



## *VII Simpósio Brasileiro de Melhoramento Animal* *São Carlos, SP, 10 e 11 de julho de 2008*

### **Modelos de regressão aleatória na estimação de componentes de variância para o tamanho de leitegada de suínos<sup>1</sup>**

Davi Nogueira Maciel Alves<sup>2</sup>, Jeffrey Frederico Lui<sup>3</sup>, Fabiano Veraldo da Costa Pita<sup>4</sup>,  
Lúcia Galvão de Albuquerque<sup>5</sup>

<sup>1</sup>Parte da dissertação de mestrado do primeiro autor, com bolsa do CNPq.

<sup>2</sup>Doutorando do Programa de Pós-graduação em Genética e Melhoramento Animal – UNESP/Jaboticabal. Bolsista Fapesp. e-mail: [davigenetic@yahoo.com.br](mailto:davigenetic@yahoo.com.br)

<sup>3</sup>Professor Doutor - FCAV - UNESP/Jaboticabal. e-mail: [jeffrey@fcav.unep.br](mailto:jeffrey@fcav.unep.br)

<sup>4</sup>Doutor em Genética e Melhoramento Animal. Pesquisador Newsham Genetics EUA. e-mail:

[fabiano\\_pita@newsham.com](mailto:fabiano_pita@newsham.com)

<sup>5</sup>Professora Adjunto - FCAV - UNESP/Jaboticabal. Pesquisadora do CNPq. e-mail: [lgalb@fcav.unep.br](mailto:lgalb@fcav.unep.br)

**Resumo** – O objetivo deste trabalho foi estimar componentes de variância para o número de leitões nascidos vivos empregando modelos de regressão aleatória (MRA), e observar divergências entre os valores estimados com os de um modelo multi-característica. Os diferentes MRA foram comparados pelo teste da razão de verossimilhança (LRT) e pelo Critério de Informação de Akaike (AIC). As herdabilidades estimadas tiveram valores entre 0,02 e 0,16. As estatísticas de ajuste dos modelos indicaram o modelo com homogeneidade de variâncias genética e de ambiente permanente com três classes de variâncias residuais como mais adequado.

**Palavras-chave:** características reprodutivas, herdabilidade, prolificidade.

#### **Random regression models to estimate of variance components of litter size in pigs.**

**Abstract** – The aim of this study was to estimate variance components for the number of piglets born alive by random regression models (RRM), and observe differences a multiple-trait model. To compare the different models the Likelihood ratio test (LRT) and Akaike's Information Criterion (AIC) were utilized. The heritability estimates ranged from 0.02 to 0.16. The model with homogeneous genetic and permanent environmental variances and residual variance modeled as a step function with three classes, presented the best fit.

**Keywords:** heritability, prolificity, reproductive traits.

#### **Introdução**

Modelos de repetibilidade e multi-características têm sido utilizados na estimação parâmetros genéticos para o tamanho de leitegada em suínos. No entanto, o primeiro modelo assume que cada parição é a mesma característica, ou seja, as correlações genéticas entre as partições são iguais a um. O segundo considera cada parição como uma característica distinta, levando em conta as correlações genéticas entre partições, porém sem empregar nenhuma estrutura para modelar as (co)variâncias

no decorrer das parições, além disto, o número de parâmetros aumenta de forma considerável a medida que novas parições são incluídas na análise.

Lukovic et al. (2004) empregaram modelos de regressão aleatória para o número de leitões nascidos vivos, tomando como função contínua da ordem do parto, o que pode ser uma alternativa aos modelos de repetibilidade e multi-características, pois ao se utilizar modelos de regressão aleatória, implicitamente se adota uma estrutura para modelar covariâncias. Na literatura ainda são escassos trabalhos que empregam esta metodologia para estimar parâmetros genéticos para prolificidade em suínos. O objetivo deste estudo foi estimar parâmetros genéticos para o número total de leitões nascidos vivos (NLNV), através de modelos de regressão aleatória (MRA), bem como observar as divergências desta metodologia com as estimativas obtidas por um modelo multi-característica (MMC).

### Material e Métodos

Os dados utilizados neste trabalho são provenientes de duas linhagens maternas Newsham<sup>®</sup>. Foram registrados entre janeiro de 2004 e setembro de 2006, em duas granjas núcleo americanas localizadas uma no estado de Wyoming e a outra no Kansas. Do arquivo original só foram consideradas leitegadas com no mínimo 1 e no Máximo 20 leitões nascidos vivos, grupos de contemporâneos com no mínimo 5 observações, dados até a sexta parição e animais com no mínimo duas observações, de forma que foram incluídas nas análises 5450 leitegadas.

Como efeitos fixos foram considerados a linhagem da matriz e o grupo de contemporâneos, formado pela combinação do rebanho, ano e época em que ocorreu a parição (jan-mar, época 1; abr-jun, 2; jul-set, 3; out-dez, 4). A trajetória média da população foi modelada por meio de uma regressão quadrática sobre polinômios ortogonais de Legendre.

Inicialmente, foram ajustados três modelos com diferentes estruturas variâncias residuais considerando: homogêneas entre as diferentes parições; agrupando-as em três classes, (Classe 1, parição 1; Classe 2, parição 2; e Classe 3, as demais parições); e em seis classes, referentes a cada parição. Nesta etapa, foi considerada uma regressão quadrática sobre polinômios de Legendre para modelar a variação genética e de ambiente permanente. Escolhida a estrutura de variâncias residuais, foram empregados quatro modelos de diferentes ordens de ajuste, sobre polinômios ortogonais de Legendre, variando de linear (ordem = 2) à cúbica (ordem = 4), bem como um modelo de repetibilidade (ordem = 1).

O modelo de regressão aleatória pode ser representado na forma matricial por:

$$y = X\beta + Z_1a + Z_2p + \varepsilon \quad (1)$$

Assumindo-se que:  $\text{Var}(a) = K_a \otimes A$ ;  $\text{Var}(p) = K_p \otimes I$ ;  $\text{Var}(\varepsilon) = R$ ;  $E(y) = X\beta$  e  $E(a) = E(p) = E(\varepsilon) = 0$ . Onde:  $y$  é um vetor de observações;  $\beta$  vetor de efeitos fixos, incluindo os coeficientes de regressão fixos;  $a$  e  $p$  são o vetor de coeficientes de regressão aleatória genéticos e de ambiente permanente, respectivamente;  $\varepsilon$  vetor de erros;  $X$ ,  $Z_1$  e  $Z_2$ , são matrizes de incidência correspondentes aos vetores.  $K_a$  e  $K_p$  são as matrizes de (co)variâncias entre os coeficientes de regressão aleatória para os efeitos genéticos e de ambiente permanente, respectivamente, e  $R$  é uma matriz contendo as variâncias residuais;  $A$  é a matriz de parentesco; e  $\otimes$  é o operador de produto direto.

Para as análises multi-características o modelo é similar ao empregado na regressão aleatória, desconsiderando-se o efeito de ambiente permanente. Nas pressuposições  $K_a$  é substituída por  $G_\theta$  que representa a matriz de (co)variâncias genéticas entre parições.

As análises empregando modelos de regressão aleatória foram realizadas com o programa DXMRR e para a análise multi-características, o programa DXMUX, ambos disponíveis no pacote DFREML (MEYER, 1998).

A escolha dos diferentes modelos foi realizada com base no Teste da Razão de Verossimilhança (LRT), no Critério de Informação de Akaike (AIC) e no Critério de Informação Bayesiano (BIC)

### Resultados e Discussão

O teste da razão de verossimilhança (LRT) e o AIC indicaram que o modelo com três classes de variâncias residuais foi o que apresentou melhor ajuste, não diferindo do modelo com seis classes quando utilizado o LRT (Tabela 1). Com base nisso adotaram-se três classes variâncias residuais nos demais modelos empregados. Devido à maior penalidade imposta, o Critério de Informação Bayesiano (BIC), indicou como melhor modelo o que considerava as variâncias residuais homogêneas.

Tabela 1. Número de parâmetros (Np), logaritmo de verossimilhança (Log L), valores de AIC e BIC, e valor probabilidade (Valor – p<sup>b</sup>) do teste da razão da verossimilhança, para os modelos com diferentes classes de variâncias residuais (*e*).

Modelo	k <sub>a</sub>	k <sub>p</sub>	<i>e</i>	Np	Log L	AIC	BIC	LRT	
								$M_i - M_{i+j}^a$	Valor – p
1	3	3	1	13	-8879.78	17785.55	17871.40	1-2	< 0.01
2	3	3	3	15	-8873.28	17776.55	17875.60	1-3	0.02
3	3	3	6	18	-8873.18	17782.36	17901.22	2-3	0.98

k<sub>a</sub> e k<sub>p</sub> – ordem do polinômio para o efeito genético e de ambiente permanente, respectivamente;

<sup>a</sup>Comparação do modelo  $M_i$  com *i* parâmetros, com  $M_{i+j}$  com *j* parâmetros adicionais;

<sup>b</sup>Calculado com *j* graus de liberdade.

Para os modelos empregados para modelar a variação genética (Tabela 2), tanto para o LRT e AIC, o modelo de repetibilidade (ordem 1) apresentou melhor ajuste, seguido pelo modelo de ordem 2 (linear).

Tabela 2. Número de parâmetros (Np), logaritmo de verossimilhança (Log L), valores de AIC e BIC, e valor probabilidade do teste da razão da verossimilhança, para os modelos ajustados para número de leitões nascidos vivos.

Modelo	k <sub>a</sub>	k <sub>p</sub>	<i>e</i>	Np	Log L	AIC	BIC	LRT	
								$M_i - M_{i+j}^a$	Valor-P <sup>b</sup>
1	1	1	3	5	-8877.70	17765.39	17798.41	1-2	0,49
2	2	2	3	9	-8875.99	17769.98	17829.41	1-3	0,55
3	3	3	3	15	-8873.28	17776.55	17875.60	1-4	0,69
4	4	4	3	23	-8870.40	17786.80	17938.68	2-3	0,49

k<sub>a</sub> e k<sub>p</sub> – ordem do polinômio para o efeito genético e de ambiente permanente, respectivamente; *e* número de classes residuais consideradas; <sup>a</sup>Comparação do modelo  $M_i$  com *i* parâmetros, com  $M_{i+j}$  com *j* parâmetros adicionais; <sup>b</sup>Calculado com *j* graus de liberdade.

As estimativas herdabilidade ( $h^2$ ), obtidas por meio do modelo de repetibilidade (ordem 1), com três classes de variâncias residuais, foram de 0,082 para a primeira parição, 0,070, para segunda e 0,079 para as demais. Estas pequenas diferenças são oriundas das diferentes estimativas de variâncias residuais para cada classe de parições adotada.

Os parâmetros genéticos estimados por meio do MRA-Linear e do MMC estão dispostos na Tabela 3. As  $h^2$  estimadas pelo MMC variaram de 0,02 a 0,16 para sexta e

terceira parição respectivamente, já pelo MRA-Linear houve uma diminuição do valor de  $h^2$  da primeira até sexta parição. Com o MRA-Linear as correlações genéticas foram elevadas entre parições adjacentes tendendo a diminuir entre parições mais afastadas. Nenhum padrão entre correlações foi observado com o MMC, provavelmente devido ao fato do MMC não adotar estrutura de covariâncias no decorrer das parições e pelo pequeno número de observações utilizados na análise.

Tabela 3. Estimativas, obtidas por modelos Multi-características (MMC) e de regressão aleatória (MRA), de variância genética (a) e de ambiente permanente (p), herdabilidades (diagonal, negrito), correlações genéticas (acima da diagonal), correlações de ambiente permanente (MRA) (abaixo da diagonal). e variâncias residuais (e), para o número de leitões nascidos vivos.

Parâmetro/ parição	1	2	3	4	5	6	1	2	3	4	5	6
	MMC						MRA (Ordem=2)					
1	<b>0,15</b>	0,72	0,79	0,98	0,26	0,71	<b>0,13</b>	0,99	0,96	0,88	0,73	0,54
2		<b>0,09</b>	0,81	0,84	0,33	0,87	0,99	<b>0,09</b>	0,99	0,93	0,82	0,64
3			<b>0,16</b>	0,80	0,75	0,98	0,96	0,99	<b>0,08</b>	0,98	0,90	0,76
4				<b>0,10</b>	0,21	0,75	0,90	0,96	0,99	<b>0,06</b>	0,97	0,88
5					<b>0,14</b>	0,74	0,84	0,91	0,96	0,99	<b>0,06</b>	0,97
6						<b>0,02</b>	0,77	0,86	0,92	0,97	0,99	<b>0,06</b>
a	1,34	0,98	1,54	0,96	1,36	0,18	1,21	0,94	0,74	0,61	0,54	0,55
p	-	-	-	-	-	-	0,86	0,84	0,85	0,91	1,00	1,13
e	7,86	9,93	8,04	8,41	8,14	9,22	7,23	9,15	7,97	7,97	7,97	7,97

Tanto com MMC como com MRA, algumas correlações entre parições apresentaram valores inferiores a 0,80 indicando um controle genético distinto em cada parição. No entanto, o modelo de repetibilidade com três classes de variâncias residuais apresentou melhor ajuste em relação aos demais, o que validaria o seu uso. É importante salientar a presença de variâncias residuais heterogêneas entre parições, o que pode influenciar na estimativa de valores genéticos dos animais.

### Conclusões

É importante considerar a heterogeneidade de variâncias residuais entre parições. Para o número de leitões nascidos vivos o modelo que adota homogeneidade de variâncias genéticas e de ambiente permanente apresentou melhor ajuste.

### Agradecimentos

À empresa Newsham Genetics, LC. Pela concessão dos dados que possibilitou a execução deste trabalho; ao CNPq pela concessão da bolsa de estudos ao primeiro autor e à FCAV-UNESP pela oportunidade de ingressar no Programa de Pós-graduação em Genética e Melhoramento Animal.

### Literatura Citada

- LUKOVIC, Z.; MALOVRH, S.; GORJANC, G. et al. A random regression model in analysis off litter size in pigs. **South African Journal of Animal Science**, v.34, n.4, p.241-247, 2004.
- MEYER, K. **DFREML - Version 3.0 b - User Notes**, 1998.