Generalized linear mixed models for the genetic evaluation of binary reproductive traits: a simulation study[1]

Diogo Anastácio Garcia[2], Idalmo Garcia Pereira[3], Fabyano Fonseca e Silva[4], Guilherme Jordão de Magalhães Rosa[5], Aldrin Vieira Pires[3], Roseli Aparecida Leandro[6]

[1] Project financed by FAPEMIG.
[2] Faculdade de Ciências Agrárias e Veterinárias/UNESP – Jaboticabal, SP.
[3] Departamento de Zootecnia – FCA/UFVJM – Diamantina, MG.
[4] Departamento de Estatística – UFV – Viçosa, MG.
[5] Department of Dairy Science, University of Wisconsin – Madison, WI, USA.
[6] Departamento de Ciências Exatas – LCE/ESALQ – Piracicaba, SP.

ABSTRACT - The objective of this study was to evaluate the use of probit and logit link functions for the genetic evaluation of early pregnancy using simulated data. The following simulation/analysis structures were constructed: logit/logit, logit/probit, probit/logit, and probit/probit. The percentages of precocious females were 5, 10, 15, 20, 25 and 30% and were adjusted based on a change in the mean of the latent variable. The parametric heritability (h^2) was 0.40. Simulation and genetic evaluation were implemented in the R software. Heritability estimates ($\hat{h}^2$) were compared with h^2 using the mean squared error. Pearson correlations between predicted and true breeding values and the percentage of coincidence between true and predicted ranking, considering the 10% of bulls with the highest breeding values (TOP10) were calculated. The mean $\hat{h}^2$ values were under- and overestimated for all percentages of precocious females when logit/probit and probit/logit models used. In addition, the mean squared errors of these models were high when compared with those obtained with the probit/probit and logit/logit models. Considering $\hat{h}^2$, probit/probit and logit/logit were also superior to logit/probit and probit/logit, providing values close to the parametric heritability. Logit/probit and probit/logit presented low Pearson correlations, whereas the correlations obtained with probit/probit and logit/logit ranged from moderate to high. With respect to the TOP10 bulls, logit/probit and probit/logit presented much lower percentages than probit/probit and logit/logit. The genetic parameter estimates and predictions of breeding values of the animals obtained with the logit/logit and probit/probit models were similar. In contrast, the results obtained with probit/logit and logit/probit were not satisfactory. There is need to compare the estimation and prediction ability of logit and probit link functions.

Key Words: beef cattle, early pregnancy, genetic parameters, GLMM, link function

Introduction

One of the factors that determine the economic viability of sustainable beef cattle enterprises is the reproductive performance of the herd. Therefore, the identification and evaluation of reproductive traits that can be easily measured and present a potential for selection are essential. In this respect, early pregnancy is an easily measured trait since it comprises only two categories (pregnant and non-pregnant). In addition, the evaluation of phenotypic expression does not result in additional costs, since the diagnosis of pregnancy is a routine practice in beef cattle farming.

However, one of the main challenges of using categorical traits in breeding programs is the development of adequate statistical methods for the estimation of parameters and the prediction of breeding values. One of the methods frequently used for this purpose is generalized linear mixed models (GLMM) (Thompson, 1979; Gianola & Foulley, 1983;

Harville & Mee, 1984; Gilmour et al., 1985, 1987; Tempelman, 1998; Pereira et al., 2006, 2007). According to McCullagh & Nelder (1989), these models can be defined based on the specification of three components: i) a random component represented by independent random variables that belong to the same distribution, which is part of an exponential family such as binomial and Poisson; ii) a systematic component, called linear predictor, in which explanatory variables enter in the form of the linear sum of their effects; iii) a link function that combines the random and systematic components, i.e., relates the mean to the linear predictor and thus permits modeling, for example, the probability of a female being precocious.

The link function for binary data most widely used in animal breeding is the probit link, also known as the threshold model (Gianola & Foulley, 1983; Abdel-Azim & Berger, 1999; Kadarmideen et al., 2000, 2001; Silva et al., 2005; Shiotsuki et al., 2009) Other functions can also be

explored, such as the logit link, which is widely used in the area of Biometrics (Breslow & Clayton, 1993; Demétrio, 2001; Nunes et al., 2004). When modeling a random variable using probit and logit functions, it is assumed that the latent variable (liability) possesses a standard normal and logistic distribution, respectively. Although important, studies comparing these functions are scarce and are based on the analysis of true data (Gilmour et al., 1987; Kadarmideen et al., 2004). Therefore, an interesting alternative to evaluate the estimation and prediction ability of probit and logit link functions is the use of simulated data.

The objective of the present study was to evaluate the application of a GLMM to the estimation of genetic parameters and prediction of breeding values for early pregnancy by comparing probit and logit link functions using simulated data.

Material and Methods

The simulated population was structured based on the random mating of 50 bulls to 1500 unrelated cows, resulting in 1500 females derived from 50 groups of 30 half-sibs, which were submitted to the evaluation of early pregnancy.

Fixed and random effects were introduced into the simulation through the specification of a linear predictor given by:

$$\boldsymbol{\eta} = X\boldsymbol{\beta} + Z\boldsymbol{u}, \qquad [1]$$

where $\boldsymbol{\eta}_{px1}$ is the linear predictor and p is the number of heifers (1500); $\boldsymbol{\beta}_{3x1}$ is the vector of fixed effects, $\boldsymbol{\beta}' = [\mu, \beta_1, \beta_2]$, whose values adopted for the simulation were $\beta_1 = 1$ and $\beta_1 = -1$, and μ is the overall mean (Table 2); $\boldsymbol{u}_{px1}$ is the vector of additive genetic values, with $\boldsymbol{u} \sim N_p(\boldsymbol{0}, A\sigma_u^2)$, where A is the relationship matrix between animals and σ_u^2 is the additive genetic variance of the population; X_{px3} and X_{pxp} are incidence matrices of effects β and u, respectively.

Considering that $\pi_i = P[Y_i = 1]$ is the probability of heifer i ($i = 1, 2, ..., 1500$) to be precocious, the logit and probit link functions are given by expressions [2] and [3], respectively:

$$g(\pi_i) = log\left(\frac{\pi_i}{1 - \pi_i}\right) \qquad [2]$$

$$g(\pi_i) = \Phi^{-1}(\pi_i), \qquad [3]$$

where $\Phi^{-1}(.)$ is the inverse function of the standard normal cumulative distribution.

According to the theory of GLMM, the probabilities of interest (π_i) can be isolated by matching the linear predictor [1] to the link functions [2] and [3]. Considering the logit and probit link functions, these probabilities are given by expressions [4] and [5]:

(i) $g(\pi_i) = log\left(\frac{\pi_i}{1 - \pi_i}\right) = \eta_i = x_i'\boldsymbol{\beta} + u_i$

$$log\left(\frac{\pi_i}{1 - \pi_i}\right) = x_i'\boldsymbol{\beta} + u_i$$

$$\frac{\pi_i}{1 - \pi_i} = exp(x_i'\boldsymbol{\beta} + u_i) \Rightarrow \pi_i = exp(x_i'\boldsymbol{\beta} + u_i) - \pi_i\, exp(x_i'\boldsymbol{\beta} + u_i)$$

$$\pi_i + \pi_i\, exp(x_i'\boldsymbol{\beta} + u_i) = exp(x_i'\boldsymbol{\beta} + u_i) \Rightarrow \pi_i[1 + exp(x_i'\boldsymbol{\beta} + u_i)] = exp(x_i'\boldsymbol{\beta} + u_i)$$

$$\pi_i = \frac{exp(x_i'\boldsymbol{\beta} + u_i)}{1 + exp(x_i'\boldsymbol{\beta} + u_i)} \qquad [4]$$

(ii) $g(\pi_i) = \Phi^{-1}(\pi_i) = \eta_i = x_i'\boldsymbol{\beta} + u_i$

$$\Phi^{-1}(\pi_i) = x_i'\boldsymbol{\beta} + u_i$$

$$\pi_i = \Phi(x_i'\boldsymbol{\beta} + u_i) \qquad [5]$$

where $\Phi(.)$ is the standard normal distribution.

Once parametric values are assumed for β and σ_u^2, the probabilities of early pregnancy can be derived from [4] and [5] and are then used in the Monte Carlo simulation to generate phenotypic expressions of the early pregnancy trait. These expressions correspond to values 0 and 1, which are generated from binomial distributions, with the probability of success being given by $P[Y_i = 1] = \pi_i$.

According to the method used, the source of variation not explained by the model in the generation of phenotypic expressions, which may be called residual, is implicit in the randomness of the values generated for Y_i by the binomial distributions with probability π_i. Therefore, when a probit link function is used, the residual variance (σ_{ep}^2), since it is unknown, is generally assumed to be 1. The objective of this value is to render the likelihood function, which is constructed based on the approach of probit regression, identifiable (Sorensen & Gianola, 2002). For the logit function, the residual variance is an approximation based on the variance of the logistic distribution (Southey et al., 2003) and is described by expression [6]:

$$\sigma_{el}^2 = \frac{3.1416^2 b^2}{3} = 3.2899 \qquad [6]$$

where b is the standard deviation of the logistic distribution, which is generally assumed to be 1 (Southey et al., 2003), as in the present study.

The parametric heritability used in the simulation was 0.40 and reflects the results of recent studies involving Nellore animals, which reported moderate to high estimates (Eler et al., 2002; Silva et al., 2003; Eler et al., 2004; Silva et al., 2005; Pereira et al., 2007; Shiotsuki et al., 2009). Different additive genetic variances were employed according to the link function to be used in the analysis of the simulated data. Thus, expressions [7] and [8] can be used to calculate the respective additive genetic variances

for the logit (σ^2_{ul}) and probit (σ^2_{up}) link functions:

$$\sigma^2_{ul} = \frac{h^2 \sigma^2_{el}}{1 - h^2} = \frac{0.4 \sigma^2_{el}}{1 - 0.4} = 2.1932 \qquad [7]$$

$$\sigma^2_{up} = \frac{h^2 \sigma^2_{ep}}{1 - h^2} = \frac{0.4 \sigma^2_{ep}}{1 - 0.4} = 0.6667 \qquad [8]$$

The simulation was structured based on the following components (Table 1): i) inverse of the link function (expressions [4] and [5]); ii) link function used for estimation and prediction (expressions [2] and [3]).

One practical aspect that can influence the estimation of genetic parameters and prediction of breeding values for categorical traits when GLMM are used is the variation in the frequencies of observations for each category (Abdel-Azim & Berger, 1999). For this reason, different percentages of precocious females were adopted in the simulation since the results obtained with the logit and probit functions may vary depending on these percentages, which were modeled according to the values assumed for μ (Table 2). The following percentages of precocious females were adopted: 5, 10, 15, 20, 25, and 30%.

The different simulation scenarios were defined by combinations of the distinct models (logit/logit, probit/logit, logit/probit and probit/probit) and percentages of precocious females. This simulation process was implemented in the R software (R Development Core

Team, 2009), considering 100 repetitions per scenario. The *mvrnorm* function of the MCMCpack was used to generate additive genetic values (u) in [1] of a multivariate normal distribution. The numerator relationship matrix A, which comprises the covariance matrix of this distribution, was obtained with the *kinship* function of the same package. The *rbinom* function was used to determine to which fixed effect of the female it would be submitted, with a probability of 0.5 of the animal being under the effect of β_1 or β_2. Values of η_i were then determined based on the sum of fixed effects and the breeding value ($x'_i \beta + u_i$). These values of η_i were used to calculate the probabilities P[$Y_i = 1$] $= \pi_i$ derived from the inverse link functions (expressions [4] and [5]) through a binomial function, considering the specifications *(link = "probit")$linkinv* and *(link = "logit")$linkinv* for probit and logit links, respectively. These probabilities corresponded to the parameters of the binomial distribution necessary for the generation of binary data obtained with the *rbinom* function.

The data sets generated for each scenario, corresponding to 100 repetitions, were analyzed by the GLMM method using the pedigreemm package (Vazquez et al., 2009) of the R software (R Development Core Team, 2009). This package uses the restricted maximum likelihood method for the estimation of variance components. In the present study, due to the adoption of a sire model, the sire

Table 1 - Simulation/analysis structures used according to the inverse function applied to obtain the probabilities of early pregnancy in heifers (π_i) and the link function used in the analysis

Simulation/Analysis[1]	Link function		σ^2_u	u	Predictor	π
	Inverse	Analysis				
LL	logit				η_{L1}	π_{LL}
		logit	σ^2_{ul}	u_l		
PL	probit				η_{L2}	π_{PL}
LP	logit				η_{P1}	π_{LP}
		probit	σ^2_{up}	u_p		
PP	probit				η_{P2}	π_{PP}

[1] Structures of the simulation/analysis: logit/logit (LL), logit/probit (LP), probit/logit (PL) and probit/probit (PP); σ^2_u is the additive genetic variance; u is the vector of breeding values.

Table 2 - Mean values (μ) used to determine the percentage of precocious females for each simulation/analysis structures

Percentage of precocious females	Simulation/Analysis structures[1]			
	PP	LP	PL	LL
5	-2.40	-3.35	-3.10	-3.95
10	-1.95	-2.60	-2.55	-3.00
15	-1.60	-2.00	-2.00	-2.32
20	-1.41	-1.69	-1.58	-1.98
25	-1.12	-1.40	-1.30	-1.56
30	-0.91	-1.07	-0.96	-1.20

[1] PP: probit/probit; LP: logit/probi; PL: probit/logit; LL: logit/logit.

variance estimates correspond to ¼ of the estimated value for the additive genetic random component.

Isolation of $\hat{h}^2$ from expressions [9] and [10], which correspond to the logit and probit link functions, respectively, was necessary to obtain the estimates of h^2 and to compare them with the parametric value of 0.40:

$$4\hat{\sigma}_t^2 = \frac{\hat{h}^2 \sigma_{el}^2}{1-\hat{h}^2} \Leftrightarrow \hat{h}^2 = \frac{4\hat{\sigma}_t^2}{\hat{\sigma}_t^2 + \sigma_{el}^2} \qquad [9]$$

$$4\hat{\sigma}_t^2 = \frac{\hat{h}^2 \sigma_{ep}^2}{1-\hat{h}^2} \Leftrightarrow \hat{h}^2 = \frac{4\hat{\sigma}_t^2}{\hat{\sigma}_t^2 + \sigma_{ep}^2} \qquad [10]$$

where $\hat{\sigma}_t^2$ is the sire variance.

The heritability estimates obtained were compared with the parametric value using the mean squared error. According to Casela & Berger (1990), it is a function that quantifies the expected value of the squared difference between the estimator ($\hat{h}^2$) and parameter (h^2). The mean squared error is defined as follows:

$$MSE = \frac{1}{N}\left[\sum_{i=1}^{N}\left(\hat{h}_i^2 - h^2\right)^2\right], \qquad [11]$$

where N is the number of repetitions per scenario, in this case 100.

The predictions of additive breeding values were compared with parametric values using Pearson correlations. In addition, the percentage of coincidence between true and predicted ranking, considering only the 10% of bulls with the highest breeding values (TOP10) was measured.

Results and Discussion

The logit/probit and probit/logit simulation/analysis structures yielded mean estimates of h^2 below and above the parametric value, respectively, for all percentages of precocious females (Table 3). In addition, the mean squared errors obtained with these models were high when compared with those obtained with the probit/probit and logit/logit models. Considering only bias, the last two structures also showed a better performance than probit/logit and logit/probit since they provided estimates that were much closer to the parametric value.

The $\hat{h}^2$ values that showed the greatest bias in the probit/probit (0.3658) and logit/logit (0.3554) simulation/analysis structures were obtained at percentages of precocious females of 25 and 30%, respectively. However, these estimates can be considered close to the simulated parametric value (0.40). For these models, the highest mean squared error was obtained at a percentage of precocious females of 5% (0.0478 for probit/probit and 0.0392 for logit/logit), indicating that a lower proportion of early pregnancy success may increase the variation in h^2 estimates.

The logit/probit and probit/logit models presented low correlations (close to zero), indicating a poor ability of prediction (Table 4). In contrast, the probit/probit and logit/logit models showed correlations ranging from moderate to high. The correlations obtained with the logit/logit model were slightly higher than those obtained with the probit/probit model, except for the percentage of precocity of 5% (0.5936 and 0.6025). These differences are probably related to the fact that inadequate likelihood functions are used when the data are analyzed as probit and logit, and vice-versa. It is understood that the use of an inadequate distribution for the data (likelihood functions) may compromise the predictions of breeding values.

Correlations between predicted and true breeding values increased with increasing percentage of precocious females when the probit/probit and logit/logit simulation models were used (Table 4). This behavior can be attributed to the fact that logit and probit link functions yield better results when the proportions of zero and one tend to be the same, i.e., 50%. According to Abdel-Azim & Berger (1999), the quality

Table 3 - Mean heritability estimates ($\hat{h}^2$) and mean squared errors (MSE) obtained with each simulation structure according to the different percentages of precocious females (%PF)

%FP	Simulation/Analysis structures[1]							
	PL		PP		LL		LP	
	$\hat{h}^2$	MSE	$\hat{h}^2$	MSE	$\hat{h}^2$	MSE	$\hat{h}^2$	MSE
5	0.5989	0.0896	0.4118	0.0478	0.4398	0.0392	0.1158	0.0948
10	0.6265	0.0862	0.3823	0.0233	0.4029	0.0275	0.1501	0.0709
15	0.6300	0.0804	0.3753	0.0190	0.3829	0.0138	0.1687	0.0603
20	0.6117	0.0695	0.3952	0.0176	0.3613	0.0167	0.1768	0.0586
25	0.6041	0.0626	0.3658	0.0135	0.3744	0.0165	0.1723	0.0602
30	0.5771	0.0484	0.3886	0.0135	0.3554	0.0112	0.1815	0.0554

[1] LP: logit/probit; PP: probit/probit; LL: logit/logit; PL: probit/logit.

of multicategorical (probit) threshold models increases as the differences between the frequencies of observations for each category decrease. Since the percentages of precocious females adopted in the simulation did not include the value of 50%, the best results were reported for the closest percentage, i.e., 30%. The values adopted here were chosen according to previous studies involving Nellore cattle (Eler et al., 2002; Silva et al., 2003; Eler et al., 2004; Silva et al., 2005; Pereira et al., 2007; Shiotsuki et al., 2009).

The percentages of animals coinciding with the TOP10 bulls obtained with the logit/probit and probit/logit models were lower than those obtained with the probit/probit and logit/logit models (Table 5). These findings suggest that comparison between logit and probit link functions is necessary for the ranking of the best animals, in agreement with the results obtained for the estimation of heritability (Table 3) and for the prediction of additive genetic effects (Table 4). In addition, the most satisfactory results were observed with increasing percentage of precocious females, a fact also demonstrated by Pearson correlations (Table 4).

Taken together, the results obtained with the simulation models proposed show the need to compare the estimation and prediction ability of logit and probit link functions, irrespective of the percentage of precocious females. Simulating data on the logit scale and smoothing them using a probit function (logit/probit), or vice-versa (probit/logit), provided less satisfactory results than those obtained when the data were generated and analyzed using the same scale and a link function (probit/probit and logit/logit).

Table 4 - Mean Pearson correlations between true and predicted breeding values obtained with each simulation/analysis structures for the different percentages of precocious females

Percentage of precocious females	Simulation/Analysis structures[1]			
	PL	PP	LL	LP
5	0.0259	0.6025	0.5936	-0.0289
10	-0.0065	0.6657	0.6878	-0.0261
15	-0.0040	0.7125	0.7353	-0.0109
20	-0.0259	0.7418	0.7623	-0.0248
25	-0.0139	0.7541	0.7809	-0.0350
30	0.0066	0.7828	0.7929	0.0169

[1] LP: logit/probit; PP: probit/probit; LL: logit/logit; PL: probit/logit.

Table 5 - Mean percentage of coincidence in top 10% for true and predicted ranking obtained with each simulation/analysis model for the different percentages of precocious females

Percentage of precocious females	Simulation/Analysis structures[1]			
	PL	PP	LL	LP
5	9.8	42.4	48.0	9.2
10	9.8	46.2	50.4	8.6
15	10.6	53.2	57.2	9.2
20	8.2	51.2	57.2	9.6
25	9.8	54.4	56.8	10.6
30	9.4	53.8	57.8	9.8

[1] LP: logit/probit; PP: probit/probit; LL: logit/logit; PL: probit/logit.

Conclusions

The generalized linear mixed models using logit and probit link functions provided similar genetic parameter estimates and predictions of breeding values of the animals when logit/logit and probit/probit simulation/analysis models were used. However, the results obtained with the probit/logit and logit/probit models were not satisfactory. It is therefore necessary to compare the estimation and prediction ability of link functions in order to determine which function is the most appropriate for the genetic evaluation of binary traits.

Acknowledgements

References

ABDEL-AZIM, G.A.; BERGER, P.J. Properties of threshold model predictions. **Journal of Animal Science**, v.77, p.582-590, 1999.

BRESLOW, N.E.; CLAYTON, D.G. Approximate inference in generalized linear mixed models. **Journal of the American Statistical Association**, v.88, n.421, p.9-25, 1993.

CASELA, G.; BERGER, R.L. **Statistical inference**. Belmont: Duxbury Press, 1990. 660p.

DEMETRIO, C.G.B. **Modelos lineares generalizados em experimentação agronômica**. Piracicaba: ESALQ, 2001. 113p.

ELER, J.P.; SILVA, J.A. II V.; FERRAZ, J.B.S. et al. Genetic evaluation of the probability of pregnancy at 14 months for Nellore heifers. **Journal of Animal Science**, v.80, p.4951-4954, 2002.

ELER, J.P.; SILVA, J.A. II V.; EVANS, J.L. et al. Additive genetic relationships between heifer pregnancy and scrotal circumference in Nellore cattle. **Journal of Animal Science**, v.82, n.9, p.2519-2527, 2004.

GIANOLA, D.; FOULLEY, J.L. Sire evaluation for ordered categorical data with threshold model. **Genetics Selection Evolution**, v.15, n.2, p.201-224, 1983.

GILMOUR, A.R.; ANDERSON, R.D.; RAE, A.L. The analysis of binomial data by generalized linear mixed model. **Biometrika**, v.72, n.3, p.539-599, 1985.

GILMOUR, A.R.; ANDERSON, R.D.; RAE, A.L. Variance components on an underlying scale for ordered multiple threshold categorical data using a generalized linear mixed model. **Journal of Animal Breeding and Genetics**, v.104, p.149-155, 1987.

HARVILLE, D.A.; MEE, R.W. A mixed model procedure for analysing ordered categorical data. **Biometrics**, v.40, p.393-408, 1984.

KADARMIDEEN, H.N., THOMPSON, R.; SIMM, G. Linear and threshold model genetic parameters for disease, fertility and milk production in dairy cattle. **Animal Science**, v.71, p.411-419, 2000.

KADARMIDEEN, H.N.; REKAYA, R.; GIANOLA, D. Genetic parameters for clinical mastitis in Holstein-Friesians: A Bayesian analysis. **Animal Science**, v.73, p.229-240, 2001.

KADARMIDEEN, H.N.; SCHWORER, D.; ILAHI, H. et al. A. Genetics of osteochondral disease and its relationship with meat quality and quantity, growth, and feed conversion traits in pigs. **Journal of Animal Science**, v.82, p.3118-3127, 2004.

McCULLAGH, P.; NELDER, J.A. **Generalized linear models**. 2.ed. London: Chapman & Hall, 1989. 511p.

NUNES, J.A.R.; MORAIS, A.R.; BUENO FILHO, J.S.S. Modelagem da superdispersão em dados por um modelo linear generalizado misto. **Revista de Matemática e Estatística**, v.22, n.1, p.55-70, 2004.

PEREIRA, I.G.; SILVA, F.F.; CARNEIRO, A.P.S. et al. Genetic evaluation of the binary reproductive trait via generalized linear mixed models (GLMM). In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 2006, Belo Horizonte. **Proceedings...** Belo Horizonte, v.8, p.233-233, 2006.

PEREIRA, I.G.; SILVA, F.F.; GARCIA, D.A. et al. Estimativa de herdabilidade para prenhez de novilhas aos 14 meses em um rebanho Nelore por meio de modelos lineares generalizados mistos. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 44., 2007, Jaboticabal. **Anais...** Brasília: Sociedade Brasileira de Zootecnia, 2007. (CD-ROM).

R Development Core Team. **R:** A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2009. Available at: <http://www.R-project.org>. Accessed on: March 5, 2009.

SHIOTSUKI, L.; SILVA, J.A. II V.; ALBUQUERQUE, L.G. Associação genética da prenhez aos 16 meses com o peso à desmama e o ganho de peso em animais da raça Nelore. **Revista Brasileira de Zootecnia**, v.38, n.5, p.1211-1217, 2009.

SILVA, J.A.II V.; MELIS, M.H.V.; ELER, J.P. et al. Estimação de parâmetros genéticos para probabilidade de prenhez aos 14 meses e altura da garupa em bovinos da raça Nelore. **Revista Brasileira de Zootecnia**, v.32, n.5, p.1141-1146, 2003.

SILVA, J.A. II V.; DIAS L.T.; ALBUQUERQUE L.G. Estudo genético da precocidade sexual de novilhas em um rebanho Nelore. **Revista Brasileira de Zootecnia**, v.34, n.5, p.1568-1572, 2005.

SOUTHEY, B.R.; RODRIGUEZ-ZAS, S.L.; LEYMASTER, K.A. Discrete time survival analysis of lamb mortality in a terminal sire composite population. **Journal of Animal Science**, v.81, p.1399-1405, 2003.

SORENSEN, D.; GIANOLA, D. **Likelihood, Bayesian, and MCMC Methods in quantitative genetics**. New York: Springer-Verlag, 2002. p.607.

TEMPELMAN, R.J. Generalized linear mixed models in dairy cattle breeding. **Journal of Dairy Science**, v.81, p.1428-1444, 1998.

THOMPSON, R. Sire evaluation. **Biometrics**, v.35, p.339-353, 1979.

VAZQUEZ, A.I.; BATES, D.M.; ROSA, G.J.M. et al. Technical Note: An R package for fitting generalized linear mixed models in animal breeding. **Journal of Animal Science**, v.88, p.497-504, 2009.