

VIII ENCONTRO DOS ALUNOS

Estatística e Experimentação Agronômica



**Anais do VIII Encontro dos Alunos em
Estatística e Experimentação Agronômica**

**Escola Superior de Agricultura Luiz de Queiroz
Piracicaba, 23 de novembro de 2018**

Realização:



Departamento de
Ciências Exatas
da ESALQ/USP



Programa de Pós-
graduação em Estatística e
Experimentação Agronômica

COMISSÃO ORGANIZADORA

(ORGANIZING COMMITTEE)

- Eduardo Elias Ribeiro Junior¹ (ESALQ/USP);
- Pórtia Piscitelli Cavalcanti (ESALQ/USP);
- Welinton Yoshio Hirai (ESALQ/USP);
- Clarice Garcia Borges Demétrio (ESALQ/USP).

COMITÊ CIENTÍFICO

(SCIENTIFIC COMMITTEE)

- Clarice Garcia Borges Demétrio (ESALQ/USP);
- Idemauro Antonio Rodrigues de Lara (ESALQ/USP);
- Rafael de Andrade Moral (Maynooth University);
- Renata Alcarde (ESALQ/USP);
- Thiago Oliveira de Paula (ESALQ/USP);
- Walmes Marques Zeviani (LEG/UFPR).

SÍTIOS ELETRÔNICOS

(WEB PAGES)

- VIII Encontro dos Alunos
<https://esalq-pggee.github.io/encontro2018/>;
- Departamento de Ciências Exatas
<http://www.lce.esalq.usp.br/>;
- Escola Superior de Agricultura Luiz de Queiroz
<http://www4.esalq.usp.br/>.

¹E-mail: jreduardo@usp.br

VIII Encontro dos Alunos

em Estatística e Experimentação Agronômica

| | |
|--|----------|
| 1 MINICURSO | 1 |
| Modelos de Regressão Não Linear | |
| <i>Prof. Dr. Walmes Marques Zeviani</i> | 1 |
| 2 CONFERÊNCIAS | 1 |
| An Extended Random-effects Approach to Modeling Repeated, Overdispersed Count Data | |
| <i>Profa. Dra. Clarice Garcia Borges Demétrio</i> | 1 |
| Análise de Dados Poisson Composto Longitudinais Multivariado | |
| <i>Prof. Dr. Afrânio Marcio Corrêa Vieira</i> | 1 |
| Planejamento para o ajuste de curvas flexíveis | |
| <i>Profa. Dra. Luzia Aparecida Trinca</i> | 2 |
| Alternative methods for modeling of the cure rate in survival studies | |
| <i>Profa. Dra. Vera Lúcia Damasceno Tomazella</i> | 2 |
| 3 COMUNICAÇÕES ORAIS | 3 |
| Seleção de formulações de bebidas prebióticas via Modelos Logitos Cumulativos Mistos | |
| <i>Short-Authors</i> | 3 |
| Imputação múltipla para matriz de interação genótipo e ambiente | |
| <i>Short-Authors</i> | 3 |
| 4 PARTICIPANTES | 4 |

1 MINICURSO

Modelos de regressão não linear: teoria e aplicações

Prof. Dr. Walmes Marques Zeviani (LEG/UFPR)

Em modelos regressão não-linear dados observados de uma variável resposta são descritos por uma função de uma ou mais variáveis explicativas que é não linear seus parâmetros. Assim como nos modelos lineares o objetivo é identificar e estabelecer a relação entre variáveis explicativas e resposta. Entretanto, enquanto os modelos lineares definem, em geral, relações empíricas, os modelos não-lineares são, em grande parte das vezes, motivados pelo conhecimento do tipo de relação entre as variáveis. Desta forma, as aplicações surgem nas diversas áreas onde relações físicas, biológicas, cinéticas, químicas, fisiológicas, dentre outras, são estabelecidas por funções não lineares que devem ter coeficientes (parâmetros) identificados (estimados) a partir de dados observados ou experimentais.

2 CONFERÊNCIAS

An Extended Random-effects Approach to Modeling Repeated, Overdispersed Count Data

Profa. Dra. Clarice Garcia Borges Demétrio (ESALQ/USP)

Non-Gaussian outcomes are often modeled using members of the so-called exponential family. The Poisson model for count data falls within this tradition. The family in general, and the Poisson model in particular, are at the same time convenient since mathematically elegant, but in need of extension since often somewhat restrictive. Two of the main rationales for existing extensions are (1) the occurrence of overdispersion (Hinde and Demétrio 1998, Computational Statistics and Data Analysis 27, 151-170), in the sense that the variability in the data is not adequately captured by the models prescribed mean-variance link, and (2) the accommodation of data hierarchies owing to, for example, repeatedly measuring the outcome on the same subject (Molenberghs and Verbeke 2005, Models for Discrete Longitudinal Data, Springer), recording information from various members of the same family, etc. There is a variety of overdispersion models for count data, such as, for example, the negative-binomial model. Hierarchies are often accommodated through the inclusion of subject-specific, random effects. Though not always, one conventionally assumes such random effects to be normally distributed. While both of these issues may occur simultaneously, models accommodating them at once are less than common. This paper proposes a generalized linear model, accommodating overdispersion and clustering through two separate sets of random effects, of gamma and normal type, respectively (Molenberghs, Verbeke and Demétrio 2007, LIDA, 13, 513-531, Molenberghs et al, 2010, Statistical Science, 25: 325-347, Vangeneugden et al, 2011, Journal of Applied Statistics, 38: 215-232, Molenberghs, Verbeke and Demétrio 2017, SORT, 41, 3-54). This is in line with the proposal by Booth, Casella, Friedl and Hobert (2003, Statistical Modelling 3, 179-181). The model extends both classical overdispersion models for count data (Breslow 1984, Applied Statistics 33, 38-44), in particular the negative binomial model, as well as the generalized linear mixed model (Breslow and Clayton 1993, JASA 88, 9-25). Apart from model formulation, we briefly discuss several estimation options, and then settle for maximum likelihood estimation with both fully analytic integration as well as hybrid between analytic and numerical integration. The latter is implemented in the SAS procedure NLMIXED. The methodology is applied to data from a study in epileptic seizures.

Análise de Dados Poisson Composto Longitudinais Multivariado

Prof. Dr. Afrânio Marcio Corrêa Vieira (UFSCar)

Distribuição Poisson Composta é uma distribuição contínua assimétrica, com massa de probabilidade positiva em $Y = 0$. Registros pluviométricos, valores pagos para apólices de seguros, dentre outras situações apresentam dados com este comportamento. Apresentaremos um problema em que múltiplas expressões bioquímicas de variedades do algodão foram mensuradas ao longo do tempo, sob um delineamento experimental planejado. Na análise, uma estratégia utilizando modelos lineares generalizados misto permite a análise multivariada das expressões bioquímicas, levando em consideração a não-normalidade, dependência temporal e estrutura do delineamento experimental.

Planejamento para o ajuste de curvas flexíveis

Profa. Dra. Luzia Aparecida Trinca (UNESP/Botucatu)

O ajuste de curvas ou superfícies sempre faz parte da análise de resultados experimentais, nos quais procura-se estabelecer relações entre a variável resposta e os vários fatores quantitativos. Os polinômios de segunda ordem são largamente empregados e suas limitações frequentes, devido a simetria imposta, não raramente levam à falta de ajuste e ao uso de modelos de alta ordem nem sempre interpretáveis ou parcimoniosos. Para curvas ou superfícies assimétricas, inclusive com assíntotas, na década de 1990, foram sugeridos os polinômios fracionários, inspirados na família de transformações Box-Tidwell, para análise de dados observacionais. Vários trabalhos mostraram que polinômios fracionários (PF) de até segunda ordem podem gerar uma grande variedade de curvas úteis para modelar as relações de interesse prático. Em princípio, os PF podem também resolver os problemas de falta de ajuste dos modelos de primeira e segunda ordem na análise de dados experimentais. No entanto, quando tentamos ajustar um PF aos dados de um experimento, esbarramos em, pelo menos, dois problemas. O primeiro é que o PF de segunda ordem, como definido originalmente, inclui dois parâmetros para cada fator, as potências, além dos coeficientes de regressão. O segundo é que o delineamento clássico apresenta pontos esparsos e simétricos na região experimental, resultando em pouca informação para estimação dos parâmetros extras do polinômio. Nesse trabalho propomos uma versão de PF de segunda ordem que restringe a estimação de uma única potência para cada fator. A ideia é que a potência determina a transformação apropriada aos níveis do fator para que o polinômio de segunda ordem seja uma boa aproximação para a relação subjacente. Sob esse modelo mais parcimonioso, estudamos o comportamento de delineamentos eficientes para estimar todos os parâmetros. Como o modelo é não linear precisamos incorporar informação a priori para a construção dos delineamentos. Resultados mostram que o delineamento resultante para o PF pode ser bem diferente do delineamento clássico, indicando que o problema de estimação das potências deve ser considerado no planejamento do experimento. O método pode ser estendido para os modelos lineares generalizados nas situações em que seja apropriado especificar o preditor linear por uma relação curva assimétrica.

Alternative methods for modeling of the cure rate in survival studies

Profa. Dra. Vera Lúcia Damasceno Tomazella (UFSCar)

In medical studies, it is common that some units under study are not susceptible to the event of interest, called immune or cured elements. A class of models, referred to as cure rate models, considers these situations and has been studied by several authors in the recent years. The cure fraction is of interest to patients and a useful measure to monitor trends and differences

in survival of curable disease. In this presentation we discuss some alternative methods for modeling cure rate in particular the Defective models. Defective models have the advantage of modeling the proportion of cured without adding any extra parameters in the model, in contrast to the most models from the literature.

3 COMUNICAÇÕES ORAIS

Seleção de formulações de bebidas prebióticas via Modelos Logitos Cumulativos Mistos

Authors

Resumo: Na área de Ciência e Tecnologia de Alimentos é comum estudos para o desenvolvimento de um produto, de acordo com a demanda, necessidades e exigências dos consumidores. As pesquisas nessa área podem envolver um elevado número de tratamentos (tipos de produtos ou marcas) com heterogeneidade, ou, ainda, limitações que restringem o tamanho dos blocos (provadores ou avaliadores). Um estudo realizado pela Universidade Federal do Ceará (UFC) desenvolveu bebidas prebióticas à base de amêndoa da castanha de caju adicionada ao suco de uva, para avaliar a impressão global utilizando uma escala hedônica estruturada de 9 pontos (variável politômica ordinal). A avaliação da aceitação sensorial das formulações desenvolvidas foram realizadas em diferentes sessões para cada tipo de bebida, com 130 julgadores não treinados. As amostras das bebidas foram servidas de forma monádica sequencial, seguindo um delineamento de blocos incompletos balanceados, em que cada provador avaliou 4 das 13 formulações de bebidas criadas. O presente trabalho apresenta uma estratégia para análise de dados categorizados ordinais. Inicialmente, utilizou-se da análise de correspondência como uma ferramenta exploratória, posteriormente foi verificado qual o modelo mais adequado e por fim foi ajustado o modelo logito cumulativo com chances proporcionais misto. Pode-se concluir que existe diferença entre a impressão global das 13 formulações de bebidas prebióticas à base de amêndoa da castanha de caju adicionadas de suco de uva. A formulação de bebida que teve melhor aceitação ao paladar dos provadores foram F4 e F6.

Selection of prebiotic beverage formulations via Mixed Cumulative Logit Models

Abstract: In the area of Food Science and Technology, studies are usually carried out for the development of a product, according to the demand, needs and demands of consumers. Research in this area may involve a large number of treatments (types of products or brands) with heterogeneity, or limitations that restrict the size of the blocks (testers or evaluators). A study conducted by the Federal University of Ceará (UFC) developed prebiotics based on cashew nut almonds added to grape juice to evaluate the overall impression using a structured hedonic scale of 9 points (ordinal polytomic variable). The evaluation of the sensorial acceptance of the developed formulations was performed in different sessions for each type of beverage, with 130 untrained judges. The beverage samples were served sequentially monadic, following a balanced incomplete block design, in which each taster evaluated 4 of the 13 beverage formulations created. The present work shows a strategy for analyzing ordinal categorized data. Initially, it was used the correspondence analysis as an exploratory tool, later it was verified which model was the most appropriate and finally the cumulative logit model with mixed proportional chances was adjusted. It can be concluded that there is a difference between the overall impression of the 13 prebiotic beverage formulations based on almond cashew nuts added with grape juice. The beverage formulation that had the best taste acceptance of the tasters was F4 e F6.

Imputação múltipla para matriz de interação genótipo e ambiente

Authors

Resumo: Nas pesquisas relacionadas à genética quantitativa, uma das maiores dificuldades é a caracterização de aspectos adaptativos que um determinado genótipo pode apresentar em diferentes ambientes, estes comportamentos são chamados de Interação Genótipo e Ambiente (IGA). Em virtude da necessidade de estimar estas características, esta linha de pesquisa utiliza-se de amplas abordagens estatística, como exemplo: modelos mistos, regressão, estatística experimental, métodos bayesianos, entre outros. Além das citas acima, uma metodologia conhecida é o modelo AMMI (Additive Main effects and Multiplicative Interaction), onde são realizadas as estimativas dos efeitos aditivos por meio da ANOVA (Analysis of Variance) para grupos de experimentos, e as avaliações dos efeitos multiplicativos a partir da DVS (Decomposição em Valores Singulares) a partir da matriz de interação. Entretanto um problema na utilização do modelo AMMI, é a ausência de determinadas respostas genotípicas que acontece por conta das adversidades naturais que o experimento pode apresentar, acarretando assim em problemas no cálculo da DVS. Uma das alternativas que pode ajudar a prosseguir na análise, é na utilização de técnicas de imputação para os valores ausente. Krzanowski (1988) propõe um método de imputação múltipla para a matriz de interação que não necessita de distribuição à sua priori. Neste trabalho foi realizado a análise de um experimento que contém 92 genótipos em 3 ambientes diferentes por meio de um delineamento inteiramente casualizado (DIC) com 3 repetições. Contudo, 6 genótipos apresentarão ausência de uma repetição, gerando um experimento desbalanceado. Desta forma, foi realizado a imputação múltipla na matriz dos efeitos multiplicativos e avaliados por meio de biplots em dois casos: sem os genótipos com repetições ausentes e com os valores imputados.

Multiple imputation for matrix genotype and environment interaction

Abstract: Research of the quantitative genetics, one of the greatest difficulties is the characterization of adaptive aspects that a certain genotype can present in different environments, these behaviors are called Genotype and Environment Interaction (GEI). Consequently with need of to estimate these characteristics, this line of research uses statistical approaches, such as mixed models, regression, experimental statistics, Bayesian methods, among others. In addition to the above, a known methodology is the AMMI (Additive Main effects and Multiplicative Interaction) model, where estimates of the additive effects through ANOVA (Analysis of Variance) for groups of experiments, and the evaluations of the multiplicative effects to from the SVD (Singular Value Decomposition) from the interaction matrix. However, a problem in the use of the AMMI model is the absence of certain genotype responses that happens due to the natural adversities that the experiment can present, thus causing problems in the calculation of the SVD. Alternatives that can help further the analysis is use of imputation techniques for missing values. Krzanowski (1988) proposes a method multiple imputation for the interaction matrix that does not need distribution. Finally this work review an experiment with 92 genotypes in 3 different environments by means of a completely randomized design (CRD) with 3 replicates. However, 6 genotypes will present miss repetition. In this way, the multiple imputation in the matrix of the multiplicative effects was performed and evaluated through biplot's in two cases: without the genotypes with miss repetitions and with the imputed values.

4 PARTICIPANTES

- Jhessica Letícia Kirch (ESALQ/USP)
- Roseli Aparecida Leandro (ESALQ/USP)

- Taciana Villela Savian (ESALQ/USP)
- Pollyane Vieira da Silva (ESALQ/USP)
- Janaína Marques e Melo (ESALQ/USP)
- Suelen Cristina Gasparetto (ESALQ/USP)
- Alessandra de Lima Goes (ESALQ/USP)
- Maria Letícia Salvador (ESALQ/USP)
- Cristiane Mariana Rodrigues da Silva (ESALQ/USP)
- Rita de Cássia de Lima Idalino (UNESP)
- Iuri Emmanuel de Paula Ferreira (UFSCAR - LAGOA DO SINO)
- Edijane Paredes Garcia (UNESP/BOTUCATU)
- Hercílio da Silva Freitas Júnior (ESALQ/USP)
- Fabrício Pilonetto (ESALQ/USP)
- Giovanni Casagrande Silvello (ESALQ/USP)
- Yuniel Tejeda Mazola (CENA/USP)
- Timóteo Herculino da Silva Barros (ESALQ/USP)
- Silvio Henrique Menezes Gomes (ESALQ/USP)
- Paula Midori Castelo (UNIFESP)
- Igor Engler Lima (ESALQ/USP)
- Cristian Marcelo Villegas Lobos (ESALQ/USP)
- Fábio Prata Vieira (ESALQ/USP)
- Mayara Aparecida Correa (ESALQ/USP)
- Humberto de Jesus Eufrade Junior (FCA/UNESP)
- Luana Tais de Freitas (ESALQ USP)
- Fagner Junior Gomes (ESALQ/USP)
- Welinton Yoshio Hirai (ESALQ/USP)
- Andréia Pereira Maria Hilário (ESALQ/USP)
- Alvaro Ramirez (ESALQ USP)
- Vivian Aparecida Brancaglioni (ESALQ/USP)
- Idemauro Antonio Rodrigues de Lara (ESALQ/USP)
- Maira Blumer Fatoletto (ESALQ/USP)
- Aldeir Ronaldo Silva (ESALQ/USP)
- Clarice Garcia Borges Demétrio (ESALQ/USP)
- Glória Cristina Vieira de Sousa (ESALQ/USP)
- Maria Beatriz Duarte Gavião (UNICAMP)
- Mariely Lopes dos Santos (UNIVERSIDADE DE SÃO PAULO ESCOLA SUPERIOR DE AGRICULTURA LUIZ DE QUEIROZ)
- Wagner Wolff (ESALQ/USP)
- Julio Cezar Souza Vasconcelos (ESALQ/USP)

- Roberto de Oliveira Rodrigues (ARYSTA (TAMBÉM ALUNO DE ENG. AGRONÔMICA NA FAESB))
- João Gabriel Ribeiro (ESALQ/USP)