

Universidade Federal do Rio Grande do Sul
Instituto de Matemática
Departamento de Estatística



Anais

III SEMANÍSTICA

III Semana Acadêmica do Departamento de Estatística

da UFRGS

<http://www.ufrgs.br/semanistica>

Porto Alegre - 22 a 25 de outubro de 2013

Organização:



Promoção:



Apoio:



Conteúdo

1	Cartaz da III SEMANÍSTICA	4
2	Introdução	5
3	Agradecimentos	6
4	Comissão Organizadora Docente	7
5	Comissão Científica	7
6	Comissão Organizadora Discente	7
7	Apresentação	8
8	Programação	9
9	Conferências	11
10	Seções de Comunicações	15

1 Cartaz da III SEMANÍSTICA

UFRGS
UNIVERSIDADE FEDERAL
DO RIO GRANDE DO SUL

III SEMANÍSTICA E
ANO INTERNACIONAL DA ESTATÍSTICA 2013
22 A 25 DE OUTUBRO

INFORMAÇÕES E INSCRIÇÕES:
www.ufrgs.br/semanistica
semanistica@gmail.com

ORGANIZAÇÃO:

PROMOÇÃO:

APOIO:

2 Introdução

A III Semana Acadêmica da Estatística (III SEMANÍSTICA), realizada de 22 a 25 de outubro de 2013, no Instituto de Matemática, Campus do Vale da UFRGS, faz parte do Ano Internacional da Estatística 2013 (Statistics2013), o qual é um evento com mais de 1.950 instituições espalhadas ao redor do mundo (122 países) que celebrarão neste ano as contribuições da Estatística para a ciência.

O Ano Internacional da Estatística, comemorado durante o ano de 2013, é uma celebração internacional que unirá universidades, faculdades, sociedades profissionais, escolas, empresas, agências governamentais, institutos de pesquisa e outros diversos grupos com o objetivo de divulgar a importância da Estatística para a comunidade científica em geral, usuários de dados públicos e privados, mídia, agentes públicos, empresários, estudantes e o público em geral.

Como instituição participante do Statistics2013, o Departamento de Estatística da UFRGS promoverá durante este ano algumas atividades especiais visando divulgar a importância da Estatística para a população em geral. Essas atividades terão como foco sensibilizar o público a compreender a força e o impacto que as estatísticas têm sobre todos os aspectos da sociedade, bem como divulgar a Estatística como uma profissão, especialmente entre os jovens, além de promover a criatividade e o desenvolvimento de atividades nas áreas da Probabilidade e da Estatística.

Para maiores informações sobre outras atividades relacionadas consulte

- i) <http://hotta0.wix.com/aiebrasil> - Ano Internacional da Estatística Brasil 2013
- ii) www.statistics2013.org - The International Year of Statistics (Statistics2013)
- iii) www.ufrgs.br/semanistica - III SEMANÍSTICA (Semana Acadêmica da Estatística 2013)

3 Agradecimentos

A III SEMANÍSTICA - Semana Acadêmica do Departamento de Estatística da UFRGS e STATISTICS2013 - Ano Internacional da Estatística não teriam sido possíveis sem o apoio das seguintes agências financiadoras e instituições:

ABE - Associação Brasileira de Estatística

DEST-UFRGS - Departamento de Estatística da UFRGS

FAPERGS - Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul

IM-UFRGS - Instituto de Matemática da UFRGS

PROEXT-UFRGS - Pró-Reitoria de Extensão da UFRGS

PROPESQ-UFRGS - Pró-Reitoria de Pesquisa da UFRGS

FURG - Universidade Federal de Rio Grande

UFPEL - Universidade Federal de Pelotas

UFSM - Universidade Federal de Santa Maria

A Comissão Organizadora da III SEMANÍSTICA e do STATISTICS2013 agradece a colaboração de todos que se dedicaram anonimamente e sem interesses pessoais, em promover a integração entre alunos, professores e profissionais em estatística.

Comissão Organizadora

4 Comissão Organizadora Docente

Cleber Bisognin (Coordenador - Departamento de Estatística-UFRGS)

Guilherme Pumi (Departamento de Estatística-UFRGS)

Liane Werner (Departamento de Estatística-UFRGS)

Márcia Elisa Soares Echeveste (Departamento de Estatística-UFRGS)

Márcio Valk (Departamento de Estatística-UFRGS)

Markus Chagas Stein (Departamento de Estatística-UFRGS)

5 Comissão Científica

Cleber Bisognin (Coordenador - Departamento de Estatística-UFRGS)

Guilherme Pumi (Departamento de Estatística-UFRGS)

Liane Werner (Departamento de Estatística-UFRGS)

Márcia Elisa Soares Echeveste (Departamento de Estatística-UFRGS)

Márcio Valk (Departamento de Estatística-UFRGS)

6 Comissão Organizadora Discente

Angélica Segala (Curso de Estatística-UFRGS)

Bruna Martini Dalmoro (Curso de Estatística-UFRGS)

Camila Thaís Weber (Curso de Estatística-UFRGS)

Gabriel da Cunha (Curso de Estatística-UFRGS)

Gustavo Thomas (Curso de Estatística-UFRGS)

Raiane Padilha Silveira (Curso de Estatística-UFRGS)

Renata Fragoso Máximo (Curso de Estatística-UFRGS)

7 Apresentação

O programa da III SEMANÍSTICA - Semana Acadêmica do Departamento de Estatística da Universidade federal do Rio Grande do Sul engloba as seguintes atividades

- 9 Conferências envolvendo pesquisas realizadas em diversas áreas da Estatística proferidas por pesquisadores convidados de diversas Universidades do Rio Grande do Sul e do Brasil;
- 2 Mesas Redondas envolvendo temas em Educação Estatística e o Profissional da Estatística;
- 1 Sessão Temática relacionada ao tema História e perspectivas da Estatística na região Sul do Brasil;
- Comunicações orais e posters apresentadas pelos participantes do evento;

8 Programação

Horários	22/10/2013	23/10/2013	24/10/2013	25/10/2013
	Terça-Feira	Quarta-Feira	Quinta-Feira	Sexta-Feira
08:30-09:15	Abertura C1	Minicurso	Minicurso	Apresentação Oral
09:15-10:00				
10:00-10:30	Coffee-Break	Coffee-Break	Coffee-Break	Coffee-Break
10:30-11:15	MR 1	C6	C8	MR 2
11:15-12:00		C7		
12:00-14:00	Almoço	Almoço	Almoço	Encerramento
14:00-14:45	C2	Passeio de Barco Porto Alegre - Guaíba	C9	
14:45-15:30	C3		ST	
15:30-16:00	Coffee-Break			
16:00-16:45	C4		Sessão Pôster	
16:45-17:30	C5		Coquetel	

Minicurso: Estatística na pesquisa experimental - Base conceitual, planejamento e análise estatística de experimentos

Ministrantes:

Amauri de Almeida Machado - UFPEL

João Gilberto Corrêa da Silva - UFPEL

(MR1) Mesa Redonda 1: A Profissão de Estatística: A visão de alguns profissionais

Coordenação: Liane Werner - UFRGS

Participantes:

Marco Siqueira Campos - Siqueira Campos

Jaimar de Barros Monteiro – CEEE

Rafael Bernardini Santos – FEE

Roberto Kock – VONPAR

(MR2) Mesa Redonda 2 : A Educação da Estatística na Era da Informação e Conhecimento: mais Criatividade no Ensino

Coordenação: Márcia Echeveste – DEST-UFRGS

Participantes:

Dinara Westphalen Xavier Fernandez

Cíntia Paese Giacomello – UCS

Guilherme Luz Tortorella

Lori Viali – DEST-UFRGS/PUCRS

Gabriela Trindade Perry – FA - UFRGS

(ST) Sessão Temática: Visão retrospectiva, situação atual e perspectiva futura da área de Estatística na região Sul do RS

Participantes:

João Gilberto Corrêa da Silva - UFPEL

Paul Gerhard Kinas - IMEF-FURG

Conferências:

(C1) Conferência 1 - Adriano Mendonça Souza - DE-UFSM

Título: Utilização da Modelagem Estatística para Avaliação de Processos Produtivos

(C2) Conferência 2 - Adriana Neumann de Oliveira – DMPA-UFRGS

Título: Probabilidade: ferramenta para a compreensão de fenômenos físicos

(C3) Conferência 3 - Lorena Vicini – DE-UFSM

Título: Processos de contagem, estimando o número excessos de ozônio, para um limiar, na Cidade do México

(C4) Conferência 4 – Rodrigo Lambert – FAMAT-UFU

Título: A distribuição da função de primeiro retorno

(C5) Conferência 5 - Michel José Anzanello – PPGEP-UFRGS

Título: Seleção de variáveis para classificação de amostras de medicamentos

(C6) Conferência 6 - Márcio Diniz - IME-USP

Título: Estatística na área de pesquisa médica

(C7) Conferência 7 - Willian Silva Barros - DME – UFPEL

Título: Métodos Estatísticos na Seleção Genômica

(C8) Conferência 8 - Paul Gerhard Kinas - IMEF-FURG

Título: Estatística Bayesiana na Ecologia de Populações Selvagens

(C9) Conferência 9 - Fábio Prates Machado – IME – USP

Título: Modelos probabilísticos para evolução de espécies

9 Conferências

Conferência 1

Utilização da Modelagem Estatística para Avaliação de Processos Produtivos

Prof. Dr. Adriano Mendonça Souza
Departamento de Estatística - UFSM

Resumo

A presença de equipamentos sofisticados e automatizados aliado as exigências de produtos com alta qualidade faz com que a avaliação de um processo produtivo seja criteriosa. Sendo assim, tem-se como objetivo apresentar métodos de modelagem estatística que auxiliem na avaliação de processos produtivos. Aborda-se, portanto as metodologias de análise de componentes principais (ACP), de modelos autoregressivos integrados e de médias móveis (ARIMA) e de modelos autoregressivos condicionais a heteroscedasticidade (ARCH), com a utilização de gráficos de controle, e por fim mostra-se uma aplicação.

Conferência 2

Probabilidade: Ferramenta para a Compreensão de Fenômenos Físicos

Prof^ª. Dr^a. Adriana Neumann de Oliveira
Departamento de Matemática - UFRGS

Resumo

Por muito tempo o homem tentou quantificar as ideias de probabilidade, para entender por exemplo os jogos de azar, mas as descrições matemáticas só apareceram muito mais tarde, com o surgimento da teoria da probabilidade. Como outras teorias, ela é uma representação dos conceitos probabilísticos em termos formais, isto é, em termos que podem ser considerados separadamente de seus significados. Esses termos formais são manipulados pelas regras da matemática e da lógica, e quaisquer resultados são então interpretados ou traduzidos de volta ao domínio do problema. Graças a este desenvolvimento teórico, podemos, por exemplo, entender o comportamento coletivo de fluídos em movimento modelando-os através de sistemas de partículas interagentes. O estudo do comportamento coletivo da interação das partículas é feito através de procedimentos limites (técnicas probabilísticas), chamado de limite hidrodinâmico. Do limite hidrodinâmico obtêm-se leis (técnicas matemáticas) determinísticas, que quando traduzidas para o modelo em questão, caracterizam a evolução coletiva do fluído que está sendo estudado.

Conferência 3

Processos de Contagem, Estimando o Número Excessos de Ozônio, para um Limiar, na Cidade do México

Prof^ª. Dr^a. Lorena Vicini
Departamento de Estatística - UFSM

Resumo

Neste trabalho, foram propostos dois modelos probabilísticos para modelar o número acumulado de vezes em que um padrão de ar é violado por determinado poluente, num determinado intervalo de tempo. Para atingir este objetivo trabalhamos com modelos sob o enfoque bayesiano. A estimação dos parâmetros desses modelos é feita utilizando uma formulação Bayesiana baseada em métodos MCMC (Monte Carlo em Cadeias de Markov). Os resultados obtidos nos permitiram verificar o comportamento dos dados diários de poluição durante dezoito anos de observações (1990 a 2008), e com isso observar que as medidas tomadas para o controle da poluição na Cidade do México, estão sendo eficazes, ou seja, se os níveis de poluição diminuíram ao longo do período analisado. Foi possível observar também que à medida que aplicamos modelos mais complexos, com a inclusão de mais parâmetros, as estimativas obtidas tendem a ser melhores do que em relação aos modelos mais simples, isso pode ser observado pelo ajuste do modelo bem como pelo critério de comparação de modelos DIC.

Conferência 4

A distribuição da Função de Primeiro Retorno

Prof. Dr. Rodrigo Lambert
Faculdade de Matemática - UFU

Resumo

Na análise estatística da recorrência de Poincaré, é clássico considerarmos o primeiro instante em que uma realização do processo atinge algum conjunto mensurável de medida positiva (tempo de entrada ou hitting time).

Quando analisamos o ínfimo entre todos os tempos de entrada, considerando como condição inicial tal conjunto mensurável (ou seja: condição inicial = conjunto alvo), temos o *Tempo de primeiro retorno*, que denotaremos por T_n . É sabido que, para processos ergódicos com entropia positiva, que T_n/n converge para 1 quase certamente [2, 3]. Aqui, n representa o tamanho das seqüências.

No presente trabalho, estudamos as flutuações do tempo de retorno ($n - T_n$). Para uma medida produto com marginais identicamente distribuídas num alfabeto enumerável, provamos que essa quantidade converge em distribuição para uma v.a. não degenerada, e apresentamos limitantes para a velocidade dessa convergência.

Referências:

- (1) M. Abadi and R. Lambert, *The distribution of the short-return function*, Nonlinearity 26 (2009) 1143.
- (2) V. Afraimovich, J.-R. Chazottes and B. Saussol, *Point-wise dimensions for Poincar recurrence associated with maps and special flows*, Discrete Contin. Dyn. Syst. 9 (2003), no. 2, 263-280. (2003).
- (3) L B. Saussol, S. Troubetzkoy and S. Vaienti, *Recurrence, dimensions and Lyapunov exponents*, J. Stat. Phys. 106 (2002), 623-634.

Conferência 5

Seleção de variáveis para classificação de amostras de medicamentos

Prof. Dr. Michel José Anzanello
PPGEP-UFRGS

Resumo

Esta conversa abordará a aplicação de ferramentas estatísticas para classificação e agrupamento de amostras de medicamentos oriundos de apreensões policiais. A sistemática a ser apresentada integra a ferramenta multivariada Análise de Componentes Principais (ACP), utilizada para definição de importância das variáveis que descrevem uma amostra, com a técnica de classificação de amostras KNN (k-nearest neighbor). Tal procedimento pode ser estendido a aplicações cujo objetivo é agrupar medicamentos falsificados com mesma procedência.

Conferência 6

Estatística na área de pesquisa médica

Márcio Diniz
IME-USP

Resumo

Uma breve discussão sobre o papel da atuação do profissional de estatística na área de pesquisa médica e apresentação de alguns dos desafios enfrentados na aplicação de técnicas estatísticas a fim de fomentar a consciência da necessidade de aliar conhecimento teórico com prático.

Conferência 7

Métodos Estatísticos na Seleção Genômica

Prof. Dr. Willian Silva Barros
Departamento de Matemática e Estatística – UFPEL

Resumo

A Seleção Genômica veio unir muitos métodos estatísticos a Genética de Populações à Genética Quantitativa. Estes dois ramos com forte orientação Biométrica tradicionalmente caminharam em separado, seja no Melhoramento Genético de Plantas e Animais ou na Genética Humana. Atualmente, a estimação de componentes da variação genética e dos valores genéticos e a predição de fenótipos usa três fontes de dados: fenotípicos, genealógicos e genotípicos em locos marcadores moleculares em desequilíbrio de ligação com os genes de interesse. Ferramenta da Genética de populações plenamente dos métodos de estimativa atualmente empregados. Dessa forma, Genética de Populações, Genética Quantitativa, Genética Molecular e Estatística são demandados simultaneamente na análise genética dos caracteres de interesse. Esse texto aborda a nova Genética Quantitativa do terceiro milênio.

Conferência 8

Estatística Bayesiana na Ecologia de Populações Selvagens

Prof. Dr. Paul Gerhard Kinas

Instituto de Matemática, Estatística e Física - FURG

Resumo

Para conhecer e preservar populações selvagens, o conhecimento de sua abundância e distribuição são fundamentais. Marcação-recaptura e de transectas lineares são dois grupos importantes de métodos para isso. Porém, no ambiente marinho há dificuldades particularmente desafiadoras para implementá-los. Com auxílio de estudos de caso, ilustraremos formas de reduzir dificuldades nos métodos de transectas lineares, por meio do uso de modelos hierárquicos Bayesianos. Concluiremos destacando as razões pelas quais consideramos que a modelagem estatística com abordagem Bayesiana tenderá a crescer entre Ecólogos.

Conferência 9

Modelos probabilísticos para evolução de espécies

Prof. Dr. Fábio Machado

Instituto de Matemática e Estatística - USP

Resumo

Mutação e seleção são responsáveis pela evolução da vida no planeta terra. Como a essência de suas dinâmicas pode ser descrita por equações, o tema evolução é também encarado como uma teoria matemática. Nesta conferência vou apresentar abordagens e contribuições recentes que a teoria da probabilidade tem dado para os modelos de evolução a partir de tópicos bem conhecidos dos processos estocásticos. Em particular vou apresentar os modelos e alguns resultados dos seguintes trabalhos:

- (1) Guiol, Hervé; Machado, Fábio P.; Schinazi, Rinaldo *On a link between a species survival time in an evolution model and the Bessel distributions*. Braz. J. Probab. Stat. 27 (2013), no. 2, 201–209.
- (2) Cox, J. Theodore; Schinazi, Rinaldo B. *A branching process for virus survival*. J. Appl. Probab. 49 (2012), no. 3, 888–894.
- (3) Guiol, H.; Machado, F. P.; Schinazi, R. B. *A stochastic model of evolution*. Markov Process. Related Fields 17 (2011), no. 2, 253–258.
- (4) Liggett, Thomas M.; Schinazi, Rinaldo B. *A stochastic model for phylogenetic trees*. J. Appl. Probab. 46 (2009), no. 2, 601–607.

10 Seções de Comunicações

Comunicações Orais

Previsão de demanda intermitente na gestão de reposição de peças para aeronaves de asas rotativas

Cátia Michele Tondolo^{1 5}

Mônica Cristina Bogoni Savian²

Conrado Couto Lara Ferreira³

Luciane Flores Jacobi⁴

Resumo: Esse trabalho teve por objetivo determinar um modelo para prever a reposição de peças de aeronaves de asas rotativas utilizadas pelo Exército Brasileiro. Os dados utilizados nesse trabalho foram obtidos junto ao Depósito Especial Alfandegário subordinado à Diretoria de Material de Aviação do Exército, DEA/DMAvEX. Até o momento já se codificou 500 peças diferentes com pedidos realizados ou não por ano, das quais serão avaliadas as 39 com maior número de itens solicitados pelo DEA/DMAvEX, que correspondem as 7,8% mais pedidas. As análises foram realizadas utilizando a planilha eletrônica Excel e o software R (R Core Team, 2012). Foram estabelecidas quais as chances de cada uma das 39 peças serem solicitadas ou não no próximo ano.

Palavras-chave: *Previsão de demanda, Helicópteros, Matriz de transição, Cadeias de Markov.*

1 Introdução

Historicamente, no início da atividade aérea, problemas relacionados ao projeto das aeronaves e a insuficiência de procedimentos operacionais eram fatores que mais contribuía para a ocorrência de acidentes. Atualmente esses fatores permanecem como sendo possíveis causadores de acidentes aéreos, porém somados a outros como, por exemplo, a atitude do comandante da aeronave. Observando-se todo processo que envolve o voo: desde o projeto da aeronave e seus componentes, passando pelo controle de qualidade na fabricação de peças até ao grau de treinamento dos pilotos e tripulação bem como dos seus aspectos psicológicos para lidar com situações de emergência (Dédalo, 2009).

¹ UFSM - Universidade Federal de Santa Maria. E-mail: cacamt@gmail.com

² UFSM - Universidade Federal de Santa Maria. E-mail: monicabogoni@yahoo.com.br

³ UFSM - Universidade Federal de Santa Maria. E-mail: conrado.engmec@gamil.com

⁴ UFSM - Universidade Federal de Santa Maria. E-mail: lucianefj8@gmail.com

⁵ Agradecimento ao FIPE/UFSM pelo apoio financeiro.

No caso específico desse estudo, concentrar-se-ão esforços no gerenciamento de materiais de manutenção usados na Aviação do Exército Brasileiro. Um dos grandes problemas encontrado no controle de pedidos de peças de reposição é fazer uma previsão adequada, pois a maioria dessas peças tem como características uma demanda muito baixa, chegando a não haver a requisição das mesmas por anos. Porém, algumas peças são importadas e todos os pedidos realizados dependem de licitações, tendo como consequência meses de espera, afetando assim a manutenção dos Helicópteros do Exército Brasileiro.

Com base no exposto anteriormente, essa pesquisa tem como objetivo determinar um modelo para prever a reposição de peças de aeronaves de asas rotativas utilizadas pelo Exército Brasileiro. Para tanto, será determinada as probabilidades de fazer ou não um pedido com base na matriz de transição calculada.

2 Materiais e Métodos

Os dados utilizados nesse trabalho foram obtidos junto ao Depósito Especial Alfandegário subordinado à Diretoria de Material de Aviação do Exército. O DEA/DMAvEX tem uma disponibilidade de onze anos de dados históricos compreendendo o período de janeiro de 2002 a dezembro de 2012, totalizando 56270 pedidos. Até o momento já se codificou 500 peças diferentes com pedidos realizados ou não por ano.

Entre essas 500 serão avaliadas as 39 com maior número de itens solicitados pelo DEA/DMAvEX, que correspondem as 7,8% mais pedidas. Para cada uma das 39 peças foi simplesmente levado em conta se houve ou não pedido no ano e não o número de peças pedidas, pois apesar de serem as peças com mais itens pedidos ainda havia muitos zeros, ou seja, muitos anos que não foram realizadas solicitações da mesma.

As análises foram realizadas utilizando a planilha eletrônica Excel e o software R (R Core Team, 2012) onde foi construído um algoritmo para estabelecer a matriz de transição.

Para determinarmos a probabilidade da realização ou não de um pedido futuro para cada peça foi utilizado cadeias de Markov, pois conforme Gomes e Wanke (2008) dentre as vantagens de se utilizar cadeias de Markov na gestão de peças de reposição, destaca-se o fato de não ser necessária a obtenção de grandes séries históricas.

Além disso, para Smart, onde existe demanda intermitente, os benefícios de soluções computadorizadas para previsões são ilusórios. E existe uma boa razão para tanto. Os métodos tradicionais de previsão não funcionam bem porque a demanda intermitente normalmente contém uma grande porcentagem de valores iguais à zero.

Antes de definirmos processo Markoviano é necessário definir processos estocásticos, que são fenômenos que variam em algum grau, de forma imprevisível, à medida que o tempo passa. Para Alves, Menezes e Zimmermann (2006) é definido como uma coleção de variáveis aleatórias $\{X_t\}$

indexadas por um parâmetro t pertencente a um conjunto T , normalmente composto de números inteiros não negativos e X_t representa uma característica mensurável de interesse no tempo t .

Processos estocásticos, conforme Silva (2009) são de interesse por descreverem o comportamento de um sistema operando sobre algum período de tempo. A variável aleatória X_t representa o estado do sistema no tempo t . Portanto, pode-se afirmar que X_t é definida em um espaço denominado Espaço de Estados. Os processos estocásticos de interesse nesse estudo são os chamados processos Markovianos.

Um processo estocástico é dito ser um Processo Markoviano se o estado futuro depende apenas do estado presente e não dos estados passados. Este tipo de processo estocástico é também denominado de processo sem memória, uma vez que o passado é “esquecido” (desprezado) (Alves, Menezes e Zimmermann, 2006; Silva, 2009).

As probabilidades condicionais $P\{X_{t+1}=j | X_t=i\} = p_{ij}$ são denominadas probabilidades de transição e representam, portanto, a probabilidade do estado X_{t+1} ser j no instante $t+1$ dado que o estado X_t é igual a i no instante t .

Será utilizado, neste trabalho, o conceito da Cadeia de Markov para a implantação do método de previsão, considerando dois estados, nulo e não nulo, ou seja, se foi ou não realizado pedido, e as probabilidades de transição são determinadas pelos valores históricos das ocorrências de demanda.

A matriz de transição que será utilizada para estimar as probabilidades de transição para os dois estados, zero ou não zero, ao longo do tempo é definida como segue:

$$P^1 = \begin{vmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{vmatrix}$$

Onde:

P_{00} = Probabilidade de ocorrer outro zero caso tenha ocorrido um zero.

P_{01} = Probabilidade de ocorrer um não zero caso tenha ocorrido um zero.

P_{10} = Probabilidade de ocorrer um zero caso tenha ocorrido um não zero.

P_{11} = Probabilidade de ocorrer um não zero caso tenha ocorrido um não zero.

P^1 = Matriz de transição no primeiro período.

3 Resultados e Discussões

A partir dos dados históricos desde 2002, foi contado o número de demandas nulas que sucederam outras demandas nulas, quantas demandas não nulas sucederam demandas não nulas, quantos zeros sucederam demandas não nulas e o contrário, fornecendo a Tabela 1 (utilizando o ITEM 234 como exemplo). Ou seja, após uma demanda nula, houve em 50% das vezes outro zero e em 50% uma demanda diferente de zero. No caso de haver uma demanda não nula a ocorrência de um zero seguinte foi de 25% das vezes, o que quer dizer que houve maiores ocorrências de anos consecutivos que essa peça foi solicitada.

Tabela 1: Contagem de ocorrências

${}_0P_0$	1	0,50
${}_0P_1$	1	0,50
${}_1P_0$	2	0,25
${}_1P_1$	6	0,75

As matrizes de transição, e consequente modelo de Markov, foram construídas considerando a realização ou não do pedido como um estado, e então foi verificada a transição de pedidos a cada dado em sequência, de ano em ano.

As transições entre demandas realizadas ou não para determinada peça podem ser modeladas como uma cadeia homogênea de Markov, e as matrizes de transição de probabilidade são dadas pela Tabela 2.

Tabela 2: Código da peça, valores das probabilidades da matriz de transição e números de iterações necessárias para convergência.

Código da peça	P_{00}	P_{01}	P_{10}	P_{11}	Número de Iterações
231	0,6923	0,3077	0,6923	0,3077	3
211	0,5714	0,4285	0,5714	0,4285	4
262, 164, 163, 204, 146, 321	0,1818	0,8182	0,1818	0,8182	3
203	1,0000	0,0000	1,0000	0,0000	4
160	0,6667	0,3333	0,6667	0,3333	4
122	0,6000	0,4000	0,6000	0,4000	0
356	1,0000	0,0000	0,2300	0,7600	1
207	0,4616	0,5384	0,4616	0,5384	8
177	1,0000	0,0000	0,4375	0,5625	1
213	1,0000	0,0000	0,2099	0,7901	1
208, 154, 124, 198, 234, 197	0,3334	0,6666	0,3334	0,6666	6
127, 214, 192, 232	0,1000	0,9000	0,1000	0,9000	4
361	0,3000	0,7000	0,3000	0,7000	3
69, 229	0,5000	0,5000	0,5000	0,5000	0
225	0,3913	0,6087	0,3913	0,6087	4
242, 23	0,2000	0,8000	0,2000	0,8000	9
231	0,6923	0,3077	0,6923	0,3077	3
159, 233, 224, 235	0,7500	0,2500	0,7500	0,2500	7
96, 199	0,2727	0,7273	0,2727	0,7273	9
310	0,4706	0,5294	0,4706	0,5294	10

Considerando os dados da Tabela 2, pode-se ter a previsão de quais peças deverão ser pedidas no ano que vem. Vale lembrar que não estamos prevendo o número de itens que serão solicitados, somente se serão ou não realizados pedidos dessa peça.

Por exemplo, o item de código número 234 possui o seu *part number* (PN), número de identificação criado pelo fornecedor e cujo propósito é simplificar a referência a esta peça, igual a

23310CA015020 e trata de um contrapino, que terá 66,66% de chance de ser pedido no próximo ano dado que tenha sido pedido nesse ano e 33,34% de possibilidade de ser feita solicitação dessa peça para o próximo período dado que nesse ano não ocorreu solicitação de contrapino.

4 Conclusão

Através do presente estudo, buscou-se modelar a ocorrência ou não de pedidos de peças de reposição para aeronaves de asas rotativas utilizadas no Exército Brasileiro usando Cadeias de Markov. Foram estabelecidas quais as chances de cada peça ser solicitada ou não no próximo ano, sendo que o próximo passo será determinar, entre as peças que serão solicitadas, uma previsão para o número de itens que será pedido para cada uma das peças solicitadas.

Referências

- [1] ALVES, Antônio C. B.; MENEZES, Marco A. F. de; ZIMMERMANN, Francisco J. P. *Pesquisa operacional: Um texto essencial para Engenharias, Computação e Ciências Econômicas*. Goiânia, 2006.
- [2] DÉDALO. *Revista de Segurança de Voo da Aviação do Exército*. nº 12, nov. 2009.
- [3] GOMES, Antonio V. P.; WANKE, Peter. *Modelagem da gestão de estoques de peças de reposição através de cadeias de Markov*. *Gestão de Produção*, v. 15, p. 57 – 72, 2008.
- [4] R Core Team (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- [5] SILVA, Gerson L. C. da. *Modelo de estoque para peças de reposição sujeitas à demanda intermitente e lead time estocástico*. 2009. 75f. Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal de Minas Gerais, Belo Horizonte, 2009.
- [6] SMART, Charles N. *Previsão de demanda intermitente para planejamento de estoques: tecnologia inovadora e resultados precisos*. Disponível: http://www.bluston.com.br/solucoes/white_paper.pdf. Acesso em: 15 ago. 2013.

Aplicação da metodologia Box e Jenkins na modelagem e previsão da série das despesas do governo com o Programa Bolsa Família.

Afonso Valau de Lima Junior¹

Fernanda Rezer²

Viviane de Senna³

Adriano Mendonça Souza⁴

Roselaine Ruviaro Zanini⁵

Resumo: O artigo objetiva estudar as características da série temporal referente aos gastos do governo com o Programa Bolsa Família- PBF, por meio de uma amostra de 115 observações mensais contendo os gastos mensais do governo com o PBF, compreendendo o período de janeiro de 2004 até mês de julho de 2013. A metodologia empregada é a modelagem Box e Jenkins. Observou-se ao final do estudo que o melhor modelo para a previsão de despesas foi o modelo ARIMA (2,1,2).

Palavras-chave: *ARIMA, Bolsa Família, Previsão.*

1 Introdução

O governo federal do Brasil possui diversos programas voltados à prestação de assistência aos necessitados, os quais tiveram como ponto de partida a Constituição Federal de 1988, que regularizou o órgão assistência social e previu a criação de novas leis com o objetivo de garantir cidadania à população considerada miserável. A primeira Lei criada foi a Lei Orgânica de Assistência Social – LOAS nº 8.742/93, sancionada após anos de discussões e debates. Mais tarde, em 2004, foi criado o Ministério do Desenvolvimento Social e Combate à Fome – MDS.

Este Ministério tem a missão de promover a inclusão social, a assistência integral e uma renda mínima às famílias em situação de pobreza. Para isso, conta com o apoio de sete secretarias responsáveis por programas e políticas públicas distintas, dentre elas, a Secretaria Extraordinária para Superação da Extrema Pobreza – SESEP, que é responsável pelo Plano Brasil Sem Miséria. Este plano consiste em um conjunto de ações efetuadas para a criação de novos programas que gerem oportunidades para a população mais pobre e assim elevar a renda e, conseqüentemente, as condições de vida da população, sendo dividido em três eixos e, no eixo 1, garantia de renda, pode ser encontrado o Programa Bolsa Família – PBF.

¹ UFSM - Pós-graduando de Especialização em Estatística. Email: avljunior@yahoo.com.br

² UFSM - Pós-graduando de Especialização em Estatística. Email: ferrezer25@hotmail.com

³ UFSM - Mestranda de Engenharia da Produção. Email: vivianedsenna@hotmail.com

⁴ UFSM - Orientador. Professor do departamento de Estatística. Email: amsouza@smail.ufsm.br

⁵ UFSM - Orientador. Professora do departamento de Estatística. Email: rzanini@smail.ufsm.br

O PBF é um dos mais conhecidos pela população e tem como função a transferência de renda básica para os cidadãos. Foi criado por meio da Lei nº 10.836 de 09/01/2004 e é regulamentado pelo Decreto 5.209 de 17/09/2004. Origina-se da união dos programas Bolsa Escola, Bolsa Alimentação, Auxílio Gás e Cartão Alimentação. Os quatro programas coexistiram de 2001 a 2004 e cada um era ofertado com critério próprio e era comum ocorrer casos nos quais uma família era contemplada por todos os e, em outras situações similares, por nenhum. A partir de 2004, com a criação do PBF e a criação de um Cadastro Único de famílias pobres, pelo Decreto nº 3.877/01, essa disparidade foi sendo eliminada.

Os valores em reais aplicados pelo governo desde a sua concepção são disponibilizados no site do MDS no formato de séries temporais com periodicidade mensal. Devido a isso, este estudo tem o objetivo descrever e verificar o comportamento da série, obtendo a compreensão necessária para a estimação dos valores para os períodos futuros, por meio da metodologia Box Jenkins.

2 Metodologia

Os dados utilizados, foram modelados com o auxílio do software Estatística 9.0, na modelagem foram retirados do site (www.mds.gov.br) da Secretaria de Avaliação e Gestão da Informação do MDS. A série contém os gastos mensais do governo com o PBF desde seu início em janeiro de 2004 até o mês de julho de 2013, totalizando 115 observações. Os valores aplicados também estão discriminados pelas cinco regiões do país.

A modelagem de Box Jenkins (1994) busca a seleção do modelo mais adequado para a aplicação do ciclo iterativo à série (SOUZA, CAMARGO, 1996). O ciclo possui duas ideias básicas, a parcimônia e a iteração entre fases. A parcimônia nada mais é do que o uso do menor número de parâmetros para estabelecer o modelo, e a iteração é repetir a ação quando necessário. As fases do ciclo, de acordo com Box, Jenkins e Reinsel (1994) são: a escolha da classe do modelo; a identificação do modelo; a estimação dos parâmetros; o diagnóstico de adequação do modelo; e o uso do modelo para a previsão e controle.

Para identificar o modelo pode-se usar a função de autocorrelação – FAC e a função de autocorrelação parcial – FACP. A modelagem Box Jenkins é o ajuste dos modelos autorregressivos e de médias móveis $ARIMA(p,q,d)$ em um conjunto de dados. Sua vantagem é a parcimônia do método e a precisão das previsões. A desvantagem da aplicação é a necessidade de conhecimento da técnica e o uso do software específico.

Para efetuar um $ARIMA(p,d,q)$ são necessárias 50 observações de uma série temporal y_t , no mínimo, com mesmo intervalo de tempo e a série deve ser estacionária. E estacionariedade se dá a medida que a média, a variância e a covariância são constantes ao longo do período. Para séries sem estacionariedade, a aplicação do processo de diferenciação pode eliminar a não linearidade.

A diferenciação é a subtração da segunda observação pela primeira demonstrado na fórmula $\Delta^d y = y_t - y_{t-1}$, o valor diferenciado é o valor de y no período $t - 1$, representado por d . A segunda diferença é dada por $\Delta^d y = y_{t-1} - y_{t-2}$, $d = 2$. A série é estacionária quando está estável ao longo do tempo e o modelo estimado se torna representativo (SOUZA, CAMARGO, 1996). Essa diferença indica a parte $I(d)$ do modelo ARIMA sua aplicação não deve ser superior a duas vezes na mesma série.

A parte $AR(p)$ autorregressiva de ordem p é explicada por $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$, ϕ indica os parâmetros reais e ε_t o ruído branco. O processo de médias móveis $MA(q)$ tem como equação $y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$. O θ indica constantes reais e ε_t o ruído branco. O ruído branco é um conjunto de variáveis aleatórias com distribuição e variância constantes e média igual a zero.

A equação na íntegra do modelo $ARIMA(p,q,d)$ é dada por $\Delta^d y = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$ (MORETTIN, 2008). Encontrados os modelos para a ordem (p,q,d) passa-se à estimação. Os critérios a serem obedecidos são, o nível de significância inferior a 5%, a presença de ruído branco, Akaike (AIC – Akaike Information Criteria) e Bayes (BIC – Bayesian Information Criteria), os dois últimos são para verificar o modelo aceitável através das equações $AIC = \ln \sigma_e^2 + (2(p + q))/n$ e $BIC = \ln \sigma_e^2 + ((p + q) \ln n)/n$; p e q são os parâmetros conhecidos, n é o tamanho da amostra, \ln é o logaritmo neperiano e σ_e^2 a variância estimada dos erros.

Quanto menor for o resultado do cálculo de AIC e BIC mais adequado será o modelo para a projeção dos valores futuros da série.

3 Resultados e Discussões

O gráfico 1 representa os dados originais das séries analisadas.

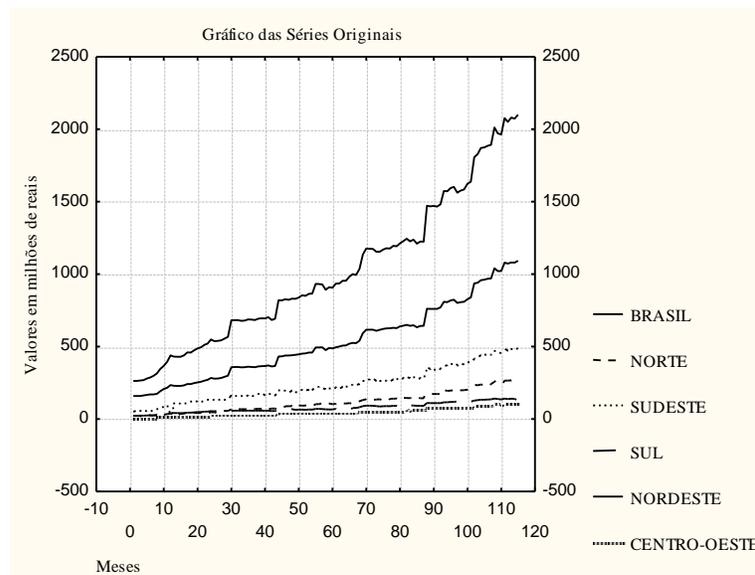


Gráfico 1: Séries originais

Ao efetuar o estudo sobre a série que representa o somatório total de gastos do governo com o PBF, o modelo que se obteve foi ARIMA(2,1,2). Contudo, ao separar a série por regiões do Brasil, devido a diferente aplicação dos recursos, os parâmetros são distintos conforme pode ser observado na Tabela 1:

Tabela 1: Resultados das modelagens aplicadas ao PBF por região do Brasil

PBF	Modelo	Coefficientes	p -value
Brasil	ARIMA(2,1,2)	$\phi_1 = 0,414991$	<0,001
		$\phi_2 = 0,584915$	
		$\theta_1 = 0,473183$	
		$\theta_2 = 0,480542$	
Região Norte	ARIMA(1,1,1)	$\phi_1 = 0,999859$ $\theta_1 = 0,845891$	<0,001
Região Sudeste	ARIMA(0,1,1)	$\theta_1 = 0,274483$	<0,001
Região Sul	ARIMA(2,2,0)	$\phi_1 = 0,425253$ $\phi_2 = -0,602462$	<0,001
Região Nordeste	ARIMA(0,2,2)	$\theta_1 = -0,683379$ $\theta_2 = 0,316620$	<0,001
Região Centro-oeste	ARIMA(2,2,0)	$\phi_1 = 0,452592$ $\phi_2 = -0,613425$	<0,001

Estes resultados foram obtidos após a análise do gráfico de cada série, nos quais pode ser constatada a tendência ascendente e a necessidade da utilização da diferenciação para torná-las estacionárias. Na Tabela 2 estão dispostas as previsões de gastos para o próximo semestre segundo a respectiva série:

Tabela 2: Previsão de despesas do governo com o PBF (em milhões de reais)

PBF	Ano/mês	Previsão	Intervalo de confiança	
			Inferior	Superior
Brasil	2013/08	2118,410	2052,425	2184,395
	2013/09	2140,859	2050,217	2231,501
	2013/10	2160,608	2047,630	2273,585
	2013/11	2181,934	2050,384	2313,483
	2013/12	2202,335	2053,101	2351,568
	2014/01	2223,275	2057,705	2388,845
Região Norte	2013/08	275,4084	266,8089	284,0078
	2013/09	278,7996	265,6684	291,9307
	2013/10	282,1903	264,9007	299,4799
	2013/11	285,5806	264,2039	306,9573
	2013/12	288,9704	263,4748	314,4659
	2014/01	292,3597	262,6678	322,0515
Região Sudeste	2013/08	487,3842	462,3347	512,4337
	2013/09	487,3842	456,4364	518,3319
	2013/10	487,3842	451,4948	523,2736
	2013/11	487,3842	447,1556	527,6128
	2013/12	487,3842	443,2409	531,5274
	2014/01	487,3842	439,6462	535,1222
Região Sul	2013/08	142,5734	134,4737	150,6731
	2013/09	141,5424	132,7408	150,3441
	2013/10	144,7884	129,2671	160,3096
	2013/11	141,7091	125,8280	157,5901
	2013/12	147,3850	123,3608	171,4092
	2014/01	142,0543	117,3768	166,7317
Região Nordeste	2013/08	1092,870	1045,337	1140,402
	2013/09	1102,250	1044,679	1159,821
	2013/10	1106,122	1007,548	1204,696
	2013/11	1112,148	994,091	1230,204
	2013/12	1119,374	956,300	1282,448
	2014/01	1122,045	932,074	1312,017
Região Centro-Oeste	2013/08	106,6217	101,0720	112,1715
	2013/09	106,4674	100,3757	112,5591
	2013/10	108,8557	98,1268	119,5847

2013/11	107,3501	96,3435	118,3567
2013/12	111,2683	94,7304	127,8062
2014/01	108,2900	91,2564	125,3236

4 Conclusão

É possível concluir que, para o governo, é muito importante a previsão dos valores que serão gastos para a manutenção dos programas como o Bolsa Família. De posse deste conhecimento é possível programar intervenções como ampliação, redução ou até mesmo exclusão do assistencialismo de renda garantida, ou seja, as previsões podem facilitar a conduta dos governos e subsidiar a criação de políticas condizentes com a realidade financeira do governo.

Referências

- [1] BOX, G.E.; JENKINS, G.M.; REINSEL, G.C. Time series analysis: Forecasting and control. 3 ed. New Jersey: Printice Hall, 1994.
- [2] BRASIL. Constituições Brasileiras: 1988. vol. VII. Brasília: Senado Federal e Ministério da Ciência e Tecnologia, Centro de Estudos Estratégicos, 2003.
- [3] _____. Ministério do Desenvolvimento Social e Combate à Fome. Histórico dos benefícios eventuais. Brasília, 2008. Disponível em: <<http://www.mds.gov.br>>. Acesso em: 28 jul. 2013.
- [4] _____. Presidência da República. Lei Orgânica da Assistência Social, Lei nº 8.742, de 7 de dezembro de 1993, publicada no DOU de 8 de dezembro de 1993.
- [5] MORETTIN, Pedro A.; TOLOI Clélia M.. Métodos quantitativos: séries temporais. São Paulo: Atual, 1986.
- [6] MORETTIN, Pedro A.. Econometria financeira: um curso de séries temporais financeiras. São Paulo: Blucher, 2008.
- [7] SOUZA, Reinaldo Castro; CAMARGO, Maria Emilia. Análise e previsão de séries temporais: os modelos ARIMA. Ijuí: Sedigraf, 1996.

Análise econômica na implantação e na gestão de uma empresa rural: estudo de caso

Fernanda Rezer de Menezes¹

Afonso Valau de Lima Junior²

Daniel Lemes³

Viviane de Senna⁴

Alfredo Skrebsky Cezar⁵

Paulo Santa Pacheco⁶

Resumo: O mercado da ovinocultura esta cada vez mais exigente, requerendo do produtor maiores investimentos. O estudo mostra a estrutura de custos, receitas e as medidas de resultados daí decorrentes obtidas a partir do fluxo de caixa. O novo empreendimento foi implantado em uma área de 2 (dois) hectares no interior do Rio Grande do Sul, com o objetivo de avaliar a viabilidade econômica de um sistema semi-intensivo de criação de ovinos de corte em ciclo completo voltado para a produção de carne de cordeiros. Para a análise de viabilidade econômica foi considerado o valor presente líquido e para avaliar o risco considerou-se o Pay-Back descontado (Período de recuperação do investimento). O VPL (Valor Presente Líquido) calculado, considerando uma Taxa Mínima de Atratividade (TMA) de 8,75% ao ano foi de R\$ 21.774,09 e o retorno do investimento através do Payback descontado ocorre no quarto ano. Portanto, o empreendimento em estudo é economicamente viável.

Palavras-chave: *ovinocultura, Valor Presente Líquido (VPL), custos, Pay-Back descontado.*

1 Instruções

No século XX, a criação de ovinos nas empresas rurais existia em virtude do preço da lã e o mercado internacional para sua comercialização estava em expansão. Os produtores de ovinos pouco comercializavam a carne ovina e seus subprodutos e tinham seus animais apenas para consumo próprio (VIANA; SOUZA, 2007). Nos anos 1980, as famílias brasileiras passam a ter seu poder aquisitivo mais alto e começam a criar o hábito de consumir cada vez mais a carne ovina para a qual a demanda passa a aumentar. A partir daí, surge o interesse por raças ovinas com aptidão para a produção de carne precoce (carne de cordeiro) e a lã passa a ser um produto secundário (BOFILL, 1996).

Com todas as exigências do mercado atual o custo para a produção de ovinos tem se elevado, pois, aumenta os custos com alimentação, mão de obra, sanidade, depreciação.

Para avaliar a viabilidade de um empreendimento é necessário calcular os indicadores de investimento que são divididos em dois grupos: os indicadores de retorno como o VPL (Valor

¹ UFSM - Pós-graduando de Especialização em Estatística. Email: ferrezer25@hotmail.com

² UFSM - Pós-graduando de Especialização em Estatística. Email: avljunior@yahoo.com.br

³ UFSM - Graduando do Curso de Zootecnia. Email: dlemes@gmail.com

⁴ UFSM - Mestranda de Engenharia da Produção. Email: vivianedsenna@hotmail.com

⁵ UFSM - Doutorando em Medicina Veterinária. Email: alfredosps@hotmail.com

⁶ UFSM - Orientador. Professor do Departamento de Zootecnia. Email: pacheco.dz.ufsm@hotmail.com

Presente Líquido, Taxa Interna de Retorno (TIR), Índice Benefício/Custo, Retorno Adicional sobre o Investimento (ROIA). e os indicadores de risco que são os Pay-Back, Taxa Interna de Retorno (TIR) e Ponto de Fischert. Sendo fundamental para a tomada de decisão de investir no empreendimento o tempo de retorno e o grau de risco associado a ele (LAPPONI, 2007).

Objetivou-se com esse estudo avaliar a viabilidade econômica da implantação de um sistema semi-intensivo de criação de ovinos de corte em ciclo completo voltado para a produção de carne de cordeiros.

2 Metodologia

A empresa rural da qual a área de campo foi arrendada para a implantação do projeto-alvo deste estudo existe a mais de 80 anos, dedica-se à produção extensiva de bovinos de corte prioritariamente e em relação à agricultura, apenas a produção de arroz irrigado.

A criação semi-intensiva de ovinos de corte foi projetada para ser executada de maneira independente à rotina e às atividades econômicas já em curso na empresa rural, constituindo-se em uma nova empresa, sem vínculo com a propriedade rural pré-estabelecida, exceto pelo empréstimo do maquinário agrícola para a implantação de pastagens cultivadas.

Este novo empreendimento foi implantado inicialmente em uma área de aproximadamente 2 (dois) hectares que não estava sendo utilizada. O investimento foi projetado por um novo produtor rural, médico veterinário, e pertencente à mesma família, porém, sem influência na administração da fazenda. Inicialmente, o rebanho foi composto por 31 (trinta e uma) animais, sendo 30(trinta) fêmeas adultas (matrizes) e 1 (um) macho (carneiro) no primeiro ano e aquisição de mais 15 animais no segundo ano.

Para classificação dos itens de custos, receitas e cálculos dos indicadores financeiros, foram utilizadas metodologias propostas por Souza; Clemente (2009). Os custos de oportunidade foram calculados considerando a taxa mínima de atratividade (TMA) de 8,75% a.a., baseada no valor dos encargos a financiamentos a pequenos e médios produtores, atualmente utilizada no mercado.

Para se chegar aos indicadores econômicos foram confeccionados os fluxos de caixa através dos custos e receitas coletados, sendo possível avaliar a rentabilidade do empreendimento utilizando o Valor Presente Líquido (VPL), pois, consideram o valor do dinheiro no tempo e o risco será contatado pelo Payback descontado, que consiste em identificar o prazo em que o montante do dispêndio de capital será recuperado através dos fluxos líquidos de caixa gerados pelo investimento (LAPPONI, 2007). As análises foram efetuadas por meio da planilha Excel.

¹ UFSM - Pós-graduando de Especialização em Estatística. Email: ferrezer25@hotmail.com

² UFSM - Pós-graduando de Especialização em Estatística. Email: avljunior@yahoo.com.br

³ UFSM - Graduando do Curso de Zootecnia. Email: dlemes@gmail.com

⁴ UFSM - Mestranda de Engenharia da Produção. Email: vivianedsenna@hotmail.com

⁵ UFSM - Doutorando em Medicina Veterinária. Email: alfredosps@hotmail.com

⁶ UFSM - Orientador. Professor do Departamento de Zootecnia. Email: pacheco.dz.ufsm@hotmail.com

Tabela 1 - Descrição do indicador econômico.

Indicador	Unidade	Equação
Valor presente líquido (VPL)	R\$	$= \sum_{i=1}^n \frac{\text{valores}_i}{(1+TMA)^i}$, onde n=número de fluxos de caixa e TMA=taxa mínima de atratividade
Período de recuperação do investimento (<i>Payback descontado</i>) descontado (PBd)	Mês	$= ((-1 * \text{valor presente dos fluxos de caixa negativos}) / \text{valor presente dos fluxos de caixa positivos}) * \text{número de fluxos de caixa}$

Fonte: SOUZA; CLEMENTE (2009).

3 Resultados e Discussões

Os custos com investimento e manutenção do rebanho para o período analisado estão apresentados na Tabela 2, os custos com aquisição de animais e mão de obra obtido no investimento são muito similares aos custos encontrados por (LIMA; ARAUJO; MACIEL, 2007 e VIANA; SILVEIRA, 2008).

Para Gouveia; Haddad e Ribeiro (2006) o item que mais influenciou na composição do custo operacional total foi o arrendamento e no referido estudo este item não foi tão impactante, pois o empreendimento foi realizado numa área de 2 (dois) hectares. Decorrente da alta concentração de animais por hectare um dos maiores desafios do produtor ao trabalhar com a criação semi-intensiva de ovinos foi controlar a verminose. Para Cezar (2010) as endoparasitoses passaram a exercer um papel central dentre as causas de prejuízos econômicos nas criações pecuárias, podendo ser consideradas o principal fator restritivo à viabilidade dos sistemas de produção intensiva ou semi-intensiva de pequenos ruminantes em diversas regiões do mundo.

De posse destes índices, pode-se afirmar que o investimento é economicamente viável para o produtor, obtendo um VPL de R\$ 21.774,09 e o período de retorno do investimento (*payback descontado*) acontece a partir do quarto ano. Para Barro et al. (2009) que obteve os animais em um sistema de confinamento para terminação dos cordeiros com dieta à base de feno de alfafa e concentrado, elevou o custo da alimentação, apresentou um VPL abaixo do espera inviabilizando o empreendimento.

¹ UFSM - Pós-graduando de Especialização em Estatística. Email: ferrezer25@hotmail.com

² UFSM - Pós-graduando de Especialização em Estatística. Email: avljunior@yahoo.com.br

³ UFSM - Graduando do Curso de Zootecnia. Email: dlemes@gmail.com

⁴ UFSM - Mestranda de Engenharia da Produção. Email: vivianedsenna@hotmail.com

⁵ UFSM - Doutorando em Medicina Veterinária. Email: alfredosps@hotmail.com

⁶ UFSM - Orientador. Professor do Departamento de Zootecnia. Email: pacheco.dz.ufsm@hotmail.com

Tabela 2 - Fluxo de caixa detalhado (valores em R\$)

Item	Períodos			
	t0	t1	t2...9	t10
Investimento	-14.462,10	0	0	5.784,84
Aquisição de animais	- 4.485,00	0	0	1.794,00
Construção (cercas e galpão)	- 3.589,60	0	0	1.435,84
Mão de obra (construções: cerca e galpão)	- 6.387,50	0	0	2.555,00
Custo Operacional Total (COT)	0	- 7054,98	- 7054,98	- 7054,98
Viagens (combustível)	0	- 743,00	- 743,00	- 743,00
Despesas Administrativas	0	- 191,30	- 191,30	- 191,30
Sanidade	0	- 452,03	- 452,03	- 452,03
Tosa	0	- 93,00	- 93,00	- 93,00
Arrendamento	0	- 370,00	- 370,00	- 370,00
Alimentação (pastagens)	0	- 922,20	- 922,20	- 922,20
Custo de matrizes de reposição	0	-1.479,33	-1.479,33	-1.479,33
Mão de obra (diarista)	0	- 120,00	- 120,00	- 120,00
Reprodução (sincronização de cio)	0	- 510,00	- 510,00	- 510,00
Depreciação	0	- 1.664,12	- 1.664,12	- 1.664,12
Remuneração do empresário	0	- 510,00	- 510,00	- 510,00
Renda Bruta	0	12.254,00	12.254,00	12.254,00
Carne	0	10.400,00	10.400,00	10.400,00
Lã	0	450,00	450,00	450,00
Venda matrizes	0	1.404,00	1.404,00	1.404,00
Resultado líquido	-14.462,10	5.199,02	5.199,02	10.983,86

Fontes: Resultados do trabalho.

Legenda: t0 – Tempo inicial, t1 – Tempo após um ano, t2...9 – Tempo de 2 a 9 anos e t10 – Tempo de 10 anos.

4 Conclusão

Verificou-se que a criação de ovinos de corte em um sistema semi-intensivo com ciclo completo voltado para a produção de carne de cordeiros é viável, apresentando um VPL de R\$ 21.774,09. O período de retorno do empreendimento em estudo ocorrerá no quarto ano.

¹ UFSM - Pós-graduando de Especialização em Estatística. Email: ferrezer25@hotmail.com

² UFSM - Pós-graduando de Especialização em Estatística. Email: avljunior@yahoo.com.br

³ UFSM - Graduando do Curso de Zootecnia. Email: dlemes@gmail.com

⁴ UFSM - Mestranda de Engenharia da Produção. Email: vivianedsenna@hotmail.com

⁵ UFSM - Doutorando em Medicina Veterinária. Email: alfredosps@hotmail.com

⁶ UFSM - Orientador. Professor do Departamento de Zootecnia. Email: pacheco.dz.ufsm@hotmail.com

REFERÊNCIAS

BARROS, C. S. et al. Rentabilidade da produção de ovinos de corte em pastagem e em confinamento. *Revista Brasileira de Zootecnia*, v.38, n.11, p.2270-2279, 2009. Disponível em: <<http://www.scielo.br/pdf/rbz/v38n11/a29v3811.pdf>>. Acesso em 12 Set. 2013.

BOFILL, Francisco Jorge. **A reestruturação da ovinocultura gaúcha**. Livraria e editora Agropecuária. Rio Grande do Sul: Guaíba. 1996.

CEZAR, Alfredo Skrebky. **Alternativas farmacológicas para a resistência parasitária múltipla em rebanhos de ruminantes: uma nova abordagem**. Dissertação (mestrado) – Universidade Federal de Santa Maria, Centro de Ciências Rurais, Programa de Pós-Graduação em Medicina Veterinária, 2010.

GOUVEIA, A. M. G; HADDAD, J. P. A. e RIBEIRO, J. G. B. L. Viabilidade econômica da criação de ovinos de corte nas regiões centro-oeste e sudeste. Brasília (DF): LK Editora e Comércio de Bens Editoriais e Autorais Ltda., 2006.

LAPPONI, J.C. *Projetos de investimento na empresa*. Rio de Janeiro: Elsevier, 2007. 488p.

LIMA, G. F. C., ARAUJO, G. G. L. e MACIEL, F. C. Produção e conservação de forragens em escala para sustentabilidade dos rebanhos caprinos e ovinos na agricultura de base familiar. Anais do III Simpósio Internacional sobre Caprinos e Ovinos de Corte João Pessoa, Paraíba, 2007.

SOUZA, A.; CLEMENTE, A. *Decisões financeiras e análise de investimentos*. 6ª ed., São Paulo: Atlas, 2009. 186 p.

VIANA, J. G. A.; SOUZA, R. S. Comportamento dos preços dos produtos derivados da ovinocultura no Rio Grande do Sul no período de 1973 a 2005. *Ciência Agrotécnica*, Lavras, vol. 31, n. 1, p. 191-199, 2007. Disponível em: <<http://www.scielo.br/pdf/cagro/v31n1/v31n1a28.pdf>>. Acesso em 10 Set. 2013.

VIANA, J. G. A.; SILVEIRA, V. C. P. Análise econômica da ovinocultura na metade sul do rio grande do sul, Brasil. In: *XLXI Congresso da Sociedade Brasileira de Economia, Administração e Sociologia Rural*, 2008. Acre: SOBER. Disponível em

¹ UFSM - Pós-graduando de Especialização em Estatística. Email: ferrezer25@hotmail.com

² UFSM - Pós-graduando de Especialização em Estatística. Email: avljunior@yahoo.com.br

³ UFSM - Graduando do Curso de Zootecnia. Email: dlemes@gmail.com

⁴ UFSM - Mestranda de Engenharia da Produção. Email: vivianedsenna@hotmail.com

⁵ UFSM - Doutorando em Medicina Veterinária. Email: alfredosps@hotmail.com

⁶ UFSM - Orientador. Professor do Departamento de Zootecnia. Email: pacheco.dz.ufsm@hotmail.com

<<http://coral.ufsm.br/extrural/vicentepp/arquivospdf/analise%20economica%20da%20ovinocultura.pdf>
f> Acesso em 12 Set. 2013.

¹ UFSM - Pós-graduando de Especialização em Estatística. Email: ferrezer25@hotmail.com

² UFSM - Pós-graduando de Especialização em Estatística. Email: avljunior@yahoo.com.br

³ UFSM - Graduando do Curso de Zootecnia. Email: dlemes@gmail.com

⁴ UFSM - Mestranda de Engenharia da Produção. Email: vivianedsenna@hotmail.com

⁵ UFSM - Doutorando em Medicina Veterinária. Email: alfredosps@hotmail.com

⁶ UFSM - Orientador. Professor do Departamento de Zootecnia. Email: pacheco.dz.ufsm@hotmail.com

Explorando um modelo de segregação alélica em tetraplóides

Felipe Grillo Pinheiro¹

Cleber Bisognin²

Resumo: Poliploidia, existência de mais de dois conjuntos de cromossomos (genomas) no mesmo núcleo, tem sido um importante mecanismo na evolução de eucariotos. Autopoliploidia e alopoliploidia são as duas principais categorias em que são classificados os poliploides e sua classificação pode ser efetuada a partir do tipo de segregação alélica que o tetraploide apresenta. O objetivo do trabalho foi em investigar as propriedades de um modelo da literatura através de simulações verificando o tamanho da amostra de gametas necessário para analisar a segregação alélica a partir de parentais com diferentes genótipos e avaliando a média e o EQM das estimativas. A análise preliminar das propriedades do modelo indicou o EQM das estimativas de $\hat{\theta}$ aumentam conforme aumenta-se a quantidade de alelos repetidos no genótipo parental, considerando os parâmetros testados. Além disso a média das mesmas estimativas convergem mais rapidamente quando se tem os quatro alelos distintos no parental. O comportamento das estimativas variam de acordo com o valor de τ .

Palavras-chave: Poliploidia, segregação alélica, distribuição multinomial, tamanho amostral.

1 Introdução

A Poliploidia, existência de mais de dois conjuntos de cromossomos (genomas) no mesmo núcleo, tem sido um importante mecanismo na evolução de eucariotos (OTTO, 2007; PARISOD et al., 2009). Estudos recentes demonstraram que o genoma de poliploides pode sofrer extensa e rápida reorganização, através de alterações funcionais e estruturais, reforçando o interesse em entender a multiplicação genômica como um processo evolutivo, bem como suas consequências (genéticas e ecológicas) sobre o estabelecimento e manutenção de poliploides na natureza (PARISOD et al., 2009; DOYLE et al., 2008).

Autopoliploidia e alopoliploidia são as duas principais categorias em que são classificados os poliploides (RAMSEY; SCHEMSKE, 1998), dependendo da origem dos genomas envolvidos na formação do tetraploide. Autopoliploides surgem a partir de parentais diploides da mesma espécie, como resultado da duplicação de genomas estruturalmente similares e homólogos ($AA + AA = AAAA$).

¹ UFRGS - Universidade Federal do Rio Grande do Sul. Email: grillopf@gmail.com

² UFRGS - Universidade Federal do Rio Grande do Sul. Email: cleberbisognin@gmail.com

Dessa forma, cada cromossomo pode parear aleatoriamente com qualquer um dos seus homólogos, formando tetravalentes e levando à herança tetrassômica, pois as combinações cromossômicas são aleatórias e todos os genótipos possíveis para os gametas, a partir do genótipo parental são produzidos em frequências iguais. Aloploidos surgem através da duplicação de genomas não-homólogos (homeólogos), ou seja, $(AA + BB = AB \rightarrow \text{duplicação} \rightarrow AABB)$. Como há o envolvimento de duas espécies diferentes existem dois genomas homeólogos entre si, cada um composto por dois conjuntos de cromossomos homólogos entre si (OLSON, 1997; PARISOD et al., 2009; SOLTIS; SOLTIS, 2000; STIFT et al., 2008). Assim, um cromossomo parecia exclusivamente com seu homólogo, ou seja, há apenas formação de bivalentes, levando a uma herança dissômica onde os genótipos possíveis dos gametas também são produzidos na mesma frequência, mas alguns genótipos jamais são produzidos.

Adicionalmente, Stebbins (1971) classificou táxons aloploidos que formam multivalentes durante a meiose como “aloploidos segmentares”, com o objetivo de indicar a duplicação de genomas parcialmente diferenciados $(AA + A'A' = AA' \rightarrow \text{duplicação} \rightarrow AAA'A')$. A diferença em relação aos aloploidos é que os subgenomas A e A' são semelhantes o suficiente para serem considerados homólogos em alguns conjuntos cromossômicos. Nesse caso a segregação alélica é intermediária.

Stift (2008), desenvolveu um procedimento, baseado na verossimilhança, para avaliar qual modelo de herança, se dissômica, tetrassômica ou intermediária, explica melhor a segregação alélica de marcadores genéticos em tetraploides. Entretanto, não se sabe o tamanho de amostra dos gametas adequado para avaliar a segregação em um marcador de um determinado genótipo parental. O objetivo deste trabalho consiste em investigar as propriedades do modelo teórico proposto por Stift (2008) através de simulações verificando o tamanho da amostra de gametas necessário para analisar a segregação alélica a partir de parentais com diferentes genótipos e avaliando a média e o EQM das estimativas $\hat{\theta}$ para os valores $\vec{\theta} = \langle \tau, \beta, \delta_1, \delta_2, \delta_3 \rangle$ do modelo fornecidos para a geração da amostra.

2 Metodologia

O modelo proposto na literatura explica as proporções (probabilidades) de todos os possíveis gametas produzidos por um indivíduo com genótipo ABCD conforme o seguinte sistema na forma matricial:

$$\begin{array}{l}
 \Pr(AA) \\
 \Pr(BB) \\
 \Pr(CC) \\
 \Pr(DD) \\
 \Pr(AB) \\
 \Pr(AC) \\
 \Pr(AD) \\
 \Pr(BC) \\
 \Pr(BD) \\
 \Pr(CD)
 \end{array}
 =
 \begin{bmatrix}
 0 \\
 0 \\
 0 \\
 0 \\
 1/6 \\
 1/6 \\
 1/6 \\
 1/6 \\
 1/6 \\
 1/6
 \end{bmatrix}
 \tau +
 \begin{bmatrix}
 1/4 \\
 1/4 \\
 1/4 \\
 1/4 \\
 -1/6 \\
 -1/6 \\
 -1/6 \\
 -1/6 \\
 -1/6 \\
 -1/6
 \end{bmatrix}
 \beta\tau +
 \begin{bmatrix}
 0 & 0 & 0 \\
 0 & 0 & 0 \\
 0 & 0 & 0 \\
 0 & 0 & 0 \\
 0 & 1/4 & 1/4 \\
 1/4 & 0 & 1/4 \\
 1/4 & 1/4 & 0 \\
 1/4 & 1/4 & 0 \\
 1/4 & 0 & 1/4 \\
 0 & 1/4 & 1/4
 \end{bmatrix}
 (1 - \tau) \langle \delta_1, \delta_2, \delta_3 \rangle \quad (1)$$

Estas equações definem um conjunto de equações não-lineares com cinco parâmetros. O parâmetro τ indica a proporção de gametas formados por associações aleatórias entre os cromossomos durante a meiose, portanto é o parâmetro que está ligado à segregação tetrassômica e seu valor pode variar de zero (segregação dissômica) à um (segregação tetrassômica). Se $\tau < 1$, as proporções dos gametas esperados depende de um conjunto de três parâmetros “dissômicos” (δ_1 , δ_2 e δ_3) que indicam, respectivamente, o grau de pareamento preferencial em cada tipo de associação alelo-genoma. O parâmetro β representa a frequência do fenômeno redução-dupla relativo à frequência de formação de quadrivalentes na meiose.

Estimação de parâmetros: A função logverossimilhança da distribuição multinomial é a função objetivo a ser maximizada. Portanto, o interesse é encontrar os parâmetros do modelo que resultem na maior logverossimilhança.

Simulação: Para avaliar o modelo e suas propriedades, uma rotina foi escrita (no R) para mimetizar a segregação dos gametas de um marcador genético, a partir das probabilidades esperadas, considerando um dado genótipo parental. Pelo caráter inicial do estudo, foram avaliados alguns valores para os parâmetros $\vec{\theta}$. O algoritmo da rotina seguiu os seguintes passos:

- 1) Escolha dos valores para os parâmetros $\vec{\theta} = \langle \tau, \beta, \delta_1, \delta_2, \delta_3 \rangle$;
- 2) Produzir re amostras de tamanho n
- 3) Obter estimativa do vetor $\hat{\vec{\theta}}$ para cada uma das re amostras de tamanho n criadas.
- 4) Avaliar a média e o EQM destas estimativas em relação ao vetor de parâmetros $\vec{\theta}$ conhecido;
- 5) Repetir os passos 2) e 3) para cada tamanho de amostra escolhido para avaliação.

Com o propósito de demonstração foram utilizados nas simulações um número de replicações $re=100$, tamanhos de amostra $n=10, 25, 50, 75, 100$ e 150 e dois conjuntos de valores iniciais para $\vec{\theta}$ com $\tau=0.9$, e 0.67 , e o restante dos parâmetros fixos com os valores $\beta=0.004$, $\delta_1=0.22$, $\delta_2=0$ e $\delta_3=0.78$.

3 Resultados e discussão

A tabela 1 apresenta a média e o EQM das 100 estimativas para os dois conjuntos de parâmetros escolhidos, com os tamanhos de amostra estudados, em dois genótipos parentais. Para o genótipo parental ABCD, que contém as quatro informações alélicas distintas, a média das estimativas dos parâmetros tendem a convergir para o valor real dos parâmetros, quando o valor $\tau=0.67$. A mesma analogia pode ser feita para as estimativas de β , δ_1 , δ_2 e δ_3 . Entretanto, quando os valores de τ são aumentados para 0.9, as estimativas de todos os parâmetros não convergem rapidamente para os verdadeiros valores dos parâmetros, mesmo com tamanhos de amostras relativamente altos.

Para o caso em que o genótipo parental é ABCA, isto é, contém alelos repetidos, o mesmo comportamento se mantém da convergência se mantém como no caso ABCD, entretanto a convergência é mais lenta, sendo o maior valor de tamanho de amostra ($n=150$) insuficiente para a obtenção da convergência adequada. Para valores de $\tau=0.67$ as estimativas dos parâmetros possuem menor EQM se compara no caso em que $\tau=0.9$ (Tabela1).

Tabela 1. Estimativas da média e do EQM para os parâmetros em 100 replicações nos diferentes tamanhos amostrais dos gametas a partir dos vetores de parâmetros conhecidos e o genótipos parentais.

Genótipo	n	Média					EQM				
		τ	β	δ_1	δ_2	δ_3	τ	β	δ_1	δ_2	δ_3
ABCD	$\vec{\theta}$	0.67	0.04	0.22	0.00	0.78					
	10	0.46	0.05	0.35	0.22	0.43	0.16	0.01	0.17	0.18	0.29
	25	0.51	0.06	0.21	0.23	0.56	0.14	0.00	0.08	0.18	0.18
	50	0.60	0.06	0.26	0.13	0.61	0.08	0.00	0.11	0.09	0.15
	75	0.58	0.07	0.25	0.10	0.65	0.07	0.00	0.09	0.06	0.11
	100	0.63	0.05	0.25	0.09	0.66	0.05	0.00	0.08	0.06	0.11
	150	0.68	0.05	0.19	0.14	0.67	0.04	0.00	0.06	0.11	0.11
	$\vec{\theta}$	0.90	0.04	0.22	0.00	0.78					
	10	0.48	0.07	0.32	0.28	0.4	0.3	0.01	0.15	0.23	0.31
	25	0.58	0.09	0.31	0.24	0.44	0.17	0.01	0.15	0.18	0.26
50	0.71	0.07	0.26	0.29	0.45	0.08	0.00	0.12	0.22	0.24	
75	0.73	0.06	0.28	0.26	0.46	0.07	0.00	0.13	0.2	0.26	
100	0.76	0.05	0.28	0.24	0.47	0.05	0.00	0.11	0.18	0.21	
150	0.81	0.05	0.26	0.19	0.55	0.02	0.00	0.11	0.13	0.18	
ABCA	$\vec{\theta}$	0.67	0.04	0.22	0.00	0.78					
	10	0.40	0.06	0.19	0.22	0.60	0.20	0.01	0.09	0.14	0.17
	25	0.44	0.07	0.19	0.22	0.59	0.17	0.01	0.08	0.16	0.16
	50	0.48	0.07	0.23	0.16	0.61	0.14	0.01	0.08	0.09	0.13
	75	0.46	0.08	0.16	0.17	0.67	0.15	0.01	0.06	0.10	0.09
	100	0.45	0.08	0.16	0.18	0.65	0.16	0.01	0.06	0.09	0.08
	150	0.46	0.09	0.17	0.15	0.67	0.13	0.01	0.06	0.07	0.07
	$\vec{\theta}$	0.90	0.04	0.22	0.00	0.78					

10	0.36	0.07	0.29	0.24	0.47	0.42	0.01	0.15	0.18	0.24
25	0.60	0.07	0.24	0.30	0.46	0.21	0.01	0.10	0.23	0.26
50	0.59	0.07	0.27	0.26	0.48	0.19	0.00	0.12	0.18	0.24
75	0.67	0.07	0.28	0.22	0.50	0.14	0.00	0.13	0.14	0.22
100	0.61	0.08	0.26	0.31	0.43	0.18	0.00	0.12	0.21	0.25
150	0.64	0.08	0.29	0.23	0.47	0.17	0.00	0.12	0.16	0.21

O gráfico1, ilustra o comportamento da média e do EQM das 100 estimativas para diferentes valores de τ e tamanhos de amostra e cenários de genótipos parentais. A convergência para os verdadeiros parâmetros é mais lenta conforme aumenta a quantidade de alelos repetidos.

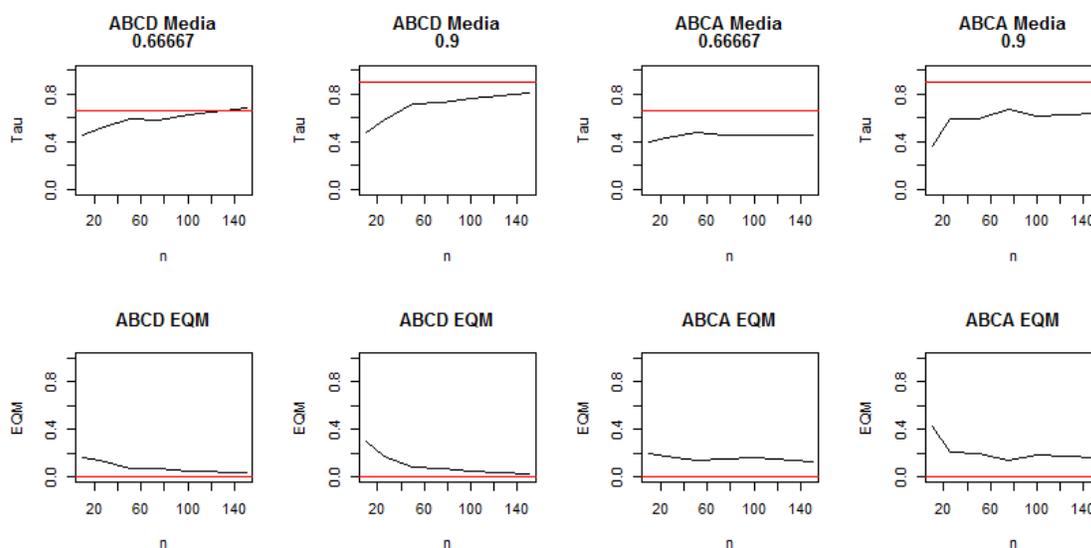


Figura 1. Comportamento da média e do EQM das 100 estimativas para diferentes valores de τ e tamanhos de amostra.

Como no caso ABCD existem várias categorias de genótipos possíveis para os gametas, conforme se aumenta o tamanho da amostra, mais refinada se tornam as probabilidades oriundas do modelo em relação àquela esperada pela amostra e a verossimilhança é maior. No caso de um parental ABCA, o alelo A está repetido, portanto, não são esperados nenhum dos genótipos envolvendo o alelo D que são encontrados nos gametas de ABCD. Como existem menos categorias de genótipos dos gametas a capacidade de discriminar as probabilidades entre diversas possibilidades de parâmetros, a convergência torna-se mais lenta ou até impossível na prática, pois a partir de $n=100$ a obtenção de mais progênie é comprometida.

3 Conclusão

A análise preliminar das propriedades do modelo indicou que a capacidade de discriminação entre as probabilidades de cada categoria de genótipo dos gametas de diferentes valores dos parâmetros estimados $\hat{\theta}$ diminui conforme aumenta-se a quantidade de alelos repetidos no genótipo parental, considerando os parâmetros testados. Isso significa que deve-se tomar cuidado ao utilizar este método quando não temos completa distinção dos alelos parentais. Além disso, as estimativas do parâmetro τ podem ser subestimadas quando o tamanho da amostra de gametas é pequeno, mesmo quando o genótipo parental não tem alelos repetidos. Uma limitação do estudo consiste no fato de que foi utilizado um número muito baixo de replicações, especialmente por se tratar de um problema envolvendo proporções. Entretanto, estas análises são preliminares e serão estendidas para outros genótipos, vetor de parâmetros, tamanhos de amostra e número de replicações.

Referências

- [1] DOYLE, J. J. *et al.* Evolutionary genetics of genome merger and doubling in plants. *Annual Review of Genetics* [S.I.], v. 42, p.443-461, 2008.
- [2] OTTO, S. P. The evolutionary consequences of polyploidy. *Cell* [S.I.], v. 131 p. 452-462, 2007.
- [3] PARISOD, C.; HOLDEREGGER, R.; BROCHMANN, C. Evolutionary consequences of autopolyploidy. *New Phytologist* [S.I.], v. 186, n. 1, p. 5–17, 2009.
- [4] RAMSEY, J.; SCHEMSKE, D.W. Pathways, mechanisms and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics*, Palo Alto, v. 29, p. 467-501, 1998.
- [5] SOLTIS, P. S.; SOLTIS D. E. The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences, USA*, v. 97, p. 7051-7057, 2000.
- [6] STEBBINS, G.L. Chromosomal evolution in higher plants. London, UK. p. 216, 1971.
- [7] STIFT, M. *et al.* Segregation models for disomic, tetrasomic and intermediate inheritance in tetraploids: a general procedure applied to *Rorippa* (yellow cress) microsatellite data. *Genetics* [S.I.], v. 179, p. 2113-2123, 2008.

Comunicações Pôster

Avaliação numérica dos testes de Wald e da razão de verossimilhanças para seleção de função de ligação em modelos para dados binários

Diego R. Canterle¹

Fábio M. Bayer²

Ao ajustar um Modelo Linear Generalizado (MLG) [1] é necessária a utilização de uma função de ligação. Torna-se prudente utilizar um teste para verificar a adequacidade dessa função de ligação selecionada. Este trabalho aborda o problema de verificação da adequacidade da função de ligação em um MLG com variáveis dependentes binárias. Particularmente será investigada a utilização do teste RESET de Ramsey [2] em amostras de tamanho finito. Para a realização desse teste são consideradas as estatísticas de Wald e da razão de verossimilhanças. Sob condições usuais de regularidade e sob hipótese nula (\mathcal{H}_0) essas estatísticas de teste convergem em distribuição para uma qui-quadrado com r graus de liberdade (χ_r^2), em que r é o número de restrições impostas sob \mathcal{H}_0 . Dessa forma, os testes podem ser realizados usando valores críticos aproximados como quantis da distribuição χ_r^2 . Contudo, em pequenas amostras a distribuição nula limite (χ_r^2) pode fornecer uma aproximação pobre à distribuição nula exata das estatísticas de teste, implicando distorções do tamanho do teste. Com isso, é de suma importância avaliar a performance desses testes em amostras de tamanho finito. Via simulações de Monte Carlo foram avaliados o tamanho e o poder dos testes de Wald e da razão de verossimilhanças, considerando como possíveis funções de ligação as funções logit, probit, cloglog e cauchit. Todas as simulações foram feitas utilizando o *software* R [3]. Pode-se verificar que apesar das distorções dos testes em pequenas amostras, eles podem ser boas alternativas para testar a adequacidade da função de ligação.

Palavras-chave: Modelos binários, Estatística Wald, Estatística da razão de verossimilhança, Função de ligação, Simulações de Monte Carlo, Teste RESET.

Referências

- [1] McCullagh, P.; Nelder, J. *Generalized linear models*, 2 ed. Chapman and Hall, 1989.
- [2] Ramsey, J.B. Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society*, v. 31, p. 350-371, 1969.
- [3] R Development Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

¹UFSM - Universidade Federal de Santa Maria, Bacharelado em Estatística. Email: diegocanterle@gmail.com

²UFSM - Universidade Federal de Santa Maria, Departamento de Estatística, Laboratório de Ciências Espaciais de Santa Maria (LACESM). Email: bayer@ufsm.br

Avaliação numérica dos estimadores do modelo $\beta ARMA$

Bruna Gregory Palm^{1 3}

Fábio M. Bayer^{2 3}

Tradicionalmente para a modelagem e previsão de variáveis ao longo do tempo utiliza-se modelos autorregressivos integrados de médias móveis. Entretanto, tais modelos tornam-se inapropriados quando a variável resposta y está restrita ao intervalo $(0,1)$. Para essas situações, foi proposto o modelo beta autorregressivo de médias móveis ($\beta ARMA$) [1], em que a variável de interesse y_t , $t = 1, \dots, n$, possui distribuição beta. A densidade beta, condicional ao conjunto de informações prévias \mathcal{F}_{t-1} , é: $f(y_t|\mathcal{F}_{t-1}) = \frac{\Gamma(\phi)}{\Gamma(\mu_t\phi)\Gamma((1-\mu_t)\phi)} y_t^{\mu_t\phi-1} (1-y_t)^{(1-\mu_t)\phi-1}$, em que $0 < y_t < 1$, $E(y_t|\mathcal{F}_{t-1}) = \mu_t$ e $Var(y_t|\mathcal{F}_{t-1}) = (\mu_t(1-\mu_t))/(1+\phi)$. Dessa forma, o modelo $\beta ARMA(p, q)$ pode ser escrito como: $g(\mu_t) = \alpha + \sum_{i=1}^p \varphi_i g(y_{t-i}) + \sum_{j=1}^q \theta_j r_{t-j}$, em que $g(\cdot)$ é uma função de ligação, φ 's e θ 's são os parâmetros autorregressivos e de médias móveis e r_t é o termo do erro. As inferências pontuais sobre os parâmetros do modelo são baseadas em estimadores de máxima verossimilhança (EMV). Os EMV possuem boas propriedades assintóticas, mas em pequenas amostras as aproximações assintóticas podem ser pobres gerando resultados inferenciais distorcidos. O presente trabalho objetiva avaliar numericamente, via simulações de Monte Carlo, os EMV do modelo $\beta ARMA$ em amostras de tamanho finito. A implementação computacional foi desenvolvida em linguagem R [2]. Para a avaliação numérica considerou-se média, viés, viés relativo, erro-padrão e erro quadrático médio dos EMV. Observou-se que os estimadores do parâmetro de precisão ϕ e dos parâmetros de médias móveis são mais viesados do que os parâmetros autorregressivos. Em trabalhos futuros pretende-se obter estimadores corrigidos via método Bootstrap.

Palavras-chave: Avaliação numérica, Estimadores de máxima verossimilhança, Modelo $\beta ARMA$, Simulações de Monte Carlo.

Referências

- [1] ROCHA, A. V., CRIBARI-NETO, F. Beta autoregressive moving average models. *Test* 18 (3), p. 529-545, 2009.
- [2] R Core Team, 2012. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*. Disponível em: <http://www.R-project.org/>. Vienna, Austria, ISBN 3-900051-07-0.

¹UFSM - Universidade Federal de Santa Maria, Bacharelado em Estatística. Email: brunagpalm@gmail.com

²UFSM - Universidade Federal de Santa Maria, Departamento de Estatística, Laboratório de Ciências Espaciais de Santa Maria (LACESM). Email: bayer@ufsm.br

³Agradecimento ao Fundo de Incentivo à Pesquisa (FIPE/UFSM)

Assessoria estatística na área da saúde na UFSM

Angelica Peripolli¹

Anaelena Bragança de Moraes²

O trabalho de assessoria estatística é uma atividade solicitada com bastante frequência no Departamento de Estatística da Universidade Federal de Santa Maria. Parte dessas solicitações é realizada por professores de outras áreas e também por alunos de pós-graduação, sendo que muitos desses auxílios não têm nenhum tipo de registro, dado que o professor que faz assessoria muitas vezes não é orientador, nem co-orientador e nem faz parte da publicação do artigo resultante.

A realização dessa atividade proporciona um grande aprendizado a todos os envolvidos, principalmente aos alunos, pois permite o contato com diversas realidades, uma vez que atende vários tipos de pesquisas e, principalmente, porque trata de dados reais, gerados pela comunidade acadêmica.

A assessoria estatística é uma atividade de grande importância na vida acadêmica, pois tem o propósito de qualificar as análises dos dados dos trabalhos científicos, além de proporcionar experiência e conhecimento da equipe de professores e alunos envolvidos, pela utilização de técnicas estatísticas eficientes. Na assessoria são utilizados vários métodos estatísticos para a análise dos dados, tais como a Estatística Descritiva, a Estatística Inferencial Paramétrica e Não-paramétrica, a Análise Multivariada, Validação de Protocolo entre outras.

O aprendizado obtido na assessoria proporciona aos participantes maior competência para o desenvolvimento de outros estudos nessa área. Acrescido a essa experiência, os benefícios também retornam a comunidade, por proporcionarem subsídios na tomada de decisões pelos gestores em saúde.

Palavras-chave: *Assessoria estatística, saúde, alunos de pós-graduação.*

¹ UFSM - Universidade Federal de Santa Maria. Email: angelicaperipolli@gmail.com

² UFSM - Universidade Federal de Santa Maria. Email: anaelena@smail.ufsm.br

Provas simuladas para estudo em Estatística

BRUNA MARTINI DALMORO¹
LISIANE PRISCILA ROLDÃO SELAU²
LUCIANA NEVES NUNES³

A tecnologia aplicada à educação é uma das principais ferramentas que estimulam uma participação ativa do aluno no aprendizado [1]. O estudante deixa de ser apenas o receptor de informações e começa a interagir com o conteúdo e, principalmente, toma a iniciativa de aprender [2]. Além disso, as tecnologias computacionais têm mudado rapidamente a forma de se ensinar Estatística [3]. Neste sentido, o projeto propõe a construção de um site que funcionará como um simulador de provas, com questões de múltipla escolha, semelhantes às questões utilizadas em concursos públicos. Estas questões estão classificadas por áreas dentro da Estatística (Descritiva, Probabilidade e Inferência) e, no momento que o aluno escolher qual área pretende estudar, uma prova simulada será gerada. O aluno responderá as questões assinalando a resposta correta e, assim que concluir a prova, o resultado aparecerá na tela, apontando os erros e, inclusive, as soluções das questões. Também deverão ser disponibilizados material de estudo, apostilas, vídeos, entre outros, além de um pequeno resumo do que se trata cada área da estatística estudada pelos alunos. A página para geração das provas simuladas está sendo construída utilizando das linguagens PHP, HTML e SQL. Esta página poderá também ser disponibilizada pelos professores via plataforma de ensino utilizada na disciplina (Moodle, por exemplo), beneficiando pelo menos 1.500 alunos por semestre.

Palavras-chave: *Educação à distância, estatística, provas simuladas.*

Referências

- [1] MORAN, J. M. Para onde caminhamos na educação? São Paulo: 2005. Disponível em: <<http://www.eca.usp.br/prof/moran/caminhamos.htm>>.
- [2] VELLEMAN, P. F.; MOORE, D. S. Multimedia for teaching statistics: Promises and pitfalls. The American Statistician, v.50, p.217-225. 1996.
- [3] PAN, W. S. The challenges of teaching statistics in the current technology environment. Journal of American Academy of Business, v.3, p.351-355, 2003.

¹ UFRGS - Universidade Federal do Rio Grande do Sul. Email: brunamdalmoro@hotmail.com

² UFRGS - Universidade Federal do Rio Grande do Sul. Email: lisiane.selau@ufrgs.br

³ UFRGS - Universidade Federal do Rio Grande do Sul. Email: lununes@mat.ufrgs.br

Avaliação da importância atribuída pelos usuários aos atributos do RU

Angelica Peripolli¹

Fernando Moreira Junior²

Bruna Gregory Palm³

Caroline Pafiadache⁴

Laís Helen Loose⁵

Rafael Piaia⁶

Vinícius Scher⁷

Grande parte dos estudantes universitários que possuem uma carga horária extensa, muitas vezes morando distante da universidade, opta por fazer suas refeições na própria, assim com professores e servidores públicos que trabalham na Universidade Federal de Santa Maria (UFSM). Essa pesquisa quantitativa tem por objetivo avaliar a importância dada aos atributos do Restaurante Universitário (RU) da UFSM - RS de acordo com a opinião dos usuários. Os usuários preencheram o questionário de avaliação formulado através do aplicativo Google Docs. Este instrumento de avaliação foi direcionado aos usuários do RU através de e-mails oriundos do Sistema de Informação para o Ensino - SIE da UFSM. Os respondentes classificaram cada atributo de acordo com seu grau de importância, de 1 a 5. Após a coleta e organização dos dados, estes foram inseridos no programa computacional SPSS 17.0 e analisados. A análise descritiva permitiu concluir que os itens mais importantes no ambiente interno do RU são a limpeza e os odores, já para o ambiente externo, o abrigo do sol e chuva e a acessibilidade. Na alimentação, os atributos relacionados a qualidade e a higiene dos alimentos servidos são de extrema importância seguidos pelo sabor, a qualidade e a variedade da carne e por fim a agilidade com que os alimentos são repostos. Os atributos que obtiveram menor classificação de importância foram a decoração e a fachada do RU e no quesito alimentação, o guisado de soja. Esses dados geram subsídios para formulação de propostas que visam melhorias nos setores avaliados.

Palavras-chave: *Restaurante Universitário, importância, atributos, análise descritiva.*

¹ UFSM - Universidade Federal de Santa Maria. Email: angelicaperipolli@gmail.com

² UFSM - Universidade Federal de Santa Maria. Email: fmjunior777@yahoo.com.br

³ UFSM - Universidade Federal de Santa Maria. Email: brunagpalm@gmail.com

⁴ UFSM - Universidade Federal de Santa Maria. Email: carolpafiadache@yahoo.com.br

⁵ UFSM - Universidade Federal de Santa Maria. Email: laisloose@gmail.com

⁶ UFSM - Universidade Federal de Santa Maria. Email: rpiaia@gmail.com

⁷ UFSM - Universidade Federal de Santa Maria. Email: profvinimat@hotmail.com

Classificação e agrupamento de séries temporais correlacionadas.

Douglas Roberto Mesquita Azevedo¹

Marcio Valk²

Neste presente trabalho abordamos o problema de classificação e agrupamento de séries temporais. Utilizando o método baseado em U-estatísticas, foi realizada a extensão de resultados já obtidos para séries temporais não correlacionadas para o caso onde há dependência entre as séries. Com estes resultados devidamente comprovados, podemos então classificar e agrupar séries em grupos com comportamento semelhante. Classificação e agrupamento de séries temporais são problemas bastante explorados na literatura atual e muitas técnicas são apresentadas para resolver esses problemas, no entanto, os procedimentos são específicos, aplicáveis somente a uma determinada classe de séries temporais.

Resultados preliminares obtidos mostram que a técnica para classificação e agrupamento de séries temporais que está sendo proposta é bastante geral, aplicável a uma ampla classe de séries. No entanto, pretende-se generalizar ainda mais a classe de séries para a qual a técnica funciona.

A partir de simulações de diferentes processos e parametrizações, são comparados grupos de séries de diferentes tamanhos e com alguns níveis de dependência, para avaliar os resultados obtidos através do método em questão.

As técnicas mais conhecidas para agrupar e classificar séries temporais baseiam-se em métricas que são utilizadas para medir a distância ou dissimilaridade entre séries temporais. No presente trabalho algumas métricas serão utilizadas como núcleo de U-estatísticas com a finalidade de observar a métrica que apresenta melhor resultado em cada uma das situações testadas.

Com os resultados obtidos através de inúmeras simulações, foi possível perceber a robustez do método mesmo quando as séries possuem dependência.

Palavras-chave: *Séries temporais, U-estatística, Classificação, Agrupamento.*

¹ UFRGS - Universidade Federal do Rio Grande do Sul. Email: douglas-mesquita@hotmail.com

² UFRGS - Universidade Federal do Rio Grande do Sul. Email: marciovalk@gmail.com

ANÁLISE DA EVOLUÇÃO DO NÚMERO DE ACIDENTES DE TRABALHO NO RIO GRANDE DO SUL DE 1999 A 2010

Anabeth Petry Radünz¹

Eduardo L. Ensslin dos Santos¹

William Jacobs²

Roselaine Ruviaro Zanini³

Acidentes de trabalho são assim classificados: típico, de trajeto e doença profissional. O acidente típico ocorre no desenvolvimento do trabalho no ambiente da companhia; o de trajeto ocorre no caminho da residência para a empresa e vice-versa; doença profissional é adquirida por condições especiais onde o trabalho é feito. O objetivo deste trabalho é realizar uma análise descritiva e analisar tendências do número de acidentes ocorridos de 1999 a 2010. Observou-se que a proporção dos acidentes típicos no RS foi de 83,43%; de acidentes de trajeto foi de 11,26%; e, de doença profissional foi de 5,31% do total do número de acidentes registrados. Do total de acidentes típicos, 17,89% correspondem ao grupo de 20 a 24 anos; 23,86% do acidente de trajeto correspondem ao grupo de 25 a 29 anos; e, 17,74% da doença profissional correspondem ao grupo de 40 a 44 anos. Verificou-se que, grande parte dos acidentes típicos e de trajeto está relacionada à faixa dos 20 a 29 anos (homens). Quanto ao adoecimento, verificou-se que a maior parte das ocorrências está relacionada à faixa de 35 a 39 anos. Além disso, foi verificada uma tendência crescente no número de acidentes típicos ($\beta=2,089$, $p<0,001$), no acidente de trajeto ($\beta=2,276$, $p<0,001$) e um declínio no número de doença profissional ($\beta=-0,546$, $p<0,001$). Destaca-se a importância da análise quantitativa dos acidentes de trabalho para um melhor entendimento sobre aumentos/declínios das taxas e os grupos que têm maior influência neles, servindo como uma informação importante para o desenvolvimento das políticas públicas.

Palavras chave: Acidentes de trabalho. Análise de séries temporais. Tendência.

¹ Acadêmico do Curso de Bacharelado em Estatística da UFSM

² Acadêmico do Mestrado em Engenharia de Produção da UFSM

³ Docente do Departamento de Estatística da UFSM

Referências

AET INFOLOG. **Base de dados históricos de acidentes do trabalho**. Disponível em: <http://www3.dataprev.gov.br/AEAT/greg/reg02/reg02.PHP>. Acessado em: 18/06/2013

CORREA, P. R. L., ASSUNÇÃO, A. A. A subnotificação de mortes por acidentes de trabalho: estudo de três bancos de dados. **Revista Epidemiologia e Serviços de Saúde**. Vol 12, número 4, p. 203-212, 2003.

FILHO, V. W. Reestruturação produtiva e acidentes do trabalho no Brasil: estrutura e tendências. **Caderno de Saúde Pública**. Vol. 15, p. 41-51, 1999.

Estimação em Processos *k-Factor* GARMA utilizando MCMC

Cleber Bisognin¹

Ian Meneghel Danilevicz^{2,3}

Os processos *k-Factor* GARMA($p, \mathbf{u}, \boldsymbol{\lambda}, q$) foram inicialmente definidos por Giraitis e Leipus (1995) e Woodward et al. (1998), onde $\mathbf{u}=(u_1, \dots, u_k)$ e $\boldsymbol{\lambda}=(\lambda_1, \dots, \lambda_k)$, p e q são os graus dos polinômios autorregressivo e média-móvel, respectivamente. Estudamos as propriedades de estacionariedade e longa dependência destes modelos. Estas propriedades são caracterizadas se u_j são distintos e $0 < \lambda_j < 0.5$, quando $|u_j| < 1$ e $0 < \lambda_j < 0.25$ e $|u_j|=1$, para $j=1, \dots, k$. Estes processos são uma generalização dos processos ARFIMA e SARFIMA, estudados por Lopes (2007) e Bisognin e Lopes (2009), respectivamente. Nos trabalhos sobre estimação dos parâmetros dos processos *k-Factor* GARMA($p, \mathbf{u}, \boldsymbol{\lambda}, q$) são utilizados métodos semiparamétricos e o método paramétrico proposto por Fox e Taquq (1986). Este último proposto inicialmente para os processos ARFIMA. Neste trabalho utilizamos os estimadores Fox e Taquq, denotado por FT, propomos uma correção empírica do vício do estimador proposto por Fox e Taquq (1986), o qual foi denotado por FT modificado, e também um estimador utilizando MCMC (*Markov Chain Monte Carlo*). Fazemos a comparação entre estes métodos de estimação para os processos *k-Factor* GARMA($p, \mathbf{u}, \boldsymbol{\lambda}, q$) via Simulações de Monte Carlo. Em algumas simulações, os processos *k-Factor* GARMA($p, \mathbf{u}, \boldsymbol{\lambda}, q$) são contaminados por *outliers* aditivos, utilizando o método de contaminação por mistura. Para comparação entre os métodos utilizamos o vício, o erro quadrático médio e a variância dos estimadores. Na estimação dos parâmetros dos processos os processos *k-Factor* GARMA($p, \mathbf{u}, \boldsymbol{\lambda}, q$), sem contaminação por *outliers*, os estimadores FT modificado e MCMC apresentaram menor vício, erro quadrático médio e variância que o estimador FT, em todos os caso, mesmo quando p e/ou q são diferentes de zero. Nos casos em que houve contaminação por *outliers* aditivos, os estimadores FT modificado e MCMC também apresentaram um excelente comportamento, apresentado menor vício, erro quadrático médio e variância que o estimador FT. Ambos os resultados foram verificados para diferentes tamanhos amostrais.

Palavras-chave: *Processos Estocásticos, Longa Dependência, Estimação paramétrica, MCMC, Outliers.*

¹ DEST – PPGMAT – UFRGS, cbisognin@ufrgs.br

² Bolsista PIBIC – UFRGS, iandanilevicz@gmail.com

³ Agradecimentos a bolsa de Iniciação Científica concedida pelo CNPq

Referências

- [1] BISOGNIN, C.; S.R.C. Lopes, Properties of Seasonal Long Memory Processes. *Mathematical and Computer Modelling*, 49, 1837-1851, 2009
- [2] FOX, R.; M.S. TAQQU, Large-sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series. *The Annals of Statistics*, vol. 14, p. 517-532, 1986.
- [3] GIRAITIS, L.; R. LEIPUS, A Generalized Fractionally Differencing Approach in Long Memory Modelling. *Lithuanian Mathematical Journal*, vol. 35(1), p. 53-65, 1995.
- [4] LOPES, S.R.C., Long-range Dependence in Mean and Volatility: Models, Estimation and Forecasting. In: *Vares M.E. e V. Sidoravicius* (eds). EBP. Rio de Janeiro: Birkhäuser, 2007.

¹ DEST – PPGMAT – UFRGS, cbisognin@ufrgs.br

² Bolsista PIBIC – UFRGS, iandanilevich@gmail.com

³ Agradecimentos a bolsa de Iniciação Científica concedida pelo CNPq

Avaliação de um teste de reconhecimento de emoções por meio da Teoria da Resposta ao Item: uma análise preliminar

Fernando de Jesus Moreira Junior¹

Felipe Argolo²

Silvio José Lemos Vasconcellos³

O ERI (*Emotional Recognition Index*) é um teste de reconhecimento de emoções composto por duas subescalas, uma para reconhecimento de emoções expressas pela face (FACIAL-I) e outra para reconhecimento por intermédio de expressões verbais (VOCAL-I). O instrumento avalia a habilidade dos indivíduos para inferir corretamente emoção expressa pela face ou estímulos verbais, com base no conceito de competência emocional (SCHERER, 2007). O objetivo do trabalho foi fazer uma avaliação do ERI por meio da Teoria da Resposta ao Item (TRI). A TRI é uma metodologia que sugere formas de representar a relação entre a probabilidade de um indivíduo dar uma certa resposta a um item, os traços latentes do indivíduo e as características dos itens, por meio de modelos matemáticos (ANDRADE; TAVARES; VALLE, 2000). A amostra do estudo foi composta por 195 estudantes universitários de diferentes cursos da UFSM, que deveriam observar 30 expressões faciais e 30 vocalizações apresentadas pelo ERI através do retroprojetor e, em seguida, assinalar em um questionário a emoção correspondente. Os dados foram analisados por meio do pacote *irt* do *Software R* e ajustados por meio do Modelo Logístico de Dois Parâmetros (ML2) da TRI. Numa análise preliminar, observou-se que a maioria dos itens se ajustou de forma apropriada, no entanto, alguns itens do ERI apresentaram um comportamento inadequado, segundo a metodologia TRI.

Palavras-chave: *Reconhecimento de emoções, Teoria da Resposta ao Item, ERI.*

Referências

- [1] ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. *Teoria da resposta ao item: conceitos e aplicações*. São Paulo: ABE - Associação Brasileira de Estatística, 2000.
- [2] SCHERER, K. R. *Component models of emotion can inform the quest for emotional competence*. In G. Matthews, M. Zeidner, & R. D. Roberts (Eds.), *The science of emotional intelligence: Knowns and unknowns*. New York, NY: Oxford University Press, p. 101–126, 2007.

¹ UFSM - Universidade Federal de Santa Maria. Email: fmjunior@smail.ufsm.br

² UFBA - Universidade Federal da Bahia.

³ UFSM - Universidade Federal de Santa Maria. Email: sjlvasconcellos@bol.com.br

Utilização da Teoria da Resposta ao Item (TRI) nos Testes Adaptativos Informatizados (TAI)

Fernando de Jesus Moreira Junior¹

Dalton Francisco de Andrade²

Esse trabalho tem como objetivo apresentar a utilização da Teoria da Resposta ao Item (TRI) na elaboração de Testes Adaptativos Informatizados (TAI), por meio de uma aplicação nas provas teóricas de habilitação para motoristas realizadas pelo DETRAN-SC. Um TAI é um teste que se caracteriza por apresentar itens adequados para o respondente, segundo o seu nível de conhecimento, proporcionando um teste personalizado que se adapta à proficiência do examinando. A TRI é utilizada no processo de calibração do banco de itens de um TAI (estimação dos parâmetros dos itens) e na estimação do traço latente. Foi utilizado o modelo logístico unidimensional de três parâmetros (ML3) da TRI, o qual considera a discriminação (a) e a dificuldade (b) do item e a possibilidade do acerto casual (c). Os dados foram coletados da base de dados do DETRAN-SC de provas aplicadas em mais de 200.000 candidatos em 2008 com um total de 462 itens com quatro alternativas de resposta. Cada prova aplicada continha 40 itens do banco, sendo que cada item tinha uma frequência diferente de aplicação, onde o item com a menor frequência foi aplicado em mais de 3.000 provas. Foram testados oito métodos de calibração por meio do software BILOG-MG. Após várias etapas de calibração, eliminando-se os itens inadequados, permaneceram os itens com $a \geq 1$. O banco final foi composto por 286 itens. Os valores médios dos parâmetros foram: 1,46 (a), -1,46 (b) e 0,07 (c). Pode-se concluir que os itens que permaneceram possuem boa discriminação e são predominantemente fáceis.

Palavras-chave: *Teoria da Resposta ao Item, Testes Adaptativos Informatizados, avaliação, proficiência.*

Referências

- [1] ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. *Teoria da resposta ao item: conceitos e aplicações*. São Paulo: ABE - Associação Brasileira de Estatística, 2000.
- [2] VAN DER LINDEN, W. J.; GLAS, C. A. W. *Computerized Adaptive Testing: Theory and Practice*. Dordrecht, Netherlands: Kluwer Academic, 2003.

¹ UFMS - Universidade Federal de Santa Maria. Email: fmjunior@smail.ufsm.br

² UFSC - Universidade Federal de Santa Catarina. dandrade@inf.ufsc.br

Estimação em Modelos de Mistura de Distribuições Independentes

Cleber Bisognin¹

Régis Nunes Vargas²

Aneli Torres Venturini³

Segundo McLachlan e Peel (2000), a história dos modelos de mistura finita remonta a mais de um século com Pearson (1894), cujos modelos são baseados em uma mistura de duas componentes normais univariadas. O Modelo de Mistura de Distribuições Independente, segundo Zucchini e Macdonald (2009), consiste de um número finito m de distribuições componentes associando-se para cada uma dessas distribuições um valor $0 < \delta_i < 1$, para $1 \leq i \leq k$, onde $\sum_{i=1}^k \delta_i = 1$, são denominadas probabilidades de mistura. Uma variável aleatória X satisfaz o Modelo de Mistura de Distribuições Independente se sua função densidade (ou massa) de probabilidade satisfaz $f_X(x) = \sum_{i=1}^k \delta_i f_{X_i}(x)$, onde X_1, \dots, X_k , são variáveis aleatórias independentes, definidas no mesmo espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, com função densidade (ou massa) de probabilidade dadas, respectivamente, por $f_{X_1}(\cdot), \dots, f_{X_k}(\cdot)$. Neste trabalho provamos que $f_X(x)$, satisfaz as propriedades de função densidade (e massa) de probabilidade, calculamos o valor esperado, momento de ordem t , $t \in \mathbb{Z}$, variância. Definimos a função de verossimilhança e os estimadores de máxima verossimilhança. Realizamos simulações de Monte Carlo, utilizando a distribuição de Poisson, isto é, $X_1 \sim P(\lambda_1)$ e $X_2 \sim P(\lambda_2)$, com $\lambda_i \in \{2, 3, 5, 7\}$, para $i = 1, 2$. Utilizamos tamanho amostral $n \in \{100, 300, 500, 1000\}$, $k = 2$ e 1000 replicações. Analisamos o vício, erro quadrático médio e variância do estimador de máxima verossimilhança para cada um dos parâmetros estimados em cada caso. Quando λ_1 e λ_2 são próximos o vício, o erro quadrático médio e a variância são maiores na estimação de ambos os parâmetros. Conforme λ_1 e λ_2 se distanciam, ambas as características analisadas diminuem sua magnitude. Esta característica ocorre para todos os tamanhos amostrais. Também fazemos uma aplicação com dados do registro dos tremores de terra de magnitude maior ou igual a sete (7:0), no mundo, que ocorreram entre 1900-2006.

Palavras-chave: Modelos de Mistura de Distribuições Independentes, Poisson, Estimação de Máxima Verossimilhança, Tamanho Amostral.

¹ DEST - UFRGS - Universidade Federal do Rio Grande do Sul. Email: cbisognin@ufrgs.br

² PPGMAT - UFRGS - Universidade Federal do Rio Grande do Sul. Email: regisnv@yahoo.com.br

³ DEST - UFRGS - Universidade Federal do Rio Grande do Sul. Email: aneliventurini@yahoo.com.br

Referências

- [1] MCLACHLAN, G. J. e PEEL, D. *Finite Mixture Models*. New York: Wiley Series in Probability and Statistics, 2000.
- [2] PEARSON, K. *Contributions to the Theory of Mathematical Evolution*. Philosophical Transactions of the Royal Society of London, 1894.
- [3] ZUCCHINI, W. e MACDONALD, I. L. *Hidden Markov Models for Time Series: an introduction using R*. Londres: Chapman and Hall, 2009.

¹ DEST - UFRGS - Universidade Federal do Rio Grande do Sul. Email: cbisognin@ufrgs.br

² PPGMAT - UFRGS - Universidade Federal do Rio Grande do Sul. Email: regisnv@yahoo.com.br

³ DEST - UFRGS - Universidade Federal do Rio Grande do Sul. Email: aneliventurini@yahoo.com.br

Análise da Autoavaliação Institucional por meio da Teoria da Resposta ao Item (TRI)

Fernando de Jesus Moreira Junior¹

Charles Rogério Paveglio Szinvelski²

A análise dos dados da autoavaliação institucional da Universidade Federal de Santa Maria (UFSM) é realizada por meio de estatísticas descritivas, tabelas de frequência e gráficos apresentados anualmente no relatório de autoavaliação institucional. Esse trabalho tem como objetivo analisar o instrumento de autoavaliação institucional respondido pelos alunos de graduação por meio da Teoria da Resposta ao Item (TRI). A TRI é uma metodologia muito utilizada nas avaliações educacionais internacionais em larga escala e, inclusive, vem sendo utilizada recentemente pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) na avaliação do Exame Nacional do Ensino Médio (ENEM). Embora a maioria das aplicações esteja relacionada com as avaliações educacionais, a TRI também tem sido utilizada em diversas outras áreas, tais como, nas ciências administrativas, nas ciências da saúde, na psicologia, nas engenharias, etc. Para executar a análise desse trabalho, foram utilizados os bancos de dados que continham as respostas dos alunos de graduação que participaram da autoavaliação institucional realizada em 2012. As respostas foram dicotomizadas para fins de utilização do Modelo Logístico Unidimensional de Dois Parâmetros (ML2) da TRI. Os dados foram analisados com o pacote *irt* do *Software R*. Os resultados mostraram que todos os itens do instrumento de autoavaliação institucional possuem uma boa qualidade, apresentando um bom poder discriminatório. O instrumento apresentou uma concentração maior de informação à esquerda da escala criada, mostrando que os alunos, em geral, consideram os itens satisfatórios.

Palavras-chave: *Teoria da Resposta ao Item, autoavaliação institucional, Modelo Logístico de Dois Parâmetros.*

Referências

- [1] ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. *Teoria da resposta ao item: conceitos e aplicações*. São Paulo: ABE - Associação Brasileira de Estatística, 2000.

¹ UFSM - Universidade Federal de Santa Maria. Email: fmjunior@smail.ufsm.br

² UFSM - Universidade Federal de Santa Maria. Email: charlessz45@gmail.com

Histórico de área plantada com soja e quantidades de bovinos no estado do mato grosso

Fernanda Rezer de Menezes¹

Afonso Valau de Lima Junior²

Viviane de Senna³

Emanuel Arnoni Costa⁴

Alfredo Skrebsky Cezar⁵

Paulo Santa Pacheco⁶

O elevado desempenho do setor agropecuário no estado do Mato Grosso favoreceu o seu desenvolvimento econômico nos últimos anos, principalmente estimulado pelos incentivos fiscais e a facilidade de créditos. Com esse propósito, o objetivo do presente estudo foi avaliar a evolução de área plantada com soja e quantidade de bovinos no estado do Mato Grosso entre o período de 2000 a 2009. A área plantada de soja, quantidade de cabeças de bovinos e o preço da terra em hectares foram obtidos de dados disponibilizados pela: Conab, IBGE e FGV. Para deflacionar o preço da terra utilizou o IGP-DI- Índice Geral de Preços - Disponibilidade Interna, com ano base de 2009. A análise estatística de correlação de Pearson permitiu verificar o grau de associação das séries. Os resultados indicaram uma alta correlação entre a área de soja plantada ($\rho = 0,78$) e a quantidade de bovinos ($\rho = 0,89$) no período. Ao considerar em média a quantidade de soja plantada por hectare encontrou um valor de 5.819,4 (mil hectares), a quantidade de bovinos foi de 24.333.672,6 (cabeças) e preço médio da terra em hectares de R\$ 5.389,30 referentes aos dez anos considerados. Portanto, o estado do Mato Grosso vem apresentando um aumento gradativo de ambas as atividades, sugerindo adotar medidas mitigadoras que promovam a sustentabilidade aliada ao melhor uso da terra.

Palavras-chave: *pecuária, soja, preço, correlação, Mato Grosso.*

Referências

- [1] COMPANHIA NACIONAL DO ABASTECIMENTO. **Indicadores agropecuários**. Disponível em <http://www.conab.gov.br> (10 Set. 2013).
- [2] FGV: FUNDAÇÃO GETULIO VARGAS. Disponível em <http://portalibre.fgv.br/> (10 Set. 2013).
- [3] FIGUEIREDO, M. C. et al. Relação econômica dos setores agrícolas do Estado do Mato Grosso com os demais setores pertencentes tanto ao Estado quanto ao restante do Brasil. **Revista de Economia Rural**. Rio de Janeiro, vol. 43, nº 03, p. 557-575. 2005. Disponível em <http://www.scielo.br/pdf/resr/v43n3/27746.pdf> (08 Set. 2013).
- [4] INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Séries históricas e estatísticas: Brasil: 1977-2009**. <http://www.ibge.gov.br>. (02 Set. 2013).

¹ UFSM - Pós-graduando de Especialização em Estatística. Email: ferrezer25@hotmail.com

² UFSM - Pós-graduando de Especialização em Estatística. Email: avljunior@yahoo.com.br

³ UFSM - Mestranda de Engenharia da Produção. Email: vivianedsenna@hotmail.com

⁴ UFSM - Doutorando em Engenharia Florestal. Email: emanuelarnonicost@hotmail.com

⁵ UFSM - Doutorando em Medicina Veterinária. Email: alfredospes@hotmail.com

⁶ UFSM - Orientador. Professor do Departamento de Zootecnia. Email: pacheco.dz.ufsm@hotmail.com