

# 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 Junior<sup>1</sup> (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-ppgaea.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/>.

---

<sup>1</sup>E-mail: [jreduardo@usp.br](mailto:jreduardo@usp.br).



# VIII Encontro dos Alunos

## em Estatística e Experimentação Agronômica

<b>1 MINICURSO</b>	<b>1</b>
Modelos de Regressão Não Linear	
<i>Prof. Dr. Walmes Marques Zeviani . . . . .</i>	1
<b>2 CONFERÊNCIAS</b>	<b>1</b>
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>	2
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>	3
<b>3 COMUNICAÇÕES ORAIS</b>	<b>4</b>
Impacto de fungos no controle de lagartas <i>Chrysodeixis includens</i> e <i>Helicoverpa amigera</i>	
<i>Ribeiro Jr, E.E., Dalessandro, C.P. &amp; Demétrio, C.G.B. . . . .</i>	4
Seleção de formulações de bebidas prebióticas via Modelos Logitos Cumulativos Mistos	
<i>Melo, J.M., Lara, I.A.R. &amp; Freitas, S.M. . . . .</i>	4
Visualização de dados utilizando ggplot2	
<i>Gasparetto, S.C., Piedade, S.M.S. &amp; Villegas, C. . . . .</i>	6
Delineamentos ótimos para experimentos com cana-de-açúcar	
<i>Goes, A.L. &amp; Sermarini, R.A. . . . .</i>	6
Imputação múltipla para matriz de interação genótipo e ambiente	
<i>Hirai, W.Y. &amp; Dias, C.T.S. . . . .</i>	7
Ajuste do modelo Gompertz aos dados de crescimento de frangos de corte	
<i>Hilário, A.P.M. &amp; Lima, C.G. . . . .</i>	8
Seleção de efeitos do tratamento usando o método bootstrap	
<i>Fatoretto, M.B., Moral, R.A. &amp; Demétrio, C.G.B. . . . .</i>	9

<b>Superdispersão em Dados Categorizados Multinomiais: uma Aplicação em Ciências Agrárias</b>	
<i>Salvador, M.L. &amp; Lara, I.A.R.</i> . . . . .	10
<b>Análise do fracionamento de fósforo em solo de cerrado por meio de técnicas multivariadas</b>	
<i>Brancaglioni, V.A., Dias, C.T.S., Coelho, M.J.A. &amp; Pavinato, P.S.</i> . . . .	11
<b>A distribuição odd log-logística gama generalizada e modelo de regressão associado para dados censurados</b>	
<i>Pratavia, F., Cordeiro, G.M., Ortega, E.M.M. &amp; Braga, A.S.</i> . . . . .	13
<b>Abordagem geoestatística para mapeamento da chuva a partir de enlaces comerciais de microondas</b>	
<i>Wolff, W., Overeem, A. &amp; Uijlenhoet, R.</i> . . . . .	13
<b>4 PARTICIPANTES</b>	<b>15</b>



# 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

#### Impacto de fungos no controle de lagartas *Chrysodeixis includens* e *Helicoverpa amigera*

*Eduardo Elias Ribeiro Junior, Celeste P. Dalessandro & Clarice Garcia Borges Demétrio*

**Resumo:** *Chrysodeixis includens* e *Helicoverpa amigera* são espécies de lagartas da família *Noctuidae* que se alimentam de uma grande variedade de plantas, incluindo grandes culturas, sendo seu controle de grande interesse em Entomologia. Para avaliar o impacto de fungos no controle dessas espécies em plantas de soja, foi realizado um experimento com 25 lagartas alimentadas com plantas inteiras inoculadas com os fungos Me 1638, Bb 3399 e If 3422, além de plantas naturais (Controle). As lagartas foram acompanhadas de três em três dias durante 18 e 21 dias para as espécies *Chrysodeixis includens* e *Helicoverpa amigera*, respectivamente. Os pesos longitudinais e os tempos até a morte e até a fase adulta foram mensurados para cada lagarta. Modelos mistos não-lineares, de crescimento logístico, foram ajustados aos pesos longitudinais, enquanto que modelos de tempo de falha acelerados foram considerados para ostempos. Os resultados indicam, em geral, que o crescimento das lagartas é inibido pela inoculação dos fungos nas plantas, e que houve uma redução do tempo até a morte e um aumento do tempo até a fase adulta.

#### Impact of fungi on control of bollworms *Chrysodeixis includens* and *Helicoverpa amigera*

**Abstract:** *Chrysodeixis includens* and *Helicoverpa amigera* are species of bollworms from the *Noctuidae* family of which feed on a wide range of plants, including important cultivated crops. To study the impact of fungi on control of these species in the soybean crop, an experiment was carried out with 25 bollworms fed with whole plants inoculated with the fungi Me 1638, Bb 3399 and If 3422 or just the natural plant (Control). The bollworms were followed by 18 and 21 days for the *Chrysodeixis includens* and *Helicoverpa amigera*, respectively. The longitudinal weights and the times to death and to adult stage were measured for each animal. Nonlinear models of logistic growth were fitted to the longitudinal weights and accelerated failure time models were fitted to time-to-event data. The results showed, in general, that the growth is inhibited by the fungi inoculation in plants, and that there is a decrease in times to death and an increase in times to adult stage.

## Seleção de formulações de bebidas prebióticas via Modelos Logitos Cumulativos Mistos

*Janaína Marques e Melo, Idemauro Antonio Rodrigues de Lara & Sílvia Maria de Freitas*

**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.

## Visualização de dados utilizando ggplot2

*Suelen Cristina Gasparetto, Sônia Maria de Stefano Piedade & Cristian Villegas*

Resumo: O ggplot2 foi criado em 2005 por Hadley Wickham, é um pacote de visualização de dados para a linguagem de programação estatística R. A partir da definição de gráficos estabelecida em *The Grammar of Graphics* livro escrito por Leland Wilkinson em 2005, Hadley escreveu um artigo em 2010 com o seguinte título: *A Layered Grammar of Graphics*, mostrando que podem ser divididas em camadas as características principais de um gráfico, o que é a essência do ggplot2. Este pacote tem como vantagens: obter gráficos mais elegantes, facilidade na criação destes gráficos, estrutura padronizada das funções deixando assim o processo ainda mais intuitivo, além da possibilidade de criar uma grande quantidade de gráficos com poucas linhas de código. Serão apresentados diversos gráficos sendo eles: gráfico de pontos, barras, dispersão e boxplot, além de suas variações. Serão utilizados para ilustrar a metodologia, os dados diários de precipitação pluviométrica da estação convencional do posto meteorológico da ESALQ (<http://www.esalq.usp.br/departamentos/leb/anos.html>).

### Data visualization using ggplot2

Abstract: The ggplot2 was created in 2005 by Hadley Wickham, is a data visualization package for the statistical programming language R. From the definition of graphics established in *'The Grammar of Graphics'* book written by Leland Wilkinson in 2005, Hadley wrote a article in 2010 with the following title: *'A Layered Grammar of Graphics'*, showing that the main characteristics of a graphic can be divided into layers, which is the essence of ggplot2. This package has the advantages of: obtaining more elegant graphics, easy creation of these graphics, standardized structure of functions, making the process even more intuitive, and the possibility of creating a large number of graphics with few lines of code. There will be several graphs being shown: dotcharts, barcharts, scatterplot and boxplot, besides its variations. It will be used to illustrate the methodology, the daily rainfall data of the conventional station of the meteorological station of the ESALQ (<http://www.esalq.usp.br/departamentos/leb/anos.html>).

## **Delineamentos ótimos para experimentos com cana-de-açúcar**

*Alessandra de Lima Goes & Renata Alcarde Sermarini*

Resumo: Um dos objetivos do programa de melhoramento genético é a seleção de novos melhores clones. Com delineamentos ótimos e modelos bem ajustados consegue-se melhorar a precisão dessa seleção, porém em fases iniciais do programa de melhoramento genético se tem uma limitação de material genético o que dificulta tal seleção devido a falta de repetição. No presente trabalho avaliamos experimentos que apresentam parte dos novos clones repetidos, ou em alguns casos, apenas uma repetição de cada, como também o critério adotado para se obter o delineamento ótimo.

### **Optimal Designs with Sugarcane**

Abstract: One of the objectives of the breeding program is the best selection of new clones. With optimum designs and well-adjusted models, the accuracy of this selection can be improved, but in the early stages of the breeding program there is a limitation of genetic material which makes this selection difficult due to lack of repetition. In the present work we evaluated experiments that present part of the new clones repeated, or in some cases, only one repetition of each, as well as the criterion adopted to obtain the optimal design.

## **Imputação múltipla para matriz de interação genótipo e ambiente**

*Welinton Yoshio Hirai & Carlos Tadeu dos Santos Dias*

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 citadas 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 ausentes. Krzanowski (1988) propõe um método de imputação múltipla para a matriz de interação que não necessita da especificação da sua distribuição à priori. Neste trabalho, foi realizada 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 apresentaram ausência de uma repetição, gerando um experimento desbalanceado. Desta forma, foi realizada 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). Necessity 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 the additive effects are estimated using the ANOVA (Analysis of Variance) model for groups of experiments, and multiplicative effects are evaluated according to the SVD (Singular Value Decomposition) of 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, causing problems in the calculation of the SVD. Alternatives that can help further analysis is used 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 presented miss repetition. In this way, the multiple imputation in the matrix of the multiplicative effects was performed and evaluated using biplot's in two cases: without the genotypes with miss repetitions and with the imputed values.

## Ajuste do modelo Gompertz aos dados de crescimento de frangos de corte

*Andréia Pereira Maria Hilário & César Gonçalves de Lima*

Resumo: O Brasil está entre os maiores produtores e exportadores mundiais de carne de frango. Nesse contexto, existe grande interesse em pesquisas que visam melhorar ainda mais a eficiência na produção da carne de frango. Nesse trabalho ajustou-se o modelo Gompertz com efeitos aleatórios para descrever

a curva de crescimento de 1080 frangos da linhagem Ross, no período de 1 a 56 dias, os dados são provenientes de um experimento cujo delineamento experimental foi casualizado em blocos com estrutura de tratamento fatorial com seis tratamentos e seis repetições de 30 aves cada. Foram testadas diferentes estruturas de variâncias e covariâncias e realizada uma análise dos resíduos para verificar a qualidade do ajuste.

### Gompertz model adjustment to the growth data of broiler chickens

**Abstract:** Brazil is among the world's largest producers and exporters of chicken meat. In this context, there is great interest in research aimed at further improving efficiency in the production of chicken meat. In this work the Gompertz model was adjusted with random effects to describe the growth curve of 1080 broilers Ross, in the period from 1 to 56 days, the data come from an experimental design randomized in blocks with treatment structure factorial with six treatments and six replicates of 30 birds each. Different covariance structures were tested and residual analysis was performed to verify the goodness of fit.

## Seleção de efeitos do tratamento usando o método bootstrap

*Maíra Blumer Fatoretto, Rafael de Andrade Moral & Clarice Garcia Borges Demétrio*

**Resumo:** O fungo *Isaria fumosorosea* é comumente encontrado no solo e infectando várias espécies de artrópodes. Este fungo tem diferentes tipos de isolados que podem ser coletados de diferentes lugares ou insetos. Biopesticidas baseados em fungos têm a capacidade de infectar um grande número de pragas e permanecer no meio ambiente. Um experimento foi conduzido no Departamento de Entomologia e Acarologia da ESALQ para estudar a tolerância à radiação UV-B em 14 diferentes isolados de *Isaria fumosorosea* para selecionar isolados com alta tolerância à radiação UV-B para o desenvolvimento de um novo micopesticida. Este experimento foi conduzido em delineamento de blocos completamente casualizados, com 4 blocos nas mesmas condições, sendo que cada bloco se refere a uma instalação do experimento. Dentro de cada bloco, uma suspensão de conídios de 14 isolados foi preparada e utilizada para inoculação em cinco placas de Petri correspondentes a cada tempo de exposição (0, 2, 4, 6 e 8h). As 70 placas de petri foram colocados em câmara de madeira com quatro lâmpadas fluorescentes e lá permaneceram até o tempo de exposição correspondente. A resposta, nesse caso, é a proporção de conídios germinados e, quanto maior essa proporção, maior é a tolerância ao isolado. Esses dados proporcionais foram analisados usando diferentes modelos. Dados de proporção de experimentos dose-resposta são freqüentemente superdispersos, então vários modelos propostos na literatura que incorporam a superdispersão serão estu-



dados. Também discutimos como selecionar o melhor modelo para descrever os dados. Além disso, a similaridade entre tratamentos será verificada usando abordagens de intervalos de confiança de bootstrap.

### Selection of treatment effects using a bootstrap procedures

Abstract: The fungus *Isaria Fumosorosea* is commonly found in soil and infect several species of arthropod. This fungus has different types of isolates that may be collected from different places or insects. Fungi-based biopesticides have the capacity to infect a large number of pests and to remain in the environment. An experiment was conducted at the Departament of Entomology and Acarology of ESALQ to study the UV-B radiation tolerance in 14 different isolates of *Isaria Fumosorosea* to select isolates with high UV-B radiation tolerance for developing of a new mycopesticide. This experiment was conducted as a randomized complete block design with 4 blocks in the same conditions, given that, each block refers to an installation of the experiment. Within each block, one conidia suspension of 14 isolates was prepared and was used for inoculation in five Petri dishes corresponding to each exposure time (0, 2, 4, 6 and 8h). The 70 dishes were placed in a wooden chamber with four fluorescent lamps and remained there until their corresponding exposure time. The response in this case, is the proportion of germinated conidia, and, higher this proportion, the higher the isolate tolerance is. This proportional data will be analysed using different models. Proportion data from dose-response experiments are often overdispersed, then several models proposed in the literature that incorporate overdispersion will be studied. We also discuss how to select the best model to describe the data. Furthermore, the similarity between treatments will be verified using bootstrap confidence intervals approaches.

## Superdispersão em Dados Categorizados Multinomiais: uma Aplicação em Ciências Agrárias

*Maria Letícia Salvador & Idemauro Antonio Rodrigues de Lara*

Resumo: Variáveis politômicas são comuns em experimentos agrônômicos, apresentando natureza natural ou ordinal. Voigt (2013), desenvolveu um experimento em casa de vegetação, cujo objetivo foi avaliar o florescimento de plantas adultas de laranjeira da variedade x11, enxertadas sobre o limão Cravo e sobre o citrumelo Swingle. Em cada planta avaliou-se o número de ramos novos, classificados como sendo: terminal, lateral, sem flor e abortada. O Modelo dos Logitos Generalizados é uma classe de modelos que pode ser empregada para a análise desses dados. Uma das características deste modelo é a pressuposição de que a variância é uma função conhecida da média e, espera-se, que a variância observada esteja próxima da variância pressuposta pelo modelo assumido. Contudo, há casos em que, os dados são mais heterogêneos do que a variância especifi-

cada pelo modelo, fenômeno este conhecido como superdispersão. Verifica-se na literatura que há uma abordagem de textos sobre superdispersão aplicados a dados de contagem e proporção para variáveis dicotômicas. Assim sendo, este trabalho tem como objetivo caracterizar o problema da superdispersão com dados multinomiais, assim como selecionar um modelo adequado para acomodar de forma satisfatória a superdispersão presente nos dados. Neste trabalho são ajustados os modelos Multinomial e Dirichlet-Multinomial e os resultados comparados, por meio da deviance e do gráfico half-normal plot. Avaliou-se forte evidências da presença de superdispersão para o modelo Multinomial. Então, realizou-se um teste de hipótese baseado na razão de verossimilhança, indicando que o modelo Dirichlet-Multinomial é o mais adequado.

### Overdispersion in Categorized Multinomial Data: an Application in Agrarian Sciences

**Abstract:** Polytomic variables are common in agronomic experiments, presenting natural or ordinal nature. Voigt (2013), developed a greenhouse experiment, whose objective was to evaluate the flowering of adult plants of orange variety 'x11', grafted on the Rangpur Lime and on the Citrumelo Swingle. In each plant, the number of new branches was evaluated, being classified as: terminal, lateral, without flower and aborted. The Generalized Logits Model is a class of models that can be used to analyze the data. One of the characteristics of this model is the assumption that the variance is a known function of the mean. It is expected that the observed variance be close to the variance assumed by the assumed model. However, there are cases in which, the data are more heterogeneous than the variance specified by the model, this phenomenon is known as overdispersion. It is verified in the literature that there is an approach of overdispersion texts applied to counting and proportion data for dichotomous variables. Thus, this work aims to characterize the problem of overdispersion with multinomial data, as well as to select a suitable model to satisfactorily accommodate the overdispersion present in the data. In this work, the Multinomial and Dirichlet-Multinomial models were adjusted, and the results are compared, through deviance and the half-normal graph. There was a strong evidence in presence of overdispersion for the Multinomial model evaluated. Then, a hypothesis test was performed based on the probability ratio, indicating that the Dirichlet-Multinomial model is the most adequate.

### Análise do fracionamento de fósforo em solo de cerrado por meio de técnicas multivariadas

*Vivian Aparecida Brancaglioni, Carlos Tadeu dos Santos Dias, Marta Jordana Arruda Coelho & Paulo Sergio Pavinato*

**Resumo:** O bioma Cerrado é uma das últimas alternativas viáveis de área para

expansão da produção com alto potencial agrícola. No entanto seus solos são descritos como sendo altamente intemperizados, ácidos e com baixa quantidade de elementos essenciais para as plantas, como por exemplo o fósforo (P). Devido à baixa disponibilidade de P no solo e a baixa eficiência do uso de fertilizantes, os agricultores utilizam, muitas vezes, fertilizantes fosfatados além das exigências das plantas. Apenas 10-20% do P aplicado como fertilizante é absorvido pelas plantas durante o ciclo de aplicação, assim, o manejo da adubação deve favorecer a absorção e diminuição dos processos de fixação pelo solo e, conseqüentemente, o aumento do aproveitamento do P pelas plantas. Um experimento de longa duração que compõe parte do trabalho de tese, desenvolvida no Programa de Pós-Graduação em Solos e Nutrição de Plantas - ESALQ/USP, que envolveu sistema de manejo da adubação de manutenção de P, doses de P e interação com a fosfatagem. O experimento esteve localizado no município de Pedra Preta, Mato Grosso (16°37'23"S e 54°28'26"O) em área de rotação das culturas de algodão e soja. O objetivo deste trabalho é avaliar a dinâmica das frações (inorgânicas e orgânicas) de P no solo em função das doses e formas de aplicação de fosfato solúvel, buscando encontrar tratamentos que sejam caracterizados por apresentarem maiores valores de fósforo em fração lábil, por meio das técnicas multivariadas empregadas, como a análise de agrupamento, análise de componentes principais e análise discriminante. Pode-se concluir que os tratamentos foram bem caracterizados de acordo com as frações de fósforo, indicando que tratamentos com doses de fósforo no sulco de 60, 90 e 120 Kg ha<sup>-1</sup> resultam em maior disponibilidade de fósforo para as culturas.

### Analysis of phosphorus fractionation in cerrado soil through multivariate techniques

**Abstract:** The Cerrado biome is one of the last viable alternatives of an area for expansion of production with high agricultural potential, however, their soils are described as being highly weathered, acidic and with a low number of essential elements for plants, such as phosphorus (P). Due to the low availability of P in soil and the low efficiency of fertilizer use, farmers often use phosphate fertilizers beyond plant requirements. Only 10-20% of the P applied as fertilizer is absorbed by the plants during the application cycle, thus, the management of fertilization should favour the absorption and reduction of soil fixation processes and, consequently, the increased in the use of P by plants. The study was a long-term experiment that composes part of the thesis work, it was developed in the Graduate Program in Soils and Plant Nutrition ESALQ/USP, this work involved maintenance management system of P support, P doses and interaction with phosphating. The objective of this work is evaluate the dynamics of the fractions (inorganic and organic) of P in the soil as a function of the doses and forms of application of soluble phosphate, seeking to find treatments that are characterized by having higher values of phosphorus in labile fraction, through the multivariate techniques employed, such as Group Analysis, Princi-

pal Component Analysis and Discriminant Analysis. It can be concluded that the treatments were well characterized according to the phosphorus fractions, indicating that treatments with phosphorus doses in the groove 60, 90 and 120 Kg ha<sup>-1</sup> resulted in higher availability to the cultures.

## A distribuição odd log-logística gama generalizada e modelo de regressão associado para dados censurados

*Fábio Pratavia, Gauss M. Cordeiro, Edwin M. M. Ortega & Altemir da Silva Braga*

**Resumo:** Neste trabalho, propomos uma extensão da distribuição gama generalizada (GG) chamada de distribuição odd log-logística gama generalizada (OLLGG), que tem flexibilidade para acomodar várias formas da função de taxa de risco (hrf), tais como crescente, decrescente, em forma de U e unimodal, e que também é apropriado para discriminar entre alguns modelos probabilísticos alternativos, como as distribuições GG, Weibull e exponencial, comumente usados na análise do tempo de sobrevivência. Além disso, é comum, em situações práticas, que o tempo de sobrevivência esteja relacionado a covariáveis que ajudam a explicar a variabilidade deste. Portanto, outro objetivo deste trabalho é propor modelos de regressão, na forma de localização e escala, denominado modelo de regressão log-odd log-logística gama generalizada (LOLLGG), para dados censurados. Ilustramos a potencialidade do novo modelo de regressão por meio de um conjunto de dados reais do nível de tensão e finalizamos com algumas considerações finais.

The odd log-logistic generalized gamma distribution and its associated regression model for censored data

**Abstract:** In this paper, we propose an extension of the generalized gamma (GG) distribution called the odd log-logistic generalized gamma (OLLGG) distribution, which has flexibility to accommodate various forms of the hazard rate function (hrf), such as increasing, decreasing, U-shaped and unimodal, and that is also appropriate to discriminate between some alternative probabilistic models such as the GG, Weibull and exponential distributions, commonly used in analyzing survival time. Besides this, it is common in practical situations for the survival time to be related to covariables that help explain the variability this time. Therefore, another objective of this work is to propose regression models, in the form of location and scale, called the log-odd log-logistic generalized gamma (LOLLGG) regression model, with censored data. We illustrate the potentiality of the new regression model by means of a real data set of the voltage level and ends with some concluding remarks.

## Abordagem geoestatística para mapeamento da chuva a partir de enlaces comerciais de microondas

*Wagner Wolff, Aart Overeem & Remko Uijlenhoet*

**Resumo:** Nas últimas décadas, o monitoramento pluviométrico por meio de medidas da atenuação induzida pela chuva a partir de enlaces comerciais de microondas (ECM) tornou-se uma forma alternativa para estimar a quantidade de chuva para grandes áreas. Para este tipo de monitoramento, a chuva média é medida ao longo de um enlace entre duas antenas celulares (transmissor e receptor). Esta topologia de linha e a interação entre elas precisa ser considerada em termo de mapeamento da chuva. Por exemplo, nas seções onde os enlaces se interceptam amostras sobrepostas são observadas e a real distância entre elas não é representada diretamente assumindo os enlaces como pontos. Assim, uma abordagem geoestatística baseada em modelos permite estimar os parâmetros e obter a variabilidade espacial de forma direta e computacionalmente eficiente considerando a característica de linha, isto é, a informação é regularizada sobre uma linha. O objetivo deste trabalho é propor um método para mapeamento da chuva a partir de ECM capaz de lidar com a topologia da informação, estimativa dos parâmetros e predição em grandes áreas. O trabalho foi realizado na Holanda (35.500 km<sup>2</sup>). Tendo uma precisão de 1db para uma amostragem de 10Hz, os dados foram disponibilizados pela T-mobile NL para amostras de mínimo e máximo sinal recebido a cada 15 minutos. Ao todo, 12 dias de Junho a Setembro de 2011 com aproximadamente 2500 enlaces por dia foram empregados aqui. Dados de radar corrigidos por estações pluviométricas foram utilizados como referência. Os resultados destacaram uma baixa melhora comparada com abordagens padrão que consideram os enlaces como pontos. No entanto, a modelagem geoestatística proposta aqui é mais robusta and contorna algumas restrições relacionadas a topologia da informação.

### Geostatistical approach to Rainfall Mapping from Commercial Microwave Links

**Abstract:** In the last decade, rainfall monitoring by using rain-induced attenuation measurements from commercial microwave links (CML) became an alternative way to estimate rainfall for large areas. For this kind of monitoring, the mean rainfall is measured along a link between two cellular antennas (transmitter and receiver). This line topology and the interaction between them need to be considered for rainfall mapping. For instance, in the sections where the links intercept each other oversampled locations are present and the distance between them can not be represented directly by assuming the links as a point. Thus, a geostatistical approach based on models allows estimating parameters and capturing the spatial variability, straightforwardly and computationally efficient by considering the line feature, i.e., the line support in geostatistics. The aim of this work is to propose a method for rainfall mapping capable of dealing

with line support, real-time parameter estimation, and large-area prediction. The work was carried out in The Netherlands (35,500 km<sup>2</sup>). Having a precision of 1 db for 10 Hz sampling, the data were made available by T-Mobile NL under minimum and maximum received powers at 15 minutes intervals. Overall, 12 days from June to September 2011 with approximately 2.500 links per day were employed here. The gauge-adjusted radar rainfall data set aggregated at 15 minutes were used as reference values. The results evidenced a few improvement compared with the default approaches which consider the links as a points. However, the geostatistical modeling proposed here are more robust and contour some constrain related to the information topology.

## 4 PARTICIPANTES

- Aldeir Ronaldo Silva (ESALQ/USP)
- Ana Maria Fioravante (ESALQ USP)
- Andréia Pereira Maria Hilário (ESALQ/USP)
- Caroline Oliveira do Nascimento (ESALQ/USP)
- Clarice Garcia Borges Demétrio (ESALQ/USP)
- Cristiane Mariana Rodrigues da Silva (ESALQ/USP)
- Cristian Marcelo Villegas Lobos (ESALQ/USP)
- Eduardo Elias Ribeiro Junior (ESALQ/USP)
- Fábio Pratavia (ESALQ/USP)
- Glória Cristina Vieira de Sousa (ESALQ/USP)
- Hercílio da Silva Freitas Júnior (ESALQ/USP)
- Janaína Marques e Melo (ESALQ/USP)
- Jhenifer Caetano Veloso (UFPR)
- Jhessica Letícia Kirch (ESALQ/USP)
- João Gabriel Ribeiro (ESALQ/USP)
- Julio Cezar Souza Vasconcelos (ESALQ/USP)
- Laura Vicuña Torres de Paula (ESALQ/USP)
- Luana Tais de Freitas (ESALQ USP)
- Maira Blumer Fatochetto (ESALQ/USP)
- Maria Letícia Salvador (ESALQ/USP)
- Pollyane Vieira da Silva (ESALQ/USP)
- Pórtia Piscitelli Cavalcanti (ESALQ/USP)

- Roberto de Oliveira Rodrigues (ARYSTA | FAESB)
- Roseli Aparecida Leandro (ESALQ/USP)
- Sidcleide Barbosa de Sousa (ESALQ/USP)
- Suelen Cristina Gasparetto (ESALQ/USP)
- Taciana Villela Savian (ESALQ/USP)
- Talita Tanaka Fernandes (ESALQ/USP)
- Valdemiro Piedade Vigas (ESALQ USP)
- Vivian Aparecida Brancaglioni (ESALQ/USP)
- Wagner Wolff (ESALQ/USP)
- Welinton Yoshio Hirai (ESALQ/USP)