors (VIF) para a análise de colinearidade. Os modelos de regressão linear múltipla evidenciaram que algumas das variáveis apresentavam p-valor muito elevado, assim algumas variáveis foram excluídas para que o modelo apresentasse um nível de confiança elevado. Contudo, pode ser observado a existência de variáveis influentes na tentativa de explicar as variações do retornos das ações da Eletrobras.

---

## CP147 – Estudo do Índice de Criminalidade contra a Vida (ICV) em Maputo/Moçambique usando séries temporais

Joel Jorge Nuvunga<sup>1</sup>, Luciano Antonio de Oliveira<sup>2</sup>, Alessandra Querino da Silva<sup>2</sup>, Carlos Pereira da Silva<sup>1</sup>, Andrezza Kellen Alves Pamplona<sup>1</sup> e Marcio Balestre<sup>1</sup>

<sup>1</sup> DEX/Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> FACET/Universidade Federal da Grande Dourados, MS, Brasil

**Email:** joelnuvungak@gmail.com

**Resumo:** Nesse trabalho analisou-se a serie histórica do índice de crime contra a vida (ICV) da cidade de Maputo/Moçambique referente ao período de janeiro de 1998 a maio de 2009, com o objetivo de estudar a evolução da criminalidade na referida cidade. Utilizou-se o modelo SARIMA, objetivando encontrar um modelo que forneça as melhores previsões. Os resultados apontaram que a sazonalidade do crime em Maputo possui estrutura trimestral relacionada aos períodos de festas e com presença de muitos imigrantes. O melhor modelo foi o SARIMA (0,0,1)x(0,1,1)<sub>4</sub>, mais parcimonioso, apresentando ruído branco e projeções satisfatórias.

## CP148 – Análise estatística no setor financeiro - estudo de um caso de uma instituição paulista

Kaue Gustavo Reis<sup>1</sup> e Vilma M. Tachibana<sup>1</sup>

<sup>1</sup> Universidade do Estado de São Paulo, SP, Brasil

**Email:** vilma@fct.unesp.br

**Resumo:** Este trabalho é referente a modelos de classificação de clientes da carteira de crédito de uma instituição financeira paulista e os dados são de clientes com a carteira já vencida e com histórico de informações conhecidas. Aplicou-se Análise de Agrupamento, Análise Fatorial e Análise de Regressão Logística. Na análise de agrupamento, realizou-se a segmentação de clientes que possuem perfis parecidos entre si e a separação de grupos de variáveis que melhor identificaram um objetivo de crédito. Na análise Fatorial observou-se o comportamento das variáveis e verificaram-se os componentes que melhor explicam os dados, numa dimensão menor de informação, mas que não comprometeram a variabilidade dos dados. Já na Regressão Logística o resultado foi um modelo que também explicou a variabilidade dos dados, contribuindo com boa capacidade preditiva de clientes inadimplentes e de forma simples. Com isso, ações de Marketing e de Crédito, em situações futuras, poderiam ser tomadas por essas segmentações e predições, a fim de melhorar o desempenho da instituição financeira perante a sua carteira de crédito.

## CP149 – Análise da precipitação anual em São João do Tigre no Cariri paraibano utilizando modelos ARMA

Kerolly Kedma Felix do Nascimento<sup>1</sup>, Fábio Sandro dos Santos<sup>1</sup>, Hérica Santos da Silva<sup>1</sup>, Ikaro Daniel de Carvalho Barreto<sup>1</sup> e Tiago Alessandro Espínola Ferreira<sup>1</sup>

<sup>1</sup> Universidade Federal Rural de Pernambuco, PE, Brasil

**Email:** daniel.carvalho.ib@gmail.com

**Resumo:** A região do Cariri paraibano fica situada no Nordeste brasileiro, apresentando acentuada variabilidade em relação à precipitação pluviométrica. Esta região apresenta alguns anos extremamente secos e poucos anos chuvosos. O período chuvoso concentra-se entre os meses de janeiro a maio podendo se estender, em alguns anos, até o mês de julho. Nos demais períodos ocorrem irregularidades de chuvas e baixos índices pluviométricos. Sendo assim, se faz necessário o uso de modelos de séries temporais para precipitação da região do Cariri paraibano. A série temporal de precipitação compreende os anos de 1935 a 2013, no município de São João do Tigre, estado da Paraíba. Para o estudo dessa série foram utilizados os testes da raiz unitária (Dickey-Fuller) e o teste de tendência (Mann-Kendall). Na sequência, ajustou-se aos dados os modelos autorregressivos e de médias móveis (ARMA), adotando-se, como critério de escolha, o valor de Critério de Informação de Akaike (AIC). Neste estudo foi possível identificar no modelo autorregressivo a presença de estacionariedade e ausência de tendência. O gráfico de previsão, obtido a partir do modelo ajustado, permitiu uma interpretação coerente do comportamento da variável precipitação para a próxima década no município analisado.

## CP150 – Análise da complexidade climática do município de Areia – PB através do SampEn

Leandro dos Santos Silva<sup>1</sup>, Ewerton Matias de Lima<sup>1</sup>, Lázaro de Souto Araújo<sup>1</sup>, Péricles Borges de Farias<sup>2</sup>, José Rodrigo Santos Silva<sup>2</sup> e Eucymara França Nunes Santos<sup>2</sup>

<sup>1</sup> Universidade Federal da Paraíba, PB, Brasil

<sup>2</sup> Universidade Federal de Sergipe, SE, Brasil

**Email:** lazaro.souto@hotmail.com

**Resumo:** Com o avanço de novas tecnologias, tornou-se mais fácil realizar previsões mais detalhadas sobre o clima. Elementos de física, estatística e biologia são utilizados para compreender a variabilidade climática que por sua vez, apresentam mudanças no espaço e no tempo. Desta forma, a fim de compreender a complexidade temporal ao longo do tempo, neste trabalho aplicamos o método Sample Entropy (SampEn) para analisar a entropia em séries temporais da temperatura do ar, umidade relativa e velocidade do vento, registradas na estação meteorológica do INMET (Instituto Nacional de Meteorologia) em Areia - PB no período de 2006 a 2012. Os valores da entropia observados para a velocidade do vento foram predominantemente maiores que os observados para as demais variáveis. Por sua vez a umidade relativa foi a série menos irregular dentre aquelas estudadas. Ao observarmos a evolução desta estatística no tempo, notamos que as séries são estacionárias, indicando que a variabilidade destas séries não foi alterada, ao longo do período em estudo.

---

## CP151 – Análise de séries temporais do número de óbitos infantis de Minas Gerais

Mariana de Souza Gonçalves<sup>1</sup> e Letícia Lima Milani<sup>1</sup>

<sup>1</sup> Universidade Federal de Alfenas, MG, Brasil

**Email:** mari.cbq@hotmail.com

**Resumo:** O objetivo deste artigo é fazer uma análise de séries temporais do número de óbitos infantis do estado de Minas Gerais, no período de janeiro de 2000 a dezembro de 2013, utilizando o modelo ajustado por meio da metodologia SARIMA. A mortalidade infantil se trata do número de crianças que morreram antes de completar um ano de idade, e a aplicação de indicadores como o de mortalidade infantil são de extrema importância, já que se fazem necessários para avaliar o desenvolvimento de regiões, além de serem parâmetros indispensáveis para que se faça o controle das condições de saúde populacional. Em geral, a mortalidade infantil tem revelado tendência decrescente em todo o mundo. Sendo assim, o presente trabalho procura explorar com a análise temporal, o estado de Minas Gerais, que foi escolhido por ser de grande relevância econômica, social e cultural, e também o quarto estado brasileiro com maior área territorial e o segundo em número de habitantes. Portanto, o estopo é estudar Minas Gerais e revelar seus níveis e tendências de mortalidade infantil, através da análise do comportamento dessa série temporal, verificando efeitos da sazonalidade e tendência. Em seguida, é feita a modelagem, para que a previsão fosse realizada.

## CP152 – Análise espectral singular aleatorizada na análise de grandes series temporais

Petala Gardenia da Silva Estrela Tuy<sup>1</sup>, Paulo Canas Rodrigues<sup>1</sup> and Rahim Mahmoudvand<sup>2</sup>

<sup>1</sup> Universidade Federal da Bahia, BA, Brasil

<sup>2</sup> Bu-Ali Sina University, Iran

**Email:** petalatuy@hotmail.com

**Resumo:** A análise espectral singular (SSA) é um método de análise de séries temporais relativamente novo e vem como uma alternativa não paramétrica para os métodos clássicos. Esta metodologia vem se mostrando eficiente na análise de séries temporais não estacionárias e com estruturas complexas de sazonalidade, uma vez que, pelo caráter não paramétrico do método, não são necessárias suposições sob estacionariedade ou normalidade dos resíduos. Permitir violações dos pressupostos, todavia, nem sempre é suficiente para lidar com séries temporais complexas. A nova era de Big Data tem trazido desafios na análise de séries de dimensões extremamente grandes. Apesar da SSA apresentar vantagens sobre os métodos convencionais, o tempo de execução do algoritmo na análise de grandes séries temporais pode inviabilizar o uso do método. Neste trabalho apresentamos a SSA aleatorizada (rSSA) que pretende ser uma alternativa para lidar com séries temporais de grandes dimensões sem que se perca qualidade na sua análise. A SSA será comparada com a SSA aleatorizada em termos de qualidade da análise e tempo de execução do algoritmo.

---

## CP153 – Gap-filling in time series by multivariate singular spectrum analysis

Rahim Mahmoudvand<sup>1</sup>

<sup>1</sup> Bu-Ali Sina University, Iran

**Email:** r.mahmodvand@gmail.com

**Abstract:** This paper, we introduce a new algorithm using Multivariate Singular Spectrum Analysis (MSSA) for gap filling in univariate time series. In this algorithm, the data before missing values and the data after (in the reverse order) missing values treated as two separate time series. Then using MSSA forecasting, two estimations of the missing values will be provided. Finally, using bootstrap we combine these estimations and produce a unique estimation for missing values.

## CP154 – Analysis of consumer price index metropolitan region of Belo Horizonte/MG using time series

Ricardo Vitor Ribeiro dos Santos<sup>1</sup>, Érica Fernanda da Cruz<sup>2</sup>, Thelma Sáfadi<sup>2</sup>, Tamara Aparecida Nogueira dos Anjos<sup>2</sup>, Charles Shalimar Felipe da Silva<sup>2</sup> and Cristian Tiago Mendes<sup>2</sup>

<sup>1</sup> Centro Federal de Educação Tecnológica de Minas Gerais, MG, Brasil

<sup>2</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** ricardo3345@gmail.com

**Abstract:** In Brazil the first studies on price indices are 1936. Inflation has been a chronic problem in the 80's and early 90. This study aimed to set a time series model to the national consumer price index in order to check for trends, seasonality, interventions. The Box & Jenkins methodology set a Autoregressive model.

---

## CP155 – Projeção da abate, rebanho e peso total de carcaças de aves em Mato Grosso pelo modelo VAR

Rondiny Carneiro<sup>1,2</sup>, Edilson Júnior<sup>1,2</sup>, Anderson Oliveira<sup>2</sup>, Marcelino Pascoa<sup>2</sup> e Juliano Bortolini<sup>2</sup>

<sup>1</sup> Instituto Mato-Grossense de Economia Agropecuária, MT, Brasil

<sup>2</sup> Universidade Federal do Mato Grosso, MT, Brasil

**Email:** rondinycarneiro@gmail.com

**Resumo:** O estado de Mato Grosso é o 7º maior produtor de aves do país, sendo responsável por 4,5% da produção de carcaças e tendo abatido cerca de 227 milhões de aves em 2014 de acordo com o IBGE. Modelos VAR foram ajustados para abate, rebanho e peso total das carcaças de aves em Mato Grosso. O critério de escolha destes modelos é o de menor AIC. A partir dos modelo VAR escolhido realizou-se projeções no período de 2015 a 2025. Para determinar o intervalo de confiança das projeções foram realizadas simulações de bootstrap, minimizando os erros não captados pelo modelo devido ao tamanho da série. O modelo mostrou-se adequado e prevê crescimento da produção, rebanho e abate de aves em Mato Grosso.

---

## CP156 – Dynamic analysis for event history data: An example of recurrent event

Rosemeire Fiaccone<sup>1</sup>, Robin Henderson<sup>2</sup> and Fabio da Silva<sup>3</sup>

<sup>1</sup> DEST/Universidade Federal da Bahia, BA, Brasil

<sup>2</sup> School of Mathematics and Statistics/Newcastle University, UK

<sup>3</sup> DECOM/Centro Federal de Educação Tecnológica, MG, Brasil

**Email:** fabiorochadasilva@yahoo.com.br

**Abstract:** In many situations, individual subjects or units may experience events that occur repeatedly. We illustrate the use of modern event-history analysis in the analysis of recurrent diarrhoea episodes in three cohorts of infants. The data are complicated by time-dependent covariates, time-dependent effects, intermittent missingness and dropout. In our approach the conditional mean based on the history is modelled as a function of possibly time-varying covariates through an additive regression model for longitudinal binary data subject to both intermittent missingness and dropout. The idea is to show how the array of additive intensity modelling techniques can be applied to longitudinal binary data, providing valuable inferences without highly computationally intensive procedures. In addition to this, it is easy to assess the effect of covariates on the event of interest, even for complex time varying effects of time varying covariates. On the other hand the more natural logistic regression approach can be unstable when events are rare, and the use of the Firth correction may lead to biased predictions, especially when, as here, we sum over a large number of ages.

---

## CP157 – Intervenções na evolução temporal do preço da cesta básica de Salvador

Tamara Aperecida Nogueira dos Anjos<sup>1</sup>, Ricardo Vitor Ribeiro dos Santos<sup>1</sup>, Érica Fernanda da Cruz<sup>1</sup> e Thelma Sáfiadi<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil  
Email: ta\_mara15@hotmail.com

**Resumo:** Todo ser humano possui, dentre outras necessidades básicas, a alimentação. A Cesta Básica, também, conhecida como Cesta Básica Nacional ou ainda Ração Essencial Mínima, foi criada pelo decreto de lei 399, de 7 de maio de 1938, e determina o consumo de alimentos necessário para um adulto, durante um mês de trabalho. A pesquisa da Cesta Básica, realizada pelo Departamento Intersindical de Estatística e Estudos Sócio-Econômicos (DIEESE), abrange dezoito capitais do Brasil, dentre elas, Salvador. Neste trabalho analisa-se a série temporal dos preços reais da cesta básica praticadas em Salvador, BA. Os dados foram obtidos por meio do DIEESE e compreendem o período de janeiro de 1995 a outubro de 2015. Foi verificada uma tendência crescente na série  $T = 52,02 + 0,76t$ . Além disso, foram utilizadas funções de transferência para a análise de intervenção devido a variação dos preços dos componentes da cesta básica. Utilizando-se o modelo verificou-se que a partir de janeiro de 2013, houve um aumento de aproximadamente 35 reais no preço da cesta básica; a partir de janeiro de 2013, o aumento foi de aproximadamente 29 reais e; em maio de 2015, houve um aumento de aproximadamente 31 reais.

---

## Sessão de Pôsters 2 / Poster Session 2

---

### Estatística Computacional [*Computational Statistics*]

---

#### CP158 – Redes neurais artificiais na identificação de padrões de ocorrência de Fusões e Aquisições (F&A) nas instituições financeiras Brasileiras

Lucas Pereira Lopes<sup>1</sup>, Gabriel Rodrigo Gomes Pessanha<sup>1</sup>, Beatriz de Souza Bernadinho<sup>1</sup> e Caio Tonolli Dias Gentile<sup>1</sup>

<sup>1</sup> Universidade Federal de Alfenas, MG, Brasil  
Email: beatrizsb08@gmail.com

**Resumo:** Na tentativa de criação de valor para os acionistas, as organizações, diante da concorrência do mercado, vêm buscando ferramentas e alternativas para auxiliar na sua gestão e desempenho. Neste contexto, o processo de Fusões e Aquisições (F&A) se destaca. Considerando o volume de transações de F&A envolvendo empresas brasileiras no setor bancário e a importância de tal estratégia de crescimento nos resultados das empresas, torna-se relevante estudar e analisar a existência de padrões de ocorrência destes processos e desenvolver metodologias de previsão a partir de variáveis que expressam as características gerais, econômicas e financeiras da firma e que são capazes de influenciar a ocorrência do processo de fusão e/ou aquisição. Com isso, o presente trabalho tem por objetivo apresentar um modelo de previsão baseado em uma rede neural artificial (RNA) para predição de operações de Fusões e Aquisições (F&A) nas instituições bancárias brasileiras. Para a modelagem considerou-se duas categorias, a primeira são os períodos em que F&A não são observados e a segunda são os períodos em que F&A são observadas. Como resultado, a RNA ajustada foi capaz de classificar corretamente na fase de teste, aproximadamente, 98,7% das observações. Revelando uma maior capacidade preditiva dos períodos em que F&A não ocorrem com 100% dos períodos corretamente previstos, já para os períodos em que as estratégias de F&A não são adotadas, a rede classificou corretamente cerca de 97,0% dos casos estudados. Os resultados mostram que as redes produzem níveis satisfatórios de previsão, tornando-se assim uma poderosa ferramenta para tomada de decisão.

---

#### CP159 – Criação de redes neurais no Statistica: apresentação dos principais recursos do software

Mariana Nolde Pacheco<sup>1</sup> e Lisiane Priscila Roldão Selau<sup>1</sup>

<sup>1</sup> Universidade Federal do Rio Grande do Sul, RS, Brasil  
Email: lisianeselau@gmail.com

**Resumo:** Redes Neurais Artificiais (RNA) se constituem em uma das técnicas mais recentes de análise de dados e podem recuperar grande quantidade de dados e reconhecer padrões baseados no acúmulo de experiência. São utilizados amplamente em diversas áreas de conhecimento tais como administração, finanças, marketing e saúde. Existem diferentes softwares disponíveis no mercado atualmente que realizam esse tipo de técnica, porém a melhor forma de realização das análises e interpretação dos resultados ainda é pouco divulgada. Dessa forma, o objetivo desse estudo é apresentar algumas formas de utilização das redes neurais com um dos softwares que apresenta o recurso de análise de redes neurais, o STATISTICA da Dell, versão 13. Sendo assim, são apresentadas as principais opções de recursos e ferramentas de análise e construção das RNA no software.

---

## CP160 – Predição da rentabilidade: Uma aplicação da metodologia de redes neurais artificiais em empresas brasileiras de capitalização

Richard Hernani Pereira<sup>1</sup>, Leandro Ferreira<sup>2</sup>, Gabriel Rodrigo Gomes Pessanha<sup>2</sup> e Lucas Pereira Lopes<sup>2</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> Universidade Federal de Alfenas, MG, Brasil

**Email:** richard.hernani@gmail.com

**Resumo:** O objetivo deste trabalho é predizer, através de modelos baseados em Redes Neurais Artificiais, o comportamento dos índices de rentabilidade, frente aos indicadores macroeconômicos e índices calculados sob o critério de avaliação CAMEL, bem como analisar o nível de importância destas variáveis sobre a rentabilidade das empresas brasileiras de capitalização. A análise de índices de rentabilidade é uma das formas de avaliar a situação econômico-financeira da empresa, pois possuem a característica de fornecerem uma visão apurada sobre a sociedade. Os indicadores macroeconômicos são estatísticas que corroboram para aumentar o poder de decisão e planejamento do mercado financeiro, visto que suas oscilações podem causar significativos impactos ou rever algumas expectativas, enquanto o critério de avaliação CAMEL trata-se de uma ferramenta analítica que possibilita um respaldo maior sobre a tomada de decisões das instituições. Redes neurais artificiais é uma das técnicas de dados mais recentes e que tem motivado um grande número de pesquisadores da área de tecnologia e da área de negócios. Os resultados apontam que os modelos propostos apresentam significativa capacidade em predizer os valores dos índices de rentabilidade, revelando que a metodologia aplicada é eficiente, quando considerados os diagnósticos dos resultados e o grau de importância das variáveis nos modelos de predição.

---

## Sessão de Pôsters 2 / Poster Session 2

---

### Estatística Bayesiana [*Bayesian Statistics*]

---

#### CP161 – Bayesian binary overdispersed longitudinal model via marginalized likelihood

Afranio M.C. Vieira<sup>1</sup> and Everton B. Rocha<sup>2</sup>

<sup>1</sup> Universidade Federal de São Carlos, SP, Brasil

<sup>2</sup> Universidade Federal de Goiás, GO, Brasil

**Email:** afranio.ufscar@gmail.com

**Abstract:** Many experiments and observational studies produce binary (or dichotomous) responses. Some of these studies have a typical longitudinal structure. Once these kind of data structure express naturally correlation between the repeated measures of a subject, usually, marginal models based on Generalized Estimating equations (GEE) or subject-specific models like Generalized Linear Mixed Models (GLMM) perform well accommodating the auto correlation or clustered correlated data. However, when both sources of correlation express themselves an extended latent structure is demanded. Molenberghs et al. (2010) have proposed an extended class of GLMM with multiplicative conjugated random effects. On a frequentist likelihood estimation, the optimization phase includes numerical integration methods like adaptive Gaussian quadrature. In this work will be presented on a Bayesian perspective the use of an analytical marginalized likelihood, what reduces the computational work.

---

#### CP162 – Avaliação de híbridos de milho utilizando o modelo AMMI com efeitos fixos e o AMMI-Bayesiano

Alessandra Querino da Silva<sup>1</sup>, Luciano Antonio de Oliveira<sup>1</sup>, Carlos Pereira da Silva<sup>2</sup>, Joel Jorge Nuvunga<sup>2</sup>, Andrezza Kellen Alves Pamplona<sup>2</sup> e Marcio Balestre<sup>2</sup>

<sup>1</sup> FACET/Universidade Federal da Grande Dourados, MS, Brasil

<sup>2</sup> DEX/Universidade Federal de Lavras, MG, Brasil

**Email:** lucianoantonio@ufgd.edu.br

**Resumo:** O modelo de efeitos principais aditivos e interação multiplicativa (AMMI) tem ampla utilização para analisar a interação entre genótipos e ambientes (GE) em programas de melhoramento genético de plantas. Nesse trabalho foi analisado um conjunto de dados provenientes da avaliação de 55 híbridos de milho em 9 ambientes distintos utilizando o modelo AMMI tradicional (efeitos fixos) e modelo AMMI-bayesiano com priori informativa para genótipos. O objetivo foi comparar essas abordagens principalmente no que se refere a interpretação biplot. Observou-se que o AMMI-bayesiano oferece uma maior flexibilidade

para atribuir efeito aleatório para genótipos e maior facilidade para implementar inferência ao biplot que tem sido uma grande limitação para a abordagem frequentista.

---

## CP163 – Análise bayesiana de delineamentos triangulares para experimentos sensoriais

Camilla Barroso<sup>1</sup> e Júlio Bueno Filho<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** camillamarques84@gmail.com

**Resumo:** Testes discriminativos como os testes triangulares são muito usados na análise sensorial para detectar diferenças entre produtos. Cada provador recebe três amostras codificadas, sendo 2 iguais e uma diferente. Ele deve identificar a amostra diferente e pelo delineamento há 1/3 chance de acerto ao acaso. A análise usual considera uma aproximação gaussiana para um modelo linear (efeitos fixos). Em um trabalho anterior, propusemos um modelo generalizado linear (GLM) para as proporções utilizando a distribuição binomial. Este modelo ignora a correlação existente entre os efeitos de provadores. Agora nós estendemos o modelo incorporando efeitos aleatórios para os provadores. Foi utilizado um exemplo utilizando dados de um experimento de contaminação de café realizado na Universidade Federal de Lavras. Um algoritmo de dados aumentados foi adaptado para amostrar a partir da distribuição a posteriori. Os resultados do modelo misto usando análise Bayesiana são apresentados e comparados aos resultados desconsiderando agrupamentos nos efeitos de provadores. O modelo proposto é de confiança e pode ser utilizado em análises de rotina em ciência dos alimentos.

---

## CP164 – Uma aplicação do modelo AMMI-bayesiano-shirincage na análise de estabilidade e adaptabilidade genotípica

Carlos Pereira da Silva<sup>1</sup>, Luciano Antonio de Oliveira<sup>2</sup>, Andrezza Kellen Alves Pamplona<sup>1</sup>, Joel Jorge Nuvunga<sup>1</sup>, Alessandra Querino da Silva<sup>2</sup> e Marcio Balestre<sup>1</sup>

<sup>1</sup> DEX/Universidade Federal de Lavras, MG, Brasil

<sup>2</sup> FACET/Universidade Federal da Grande Dourados, MS, Brasil

**Email:** lucianoantonio@ufgd.edu.br

**Resumo:** Nesse trabalho aplicamos o modelo AMMI-bayesiano-shrincage a um conjunto de dados provenientes do ensaio de 12 genótipos de mostarda a 6 ambientes distintos. O delineamento foi o de blocos completos casualizado com três repetições. Observou-se efeito de encolhimento para os valores singulares, porém não muito expressivos para os três primeiros eixos principais em comparação com outros dois métodos. Certamente isso se deve pela presença de interação complexa. Regiões de credibilidade foram implementadas para efeitos principais de genótipos, bem como para escores genotípicos e ambientais que permitiram: separar subgrupos homogêneos de genótipos e ambientes; identificar genótipos e ambientes estáveis e observar adaptabilidade de genótipos a ambientes específicos.

## CP165 – Cálculo das condicionais a posteriori para a análise bayesiana do modelo fatorial dinâmico para um vetor de séries temporais utilizando distribuição t multivariada

Larissa Ribeiro de Andrade<sup>1</sup>, Daniel Furtado Ferreira<sup>1</sup> e Thelma Sáfyadi<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** larissa.mat@gmail.com

**Resumo:** Os modelos t multivariados são simétricos e com caudas mais pesadas que a distribuição normal, característica importante na análise de séries financeiras. A proposta deste trabalho é obter as condicionais a posteriori, para a análise Bayesiana do modelo fatorial dinâmico, na classe de modelos t multivariados. Devido à complexidade da distribuição t multivariada, essa variável foi representada, neste trabalho, como uma mistura da distribuição normal multivariada com uma raiz de qui-quadrado. Este artifício permitiu o cálculo das distribuições a posteriori de interesse.

---

## CP166 – Ajuste de modelos não lineares com enfoque bayesiano aos dados de comprimento do fruto da pereira Asiática

Rick Mangueira<sup>1</sup>, Taciana Savian<sup>1</sup> e Erasnilson Vieira Camilo<sup>1</sup>

<sup>1</sup> ESALQ/Universidade de São Paulo, SP, Brasil

**Email:** erasnilson@hotmail.com

**Resumo:** O estudo de curvas de crescimento do fruto são importantes para obter conhecimento sobre seu crescimento e seus estádios fenológicos, podendo-se estabelecer o período adequado para colheita, estratégias para conservação do fruto e controle de pragas. Pode-se obter esse conhecimento fazendo um estudo utilizando modelos não lineares, para se obter informações por meio de parâmetros com interpretações biológicas que trazem consigo um resumo sobre a curva característica do crescimento do fruto. Este trabalho foi realizado com o objetivo de descrever a curva de crescimento do comprimento do fruto da pereira Asiática da cultura Shinsseiki, ao longo do tempo, por meio de uma abordagem bayesiana. Os modelos de crescimento utilizados para este fim foram o logístico e o Gompertz. A metodologia bayesiana foi usada para estimação dos parâmetros dos modelos estudados. Os dados utilizados no trabalho são referentes ao comprimento (mm) do fruto da pereira asiática da cultivar Shinsseiki, no qual foram mensurados 20 frutos em cada dia de observação, totalizando 11 observações ao longo de 154 dias. Em cada dia de observação foi calculada a média do comprimento do fruto e realizado o ajuste das curvas. Obteve-se um bom ajuste para os modelos logístico e Gompertz. Considerando o critério de informação de qualidade de ajuste DIC, o modelo logístico apresentou o melhor ajuste a curva de crescimento do comprimento do fruto da pereira Asiática.

## CP167 – Método Bayesiano para estimação de efeitos latentes com desfechos distais

Silvia Regina Morais<sup>1</sup>, Lilia Carolina Costa<sup>1</sup> e Leila D.A.F. Amorim<sup>1</sup>

<sup>1</sup> Universidade Federal da Bahia, BA, Brasil

**Email:** slemos@ufba.br

**Resumo:** Nos últimos anos, pesquisadores como Lanza et al. (2013) e Asparouhov & Muthén (2014) têm proposto técnicas com abordagem frequentista para estimação de efeitos latentes em desfechos distais. Em problemas deste tipo é de interesse avaliar o efeito de variáveis não observadas, geralmente provenientes do uso de análise de classes latentes (Latent Class Analysis, LCA, em inglês) em variáveis respostas observadas que podem ser binárias, de contagem ou contínuas. Estudos recentes consideram modelagem simultânea da classe latente e o seu efeito no desfecho distal através do uso de regra de Bayes a partir dos resultados de LCA com covariáveis (Lanza et al., 2013) ou a incorporação dos erros de mensuração obtidos no LCA diretamente na modelagem com o desfecho distal (Asparouhov e Muthén, 2014). O objetivo deste trabalho foi apresentar uma terceira estratégia para estimação de efeitos latentes em desfechos distais usando uma abordagem de inferência bayesiana. Os resultados numéricos da simulação de Monte Carlo revelaram que as estimativas encontradas tendem a ser próximas do verdadeiro valor, dos parâmetros do modelo proposto, e precisas. Além disso, o erro-padrão e raiz quadrada do erro quadrático médio das estimativas diminuíram à medida que os tamanhos amostrais cresceram. Portanto, os resultados da simulação demonstraram que a estratégia de estimação de efeitos latentes em desfechos distais usando uma abordagem bayesiana pareceu satisfatória para o caso particular apresentado no presente trabalho.

## CP168 – Construção de prioris utilizando o conceito de máxima entropia: aplicação a dados genéticos

Tais Alvarenga<sup>1</sup>, Ernandes G. Moura<sup>1</sup> e Marcio Balestre<sup>1</sup>

<sup>1</sup> Universidade Federal de Lavras, MG, Brasil

**Email:** ernandesfederal@hotmail.com

**Resumo:** Com o advento das novas técnicas de genotipagem, a análise genômica ampla tem ganhado espaço, principalmente dado ao fato de que grande quantidade de informação tem sido disponibilizada. No entanto, com o aumento no número de marcas problemas como a multicolinearidade e/ou sobreajuste de modelos tem dificuldade a abordagem estatística e interpretação genética de modelos usados na análise genômica ampla, com isso modelos bayesianos tais como, Bayes A e Bayes B tem sido alvo de críticas. Este trabalho foi realizado com o objetivo de estabelecer uma justificativa genética na derivação de prioris para os efeitos genéticos e suas variâncias de modo que o aprendizado bayesiano possa ser garantido, fazendo com que modelos de marca específicas possam ter algum significado genético, abordados através de derivação de prioris sobre a ótica da máxima entropia. Utilizou-se de uma simulação de 6000 SNPs distribuídos em 6 grupos de ligação em 500 indivíduos pertencentes a uma população F2, com o intuito de simular o máximo de situações dos comportamento dos dados para posteriormente comparar com o modelo de máxima entropia.

# Index

- Afonso, S.D.J., 92, 145  
Albertini, T.Z., 117  
Albuquerque, M., 89, 99  
Almeida, A.C.L., 83  
Almeida, E.H.P., 120  
Almeida, I.S.F., 139  
Almeida, M.F.F., 97, 154, 158  
Almeida, T., 142  
Alvarenga, T., 180  
Alvarenga, T.C., 153  
Alves, A., 155  
Alves, A.X.P., 147  
Alves, B.M., 118, 121, 137, 138, 144  
Alves, E., 96  
Alves, J.A., 105, 107, 128  
Amaral, M.T.R., 110  
Amorim, L.D.A.F., 90, 93, 121, 139, 143, 155, 180  
Ananias, M., 164  
Andrade, B., 164  
Andrade, F., 132  
Andrade, H., 98, 106  
Andrade, J., 96  
Andrade, J.M.L., 164  
Andrade, M., 56, 78  
Andrade, M.G., 106  
Andrade, M.J.B., 103, 109  
Andrade, T.C.M., 102  
Araújo, E., 94  
Araújo, L.B., 150, 151  
Araújo, L.S., 171  
Araújo, M.F.C., 150  
Ashcroft, D., 124  
Assis, D.F., 156  
Assumpção, R.A.B., 132  
Azevedo, C., 152
- Bailey, T., 143  
Balduino, R., 132  
Balestre, M., 115, 131, 169, 177, 178, 180  
Barbosa Jr, F., 85  
Barbosa, G., 155  
Barbosa, M., 115, 131, 145  
Barbosa, T., 96  
Barg, S., 131  
Barioni Jr., W., 67  
Barioni, L.G., 117  
Barreto, I.D.C., 170  
Barreto, M., 155  
Barreto, M.L., 155  
Barros Neto, J.F., 156  
Barros, W., 149  
Barroso, C., 178  
Batista, A.P., 48  
Batista, E.R., 118
- Batista, M.E.O., 97  
Benz, V., 80  
Bereta, E.M.P., 91  
Bernadinho, B.S., 165, 166, 175  
Bertonha, A., 77  
Bianco, C.I., 139  
Biase, A.G., 117  
Bicalho, C.C., 156  
Bicalho, E.S., 162  
Bizerra, E., 99  
Blanck, E., 106  
Boer, M., 53  
Bolaina, G.M.R., 95, 111  
Bortoloni, J., 173  
Boruch, F., 130  
Bracco, M.M., 120  
Braga Jr, R.L.C., 153  
Braga, J.S., 109, 128  
Brancaglioni, V.A., 97, 154, 158  
Brevi, F.P., 85  
Brien, C.J., 64  
Brito, A., 89, 99  
Brito, F., 126  
Brito, M.C.M., 141  
Bueno Filho, J., 178  
Bueno, J., 152
- Calegario, N., 128  
Calegario, N., 105, 107  
Callegari, F.V.R., 104  
Calsavara, V., 76  
Camilo, E.V., 112, 179  
Campos, A.C., 90, 121, 143  
Candolo, C., 104  
Cannavan, F.S., 162  
Cargnelutti Filho, A., 118, 121, 137, 138, 144  
Carneiro, R., 173  
Carneseca, E.C., 159  
Carvalho, J.B., 128  
Carvalho, M.L., 157  
Carvalho, M.M., 153  
Cavalcanti, C., 151  
Cazorla, I., 65  
Chamapiwa, E., 124  
Chaves, A.D.C.G., 93  
Chaves, G.G., 137, 144  
Chaves, L., 168  
Chiaravalloti Neto, F., 149  
Chinelato, P., 112  
Coatti, G., 77  
Colaço, W., 89  
Colosimo, E., 76  
Conceição, A.L.S., 92, 145  
Conceição, K., 81

- Corrêa, E.R., 149  
 Correia, L.C.L., 72  
 Corrente, J.E., 127  
 Cortese, G., 75, 131  
 Costa, A.C., 113  
 Costa, D.S., 122, 127, 160  
 Costa, E., 80  
 Costa, L.C., 143, 158, 180  
 Costa, R.C., 118  
 Costanzi Filho, A., 101  
 Cristino, C.T., 119  
 Cruz, A.C.C., 91
- da Conceição, J.S., 72, 141  
 da Cruz, E.F., 173, 174  
 da Silva, A.L., 141  
 da Silva, A.Q., 169, 177, 178  
 da Silva, A.R., 122, 127, 160  
 da Silva, C.C., 147  
 da Silva, C.P., 169, 177, 178  
 da Silva, C.S.F., 168, 173  
 da Silva, F., 161, 173  
 da Silva, H.S., 170  
 da Silva, J.P., 150  
 da Silva, M.J.C., 147  
 da Silva, M.S., 92, 145  
 da Silva, P.V., 114  
 da Silva, S.B., 83  
 Damaceno, N., 145  
 de Almeida, I.G., 144  
 de Almeida, P.L., 125  
 de Almeida, R.R.P., 93  
 de Almeida, S.S., 96  
 de Andrade Filho, M.G., 166  
 de Andrade, A.P., 85  
 de Andrade, B.S., 166  
 de Andrade, L.R., 179  
 de Andrade, M.J.B., 108  
 de Bem, C.M., 137  
 de Castro, V.L.S.S., 110  
 de Freitas, R.T.F., 113  
 de Jesus, J.A., 139  
 de Lima, C.G., 104, 110, 131  
 de Lima, E.M., 171  
 de Lima, I.M.B., 109  
 de Lima, K.P., 108, 159  
 de Lima, P.C.G., 130  
 de Lima, R.R., 90, 153  
 de Lima, V.A., 130  
 de Medeiros, E.S., 125  
 de Medeiros, S.R., 117  
 de Mello, J.M., 48  
 de Oliveira, I.R.C., 98  
 de Oliveira, L.A., 177, 178  
 de Oliveira, L.P., 110  
 de Oliveira, M.S., 168  
 de Oliveira, T.B., 125  
 de Sousa, D.J.A., 93  
 de Sousa, S.B., 108, 159, 168
- de Souza, E.M., 124  
 de Souza, L.D.A., 109  
 de Souza, T.V., 113  
 Delfino, A.C., 167  
 Demétrio, C.G.B., 48, 64  
 Destro Filho, J.B., 126  
 Dias, C.T.S., 117, 143, 150, 151  
 Diggle, P.J., 40, 71  
 do Amaral, M.F., 95, 111  
 do Nascimento, K.K.F., 170  
 do Prado, J.R., 150  
 Domingues, J., 97, 154, 158  
 Dore, L., 164  
 dos Anjos, T.A.N., 173, 174  
 dos Reis, J.S., 98  
 dos Santos, A., 64  
 dos Santos, C.C., 145  
 dos Santos, C.S.A., 93  
 dos Santos, F.S., 170  
 dos Santos, M.L.A., 124  
 dos Santos, R.V.R., 168, 173, 174  
 dos Santos, S.E.G., 147  
 dos Santos, V.K., 119  
 dos Santos, W.V., 122, 127, 160  
 Drebes, L., 122, 123  
 Dreifke, L.G., 122, 123  
 Duarte, A., 83  
 Duarte, N.E., 78  
 Duczmal, L., 83
- Edilson Jr., 173  
 Elisete, A., 145  
 Espínola, T., 155  
 Etienne, Q., 84
- Facco, G., 118, 137  
 Farias, P.B., 171  
 Feitosa, J.V., 104  
 Fernandes, A., 90, 109  
 Ferraudo, A.S., 162  
 Ferreira Filho, P., 91  
 Ferreira, C., 83  
 Ferreira, D.F., 101, 149, 179  
 Ferreira, I., 161  
 Ferreira, L., 176  
 Ferreira, P.H., 59  
 Ferreira, R., 76  
 Ferreira, T.A.E., 170  
 Ferreira, T.A.P., 101  
 Fiaccone, R., 113, 155, 173  
 Figueiredo Neto, A., 145  
 Filzmoser, P., 58  
 Finger, F., 149  
 Follmann, D.N., 144  
 Fonrozo, S.E., 153  
 Fonseca, M., 61  
 Fontes, D., 72  
 França, R.W., 132  
 Francisco, L.B., 129

- Freato, T.A., 113  
Freitas, M.C., 145  
Freitas, U., 95  
Frighetto, R.T.S., 118  
Fumes, G., 127
- Galletti, A., 164  
Gandasi, N., 131  
Gbete, S.D., 84  
Gentile, C.T.D., 165, 166, 175  
Giampaoli, V., 55, 145  
Gomes, G.S.S., 144  
Gonçalves, L.R., 97, 169  
Gonçalves, M.S., 171  
Gonzalez, F., 157  
Gonzatto Jr, O.A., 135  
Govone, J.S., 97, 154, 158  
Guedes, T.A., 135  
Guimarães, E.C., 95, 111  
Guimarães, P.H.S., 112, 162
- Hamedani, G., 75, 142  
Hebling, L., 81  
Henderson, R., 173  
Henriques, M.J., 135  
Higa, R.H., 68  
Hinde, J., 71  
Holmboe, S., 75  
Hongyu, K., 147, 151  
Horimoto, A.R.V.R., 78  
Hrůzová, K., 58  
Hron, K., 58  
Hubert, M., 57
- Isaac Jr, M.A., 105, 128
- Janeiro, V., 135  
Jesus, E.S., 102
- Kleinpaul, J.A., 137, 138, 144  
Kontopantelis, E., 124  
Krieger, J.E., 78  
Krzanowski, W., 143
- Léskow, J., 166  
Lúcio, A.D., 80, 122, 123  
La Scala Jr, N., 152, 154, 162  
Lachos, V.H., 83  
Lanna, D.P., 117  
Lassance, R., 164  
Lavezo, A., 137, 144  
Leandro, R.A., 63, 112  
Ledo, C.A.S., 67, 92, 145  
Lee, D.-J., 53  
Leite, I., 157  
Lemes, A.M., 91  
Lima, P., 90  
Lima, V., 107  
Link, C.F., 132  
Lira, A.S., 102, 119
- Lisboa, J.L.C., 96  
Lopes, B.A., 165  
Lopes, L.P., 103, 146, 165, 166, 175, 176  
Lopes, R., 80  
Lourenço, P., 142  
Lourenço, V.M., 57, 61  
Louzada, F., 59, 60, 66, 85, 115  
Ludeña, C., 155  
Luiz, A.J.B., 118  
Luna, J.G., 112
- Madeira, R.A.V., 156  
Maestre, M.R., 125, 129  
Mafra, A.C.C.N., 102, 120  
Mahmoudvand, R., 172  
Maia, A., 89  
Maia, A.H.N., 110  
Maia, A.L.S., 144  
Maia, J., 107  
Maia, L., 151  
Maia, R.P., 63  
Malaquias, J.B., 112  
Malhaeiros, E.B., 104  
Manfrim, A.M.P., 120  
Mangueira, R., 179  
Marchezini, B.R., 103, 146  
Marques Jr, J., 157  
Martins Filho, M.V., 123  
Martins, A.B.T., 135  
Martorano, L.G., 117  
Masuda, E.T., 149  
Mateus, A.L., 167  
Mayer, F., 47  
Medeiros, E.S., 129  
Medina, R.R.S.T., 120  
Melo, E.A., 105, 107  
Melo, I.R.S., 109, 128  
Mendes, C.T., 173  
Mendonça, E., 112, 125  
Meyer, A.S., 120  
Milani, H., 78  
Milani, L.L., 165, 167, 171  
Millani, L.L., 167  
Moitinho, M.R., 152, 162  
Molenberghs, G., 45, 62  
Monteiro-Da-Rocha, F.M., 77  
Morais, A., 103, 109  
Morais, A.R., 108, 114  
Morais, E., 103  
Morais, S.R., 180  
Moreira, E., 136  
Moreira, R.F.C., 92  
Moura, E., 132  
Moura, E.G., 105, 180  
Mukherjee, B., 45, 61  
Muniz, J., 109, 114, 115
- Nardino, M., 149  
Natalin, C.A.M., 120

- Negrão, A.B., 78  
 Nojosa, R.T., 108  
 Nunes, E., 148  
 Nunes, M., 49  
 Nuvunga, J.J., 169, 177, 178  
  
 Olinda, R., 99, 112, 125  
 Oliveira, A., 147, 173  
 Oliveira, D., 103  
 Oliveira, D.P., 108  
 Oliveira, E., 106  
 Oliveira, F.L.P., 83  
 Oliveira, G., 115  
 Oliveira, I.R.C., 149  
 Oliveira, L.A., 169  
 Oliveira, P., 81  
 Oliveira, P.P.A., 117  
 Oliveira, T., 125  
 Oltjen, J.W., 117  
 Ossani, P.C., 153  
 Ozelame, C., 147  
  
 Pacheco, M.N., 111, 175  
 Padovani, C.R., 110  
 Paes, A.T., 72  
 Paes, N., 151  
 Paes, N.A., 94, 128  
 Paiva, A., 155  
 Pamplona, A.K.A., 169, 177, 178  
 Panosso, A.R., 152, 154  
 Pascoa, M., 173  
 Pavan, J., 101  
 Peccin, F., 94, 101  
 Pedersen, M.G., 131  
 Pedroso-De-Lima, A.C., 77  
 Perdoná, G., 85  
 Perdoná, G.S.C., 60  
 Pereira, A.C., 78  
 Pereira, E.Q., 124, 135  
 Pereira, G., 82  
 Pereira, G.M., 149  
 Pereira, G.T., 123, 157  
 Pereira, I., 142  
 Pereira, L.A., 136  
 Pereira, M.B., 84  
 Pereira, O., 77, 78  
 Pereira, R., 90  
 Pereira, R.H., 176  
 Pereira, S.R., 143  
 Peria, F., 85  
 Peticarrari, A.L.P., 120  
 Pessôa, G.C.M., 93  
 Pessanha, G.R.G., 97, 103, 146, 165, 166, 169, 175, 176  
 Peternelli, L., 136  
 Pezzato, A.C., 129  
 Pezzini, R.V., 137, 138  
 Philippsen, A.S., 91, 106  
 Pinheiro, C., 82  
 Pinto, E.R., 136  
  
 Pires, A.M., 61  
 Pires, M., 76  
 Possato, E.L., 105, 107, 128  
 Prado, D., 168  
 Previdelli, I., 77, 78  
 Protazio, J.M., 99  
  
 Queiroz, S.E.E., 162  
  
 Rabelo, M., 90  
 Ramos, E.M.L.S., 96  
 Raymaekers, J., 57  
 Reeves, D., 124  
 Rego, N., 98  
 Reis, K.G., 170  
 Reis, M., 100  
 Resende, E., 126  
 Rezende, B.R., 122, 127, 160  
 Rezende, J.C.R., 129  
 Ribeiro Jr, E.E., 46  
 Ribeiro Jr, G., 114  
 Ribeiro, A., 101  
 Ribeiro, A.H., 49  
 Ribeiro, C.M., 167  
 Ribeiro, M.H.D.M., 78, 130  
 Ribeiro, P.C.M., 168  
 Ribeiro, R., 99, 126  
 Ribeiro, T., 103, 109, 114, 115  
 Rocha, A.C.P., 104  
 Rocha, E.B., 177  
 Rodríguez-Álvarez, M.J., 53, 54  
 Rodrigues, A., 76  
 Rodrigues, A.A.A., 126, 136  
 Rodrigues, A.M., 122, 123  
 Rodrigues, J.A., 75, 141, 142  
 Rodrigues, M.M., 55  
 Rodrigues, P.C., 61, 71, 148, 172  
 Rosa, F.H., 85, 94  
 Rosa, K., 136  
 Rosa, L., 161  
 Rousseeuw, P.J., 57  
 Russo, S., 135  
  
 Sáfadi, T., 150, 157, 173, 174, 179  
 Saad, F.M.O.B., 98  
 Santana, C.S., 122, 123  
 Santana, H.A., 122, 127, 160  
 Santana, L., 164  
 Santoro, K.R., 102, 119  
 Santos, A., 95, 161  
 Santos, B., 121  
 Santos, C.O., 167  
 Santos, E.F.N., 171  
 Santos, G., 161  
 Santos, G.O., 102  
 Santos, J.E., 95  
 Santos, L.N., 102  
 Santos, P.R., 98  
 Santos, S., 99, 100  
 Santos, V., 100

- Sari, B.G., 122, 123  
Sartori, J.R., 129  
Sartori, M.M.P., 129  
Sartorio, S.D., 104, 139  
Savian, T., 179  
Scalon, J.D., 113, 156, 159  
Schabarum, D.E., 138  
Scheike, T.H., 75  
Schuelter, A.R., 149  
Segaert, P., 57  
Selau, L.P.R., 111, 175  
Sena, S., 139, 143, 155  
Senna, V., 82  
Sermarini, R.A., 48, 64  
Sicchieri, M.P.L.S., 159  
Silva Jr, J.F., 123  
Silva, A., 75, 142, 143, 164  
Silva, A.P., 75, 142  
Silva, A.R., 119  
Silva, C., 135  
Silva, D., 151  
Silva, E., 142  
Silva, F., 151  
Silva, G., 77, 132  
Silva, G.A., 107, 128, 153  
Silva, G.C.C., 105  
Silva, G.O., 141  
Silva, I.R., 149  
Silva, J., 85, 94, 95, 101, 132, 135  
Silva, J.R.S., 171  
Silva, J.V.R., 150  
Silva, L., 103, 109, 114  
Silva, L.S., 171  
Silva, M.L., 99  
Silva, P.L.N., 41, 65, 71  
Silva, R., 98, 99, 111  
Silveira, D.L., 137  
Simões, F.M., 138, 144  
Singer, J., 77, 80  
Siqueira, D.S., 157  
Soler, J.M.P., 49  
Sousa, I.F., 148  
Sousa, S.B., 108  
Souza Jr, C., 143  
Souza, C.N.P., 93, 96, 148  
Souza, D., 168  
Souza, E., 169  
Souza, G.M., 119  
Souza, G.S., 67  
Souza, H., 85  
Souza, L.J.N., 119  
Souza, P., 132  
Souza, T., 99  
Stefanelo, L.S., 137, 144  
Stosic, T., 164  
Tachibana, V.M., 170  
Taconeli, C.A., 46  
Taporoski, T.P., 78  
Teixeira, D.B., 123, 152, 157, 162  
Thomasi, R.M., 137, 138  
Thompson, E.A., 39, 71  
Tischer, A.L., 122  
Tischler, A.L., 123  
Todorov, V., 58  
Toebe, M., 118  
Tomazella, V., 65  
Torres Jr, R.A.A., 117  
Trinca, L.A., 55, 63  
Tsai, S.M., 162  
Tuy, P.G.S.E., 172  
Tuzine, M.S., 48  
Uliana, D.B., 137, 138  
Vallada, H., 78  
van den Bossche, J.W., 57  
Vasconcellos, V., 157  
Vedovato, B., 129  
Veiga, E.P., 46  
Veloso, G., 76  
Veloso, M.V.S., 152, 153  
Venturini Filho, W.G., 129  
Verruma-Bernardi, M.R., 139  
Victor, E.S., 102  
Vieira, A.M.C., 177  
Vieira, N., 109  
Vigas, V., 115  
Vivanco, M.J.F., 46  
von Schantz, M.V., 78  
Waldman, E.A., 149  
Wanderley, W., 99  
Wartha, C.A., 137, 138, 144  
Yamanouchi, T.K., 114  
Zatz, M., 77  
Zeviani, W., 47  
Zeviani, W.M., 46