

ESTRUTURA DE COVARIÂNCIAS EM DADOS DE MEDIDAS REPETIDAS: UMA APLICAÇÃO A DADOS DE PESAGENS DE ZEBUÍÑOS

ALFREDO RIBEIRO DE FREITAS¹, ALEXANDRA FARIA VIEIRA², PEDRO FERREIRA FILHO³

¹ PESQUISADOR DA EMBRAPA PECUÁRIA SUDESTE, CAIXA POSTAL 339, 13560-970 - SÃO CARLOS - SP, BOLSISTA DO CNPQ

² BACHARELANDO DO CURSO DE ESTATÍSTICA DA UFSCAR, SÃO CARLOS, SP

³ PROF., DO CURSO DE ESTATÍSTICA DA UFSCAR, CAIXA POSTAL 676, 13565-905, SÃO CARLOS, SP

RESUMO: Dados de pesos de bovinos, analisados como medidas repetidas, é de interesse para estabelecer curvas de crescimento, comparar tratamentos em tempos específicos ou ponderados sobre o tempo, estimar parâmetros genéticos por regressão aleatória, pois as estruturas repetidas consideram as oscilações de variâncias genéticas e residuais com o tempo. O objetivo deste trabalho foi escolher a estrutura de covariância mais adequada das observações dentro de um mesmo animal, por meio do procedimento MIXED do SAS e critérios de AIC *Akaike's Information Criterion* e SBC *Schwarz's Bayesian Criterion*. Foram analisados dados de nove pesagens, do nascimento até os dois anos de idade, de bovinos da raça Indubrasil, machos e fêmeas, nascidos de 1989 a 1998. A estrutura de variância mais adequada, ou seja, a que possui o maior valor do AIC e do SBC foi a Auto Regressivo Heterogêneo de Primeira Ordem - ARH(1), seguida da Huynh-Feldt e Composto simétrico.

PALAVRAS-CHAVE: bovinos de corte , critérios de informação de Akaike e de Schwarz, dados de pesos , procedimento MIXED

(The authors are responsible for the quality and contents of the title, abstract and keywords)

COVARIANCES STRUCTURES IN REPEATED MEASURES DATA: AN APPLICATION TO WEIGHING DATA OF ZEBU CATTLE

ABSTRACT : Cattle's weight data analysed as repeated measures are importante for estimating growing curve, to compare treatments in specific time or weighed on time and estimate genetic parameters by random regresion because repeated structures considers the oscilations of genetic and residuals variance on time. The objective of this work was to modelate the covariance structure considering measures at different ages on the same animal. Data from nine weighing, from birth to 24 months of age, of Indubrasil purebreed cattle, male and female, birth from 1989 to 1998 were analised. It was used the SAS MIXED procedures and Akaike's Information Criterion and SBC Schwarz's Bayesian Criterion. The selected covariance structure was the Auto Regressive Heterogeneous of first order - ARH(1), folowed by Huynh-Feldt and compound symmetric (CS).

KEY WORDS: Akaike's Information Criterion , beef cattle , MIXED procedure , Schwarz's Bayesian Criterion, weight data

INTRODUÇÃO

Na pesquisa agropecuária é comum o uso de dados medidas repetidas, em que os indivíduos são classificados em diferentes tratamentos, e avaliados em diversas condições de avaliação ou tempo. A análise de dados do desenvolvimento ponderal de bovinos, por exemplo, na estrutura de medidas repetidas, tem várias finalidades (KESELMAN et al 1998; LITTEL et al. 1998): estabelecer curvas de crescimento; estimar parâmetros genéticos, comparar tempo dentro de um determinado tratamento, comparar tratamentos em tempos específicos ou ponderados sobre o tempo. No caso de bovinos, as avaliações de pesagens sobre o mesmo animal requer atenção especial, pois: a) as medidas repetidas são irregulares no

tempo - o intervalo entre duas medidas consecutivas quaisquer não é constante; b) possuem estrutura incompleta - possuem observações perdidas; c) variabilidade crescente na resposta dos indivíduos em função do tempo; d) não são balanceadas com relação ao tempo - as observações não são feitas com os animais possuindo a mesma idade. Os modelos usualmente utilizados para analisar de medidas repetidas são os lineares mistos, os quais requerem que a estrutura de covariância das observações de um mesmo indivíduo seja adequadamente escolhida. O procedimento MIXED do SAS disponibiliza ao usuário diferentes estruturas de matriz de covariâncias (WOLFINGER, 1993), além de permitir selecionar a mais adequada. O objetivo foi escolher a estrutura de covariância mais adequada das observações dentro de um mesmo animal por meio do procedimento MIXED do SAS e critérios de AIC e SBC (BOZDOGAN, 1987). Foram analisados dados de nove pesagens, do nascimento até os dois anos de idade, de bovinos da raça Indubrasil, machos e fêmeas, nascidos de 1989 a 1998 .

MATERIAL E MÉTODOS

Foram analisados dados de nove pesagens, do nascimento até os dois anos de idade, de bovinos da raça Indubrasil, machos e fêmeas, nascidos de 1989 a 1998, pertencentes ao Arquivo Zootécnico Nacional das raças zebuínas. O modelo matemático foi $y_{ijk} = m + S_i + A_{ij} + M_k + (SM)_{ik} + e_{ijk}$, em que y_{ijk} é k-ésimo peso ou condição de avaliação do j-ésimo animal aleatório no i-ésimo sexo; m é o efeito médio, S_i é o efeito i-ésimo sexo; A_{ij} é o efeito do j-ésimo animal dentro do i-ésimo sexo; M_k é o efeito da k-ésima medida; $(SM)_{ik}$ é o efeito de interação de sexo e medida e e_{ijk} , o erro aleatório suposto normalmente distribuído, com média zero, identicamente distribuído e com variância constante. Dentre as estruturas de covariâncias (SEARLE et al. 1992; WOLFINGER, 1993; LITTELL al. 1996,), optou-se por três: Composto Simétrico (CS) - todos os níveis de tempo têm a mesma variância e as covariâncias entre dois quaisquer níveis de tempo são iguais; Huynh-Feldt (HF) - tem heterogeneidade não estruturada ao longo da diagonal principal, esta estrutura constrói os elementos fora da diagonal calculando uma média aritmética; Auto Regressivo Heterogêneo de Primeira Ordem: ARH(1) - Esta estrutura heterogênea é uma generalização direta da estrutura básica AR(1), ela é obtida por permitir variâncias distintas ao longo da diagonal principal. A estrutura de covariância mais apropriada será selecionada por meio dos critérios de AIC (*Akaike's Information Criterion*) e SBC (*Schwarz's Bayesian Criterion*) e uso do procedimento MIXED do SAS. A escolha das estruturas de covariâncias foi feita considerando-se a maximização dos critérios de informação da REML: $AIC = FV - q$ e $SBC = FV - (q/2) \log(n - p)$, em que FV é a função de verossimilhança por REML; q é o número de parâmetros de covariância, n é o número de observações e p é o posto da matriz de planejamento X. As opções do MIXED (SAS, 1996), para o presente trabalho foram:

| | |
|------------|--------------------------------------|
| Proc Mixed | Metodo=REML; |
| Class | Sexo Animal Medida ; |
| Model | Peso = Sexo Medida Sexo*Medida; |
| Random | Animal(sexo); |
| Repeated | Medida /Sub= Animal(Sexo) Type = cs; |

A opção RANDOM determina a estrutura de variância da variação aleatória entre animal e considera as avaliações no mesmo animal independentes; a opção REPEATED, especifica a estrutura de covariância dentro das avaliações no mesmo animal, ou seja, considera a dependência entre os e_{ijk} ; a opção TYPE especifica o tipo de estrutura de covariância a ser determinada (G = CS, no exemplo). No PROC MIXED, REML é o método de estimação default para se estimar os componentes de variância.

RESULTADOS E DISCUSSÃO

As estimativas dos parâmetros das estruturas de covariâncias do tipo CS, HF e ARH(1) obtidas por REML e procedimento MIXED estão nas Tabelas 1. A Tabela 2 fornece os resultados dos testes de AIC (*Akaike's Information Criterion*) e SBC (*Schwarz's Bayesian Criterion*) segundo as estruturas de covariâncias adotada. Verificou-se que o processo numérico para obtenção das estimativas de variâncias REML convergiu na quarta iteração para a estrutura CS, na quinquagésima iteração para a estrutura HF e na vigésima iteração para estrutura ARH(1).

Observando-se a Tabela 2 pode-se concluir que a estrutura de variância mais adequada, ou seja, a que possui o maior valor do AIC e do BIC, é a estrutura de G = ARH(1), seguida da estrutura HF e CS. Observando-se a estrutura ARH(1) na Tabela 1, verifica-se que as variâncias crescem em função da idade ou tempo, comprovando-se o fenômeno da inflação de variância. Pode-se inferir também por meio da Tabela 1 que em dados de medidas repetidas de bovinos, as observações adjacentes são mais altamente

correlacionadas do que as demais. Portanto, esta estrutura de variância, quando considerada em um modelo misto e sob o enfoque de análise de dados de medidas repetidas, proporciona estimativas de parâmetros genéticos mais eficientes em dados de pesagens de bovinos.

CONCLUSÕES

A estrutura de covariância mais adequada para modelar a variabilidade das pesagens dentro de um mesmo animal é a Auto Regressivo Heterogêneo de Primeira Ordem. Essa estrutura mostra que as variâncias crescem em função da idade ou tempo.

REFERÊNCIAS BIBLIOGRÁFICAS

- BOZDOGAN, H. 1987. Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. *Psychometrika.*, 52(3):345-370
- KESELMAN, H.J.; ALGINA, J.; KOWALCHUK, R. K. et al . , 1998. A comparison of two approaches for selecting covariance structures in the analysis of repeated measurements. *Communication in Statistics, New York.*, 27(3):591– 604.
- LITTELL, R.C.; HENRY, P.R.; AMMERMAN, C.B. 1998 . *Statistical analysis of repeated measures data using SAS procedures.* J. Anim. Sci., 76:1216-1231
- LITTELL, R.C.; MILLIKEN, G..A .; STROUP, W.W.; WOLFINGER, R.D. 1996. *SAS System for Mixed Models.* Cary: Statistical Analysis System Institute, 633p.
- PERRI, S.H.V. . *Ajuste de modelos mistos de classificação dupla: procedimentos do sistema estatístico SAS:* Piracicaba, SP, USP, 1999. 126p Tese (Doutorado em Agronomia) – Universidade de São Paulo, 1999.
- SAS INSTITUTE. 1996 . *Advanced general linear models with an emphasis on mixed models. Course notes.* Cary: Statistical Analysis System institute. 614p.
- WOLFINGER, R. 1993 . Covariance structure selection in general mixed models. *Commun. Statist. Simula.*, 22(4):1079-1106

TABELA 1 - Estimativas de parâmetros por REML das estruturas de covariâncias: Huynh-Feldt (HF), Auto Regressivo Heterogêneo de Primeira Ordem: ARH(1) e simetria Composto (CS).

| G= HF | | G = ARH(1) | | G = CS | |
|--------------|------------|--------------|------------|--------------|------------|
| Parâmetros | Estimativa | Parâmetros | Estimativa | Parâmetros | Estimativa |
| σ_1^2 | 20,15 | σ_1^2 | 33,74 | σ_1 | 787,17 |
| σ_2^2 | 2121,67 | σ_2^2 | 1033,05 | σ_1^2 | 1054,77 |
| σ_3^2 | 2151,79 | σ_3^2 | 1293,90 | | |
| σ_3^2 | 2151,59 | σ_4^2 | 1564,69 | | |
| σ_4^2 | 2149,30 | σ_5^2 | 2001,95 | | |
| σ_5^2 | 2137,98 | σ_6^2 | 2403,60 | | |
| σ_6^2 | 2135,91 | σ_7^2 | 3048,16 | | |
| σ_7^2 | 2135,65 | σ_8^2 | 3744,35 | | |
| σ_8^2 | 2126,06 | σ_9^2 | 5537,47 | | |
| HF | 1061,89 | ρ | 0,70 | | |

TABELA 2 – Resultado dos testes de AIC e BIC de acordo com estrutura da Estrutura de covariância adotada .

| Tipo de Estrutura de covariância | Q | Valor do AIC | Valor do BIC |
|----------------------------------|----|--------------|--------------|
| CS | 2 | -112693 | -112699 |
| HF | 10 | -105148 | -105179 |
| ARH(1) | 10 | -101834 | -101866 |

q = número de parâmetros