Heritability Estimates For Visual Score Data Using Linear And Threshold Models – A Simulation Study

F. Barichello[1], M.M. de Alencar[2], R.A.A. Torres Júnior[3], and L.O.C. Silva[3]

Introduction

Genetic improvement programs in Brazil involve the monitoring of several traits, such as body structure, precocity, and muscling, which are evaluated using visual scores that have a discrete nature. The systems used to evaluate animals vary in terms of the number of classes, the referential from which the scores are given, and the dispersion of observations in score classes.

These traits are usually analyzed as if they were continuous using linear models, a methodology that is not optimal. More adequate methodologies, such as threshold models, have been described for these traits. Another approach to analyze discrete traits is using transformation procedures, which makes the use of a linear model more appealing.

The objective in this study was to use simulation in order to evaluate the impact of different forms of assigning visual scores to a continuous trait, genetic parameters, and methodologies of analysis on variance component estimates.

Material and methods

Two values for the heritability (h^2) and contemporary group variance (σ^2_{CG}), were evaluated in four different ways of assigning scores to obtain the observed trait (Y). A simulated herd with 40 sires and 1,200 dams was run for 20 years, keeping 1,000 dams after the breeding season, with 90% of them leaving a calf with record at weaning. This process resulted in around 18,000 phenotypes and 21,040 animals per data set. The results of five simulation programs for each of these variables were condensed into a single data set, which was evaluated using three different methodologies to estimate the variance components.

To simulate the categorical trait, an underlying continuous variable was first simulated and then classified based on thresholds to obtain the observed trait. The underlying variable was simulated from a linear model with an overall mean, a contemporary group effect, and an age of dam effect, considered fixed in the analysis, and additive direct and maternal, maternal permanent environmental, and an uncorrelated residual effect, considered random in the analysis.

Simulated effects of sex (2), year of calf's birth (20), and management group (5), were combined to build the contemporary groups, making 200 contemporary groups with around

---

[1] Universidade Estadual Paulista, FCAV/UNESP, Jaboticabal, SP, Brazil; FAPESP and CAPES' scholarship
[2] Embrapa Southeast Livestock, São Carlos, SP, Brazil; CNPq fellow
[3] Embrapa Beef Cattle, Campo Grande, MS, Brazil

90 calves in each. The contemporary group effect was simulated from a normal distribution with zero mean and variance equal to σ^2_{CG}, where σ^2_{CG} assumed two values of 0.25 or 1.00.

Direct and maternal additive effects were generated from a bivariate normal distribution with zero mean and variances $\sigma^2a = h^2$ (h^2 equals either 0.25 or 0.49), $\sigma^2m = h^2m$ ($h^2m = 0.09$), and the correlation between them $r_{am} = -0.20$. Maternal permanent environment effect was generated from a normal distribution with zero mean and variance $\sigma^2c = c^2$ ($c^2 = 0.09$). The residual variance was set to 1.00.

After generating the underlying continuous variable, it was categorized based on five thresholds determined according to the desired frequency distribution (Y). Four different sets of thresholds were applied, named: relative uniform distribution (Y_1 – the thresholds shifted from one contemporary group to another and were determined to produce on average the same frequency in each of the six classes for each contemporary group); relative normal distribution (Y_2 – similar to Y1 but with higher frequencies in the intermediate classes, 10%, 15%, 25%, 25%, 15%, and 10%); and relative asymmetric distribution (Y_3 – similar to Y2 but skewed toward the higher classes, 2%, 5%, 13%, 28%, 33%, and 19%); normal fixed distribution (Y_4 – similar to Y2 but the thresholds were kept fixed across contemporary groups causing the class frequencies to change according to contemporary group effect).

Visual scores were analysed with three methodologies (M): linear animal model (LN), threshold animal model (TH), and linear animal model applied to transformed data (TR – where the scores were replaced by the mean of a standard normal distribution sliced at the thresholds that would result in the class frequencies observed in the contemporary group). Variance components were estimated using the GIBBS2F90 and THRGIBBS1F90 programs (Misztal et al., 2002), based on the same model used to generate the underlying response variable. Non informative priors were used for the variance components. A chain of 1,100,000 cycles was run for all models, where the first 100,000 cycles were discarded as burn-in, and one sample was kept every 200 cycles. The POSTGIBBSF90 was used to verify chain convergence and to calculate posterior means.

The model comparisons for estimated heritability (h^2e) were made with an ANOVA procedure considering the actual h^2 value, σ^2_{CG}, distribution (Y), methodology (M), and their interactions as sources of variation.

Results and discussion

Table 1 presents the effect of direct heritability and methodology (h^2 x M) interaction on h^2e. For both levels of heritabilities studied, the threshold model presented higher estimated heritability (h^2e) than linear and linear with transformed data models, and the estimated heritabilities with linear and linear with transformed data were the same for both levels. Similar results were observed in a simulation study by Abdel-Azim & Berger (1999), who found out that the threshold model provided slightly greater heritability estimates than linear model. Visual scores for muscling, body structure and conformation of Nellore animals were analysed under threshold and linear models by Faria et al. (2008), who reported that threshold model resulted in higher heritability than linear model for visual score of muscling, while for body structure and conformation both approaches provided similar heritabilities.

For stillbirth and calving difficulty in Swedish Holstein, Steinbock et al. (2003) reported similar heritabilities derived from the threshold model and the heritabilities from the linear model when it was transformed to the underlying scale. No transformation was applied here.

Table 1: Effect of direct heritability (h^2) and methodology of analysis (M) interaction on estimated direct heritability (h^2e)

h^2	Methodology*		
	LN	TR	TH
0.25	0.222^b	0.222^b	0.248^a
0.49	0.436^b	0.437^b	0.486^a

* LN, TH, and TR: linear, threshold, and linear with transformed data models.

According to Matos et al. (1997), the threshold models, at least theoretically, seem to be more appropriated for discrete traits, and thus may capture a higher portion of genetic variability than linear models.

Many simulation studies about discrete traits (Meijering & Gianola, 1985; Hoeschele & Gianola, 1988; Abdel-Azim et al., 1999) have showed advantages of threshold over linear models. Although, studies carried out with field data (Steinbock et al., 2003; Faria et al., 2008) have not showed threshold model superiority over linear models. Some studies showed that when the discrete trait has four or more levels of response estimates from threshold and linear models are quite similar (Varona et al., 1999; Ramirez-Valverde et al., 2001; Lee et al., 2001), which is the case here with six classes.

In Table 2 the effect of contemporary group variance (σ^2_{CG}) and Y on the estimated direct heritability (h^2e) is presented. Both sources of variation had no effect on the estimated heritability when the true heritability was low ($h^2 = 0.25$), but at a high level of heritability ($h^2 = 0.49$) larger estimates of heritability were found with $\sigma^2_{CG} = 0.25$, independently of the way that the score were assigned (Y).

Table 2: Effect of direct heritability (h^2), contemporary group variance (σ^2_{CG}), and frequency data distribution (Y^*) on estimated direct heritability

σ^2_{CG}	Y_1	Y_2	Y_3	Y_4
	$h^2 = 0.25$			
0.25	0.231^{aA}	0.234^{aA}	0.230^{aA}	0.227^{aA}
1.00	0.231^{aA}	0.236^{aA}	0.231^{aA}	0.227^{aA}
	$h^2 = 0.49$			
0.25	0.460^{aA}	0.467^{aA}	0.461^{aA}	0.469^{aA}
1.00	0.439^{bB}	0.456^{aB}	0.443^{bB}	0.428^{cB}

*Y_1= uniform relative scale; Y_2 = normal relative scale; Y_3 = asymmetric relative scale; Y_4 = normal fixed scale.

Table 2 also shows that when the heritability is high and there is large contemporary group effect, the highest estimated heritability (h^2e) was obtained for scores assigned with a normal relative scale (Y_2), followed by an asymmetric scale (Y_3), and an uniform scale (Y_1) (no difference was found between asymmetric and uniform scales), while the fixed normal scale

(Y_4) provided the lowest estimated heritability. In a simulation study with discrete traits analyzed with a sire model under threshold methodology, Abdel-Azim & Berger (1999) found that correlations between the true and estimated heritability increase as the frequency of data distribution became closer to normal compared to skewed distributions. For this situation, when the trait is under high environmental influence, and the scores are assigned based on fixed thresholds, regardless of the contemporary group mean, many contemporary group presented highly skewed class distributions that may have caused the higher decrease on the estimated heritability, showing a larger difficulty to separate the additive genetic effect from the data.

Conclusion

Threshold models provide higher estimated heritability, than linear models. A scale of assigning visual scores relative to the contemporary group with higher frequencies for the intermediate classes seem to retain more information to perform a genetic evaluation based on the higher heritabilities provided. Additionally, the use of a fixed scale seems to result in a less accurate result, mostly when heritabilities are high and large contemporary group variations occur.

References

Abdel-Azim, G.A., and Berger, P.J. (1999). *J. Anim. Sci.*, 77:582-590.

Faria, C.U., Magnabosco, C.U., Albuquerque, L.G. *et al.* (2008). *Pesq. Agropec. Bras.,* 43:835-841.

Hoeschele, I., and Gianola, D. (1988). *J. Dairy Sci.,* 72:1569-1577.

Lee, D., Misztal, I., Bertrand, J.K. *et al*. (2002). *J. Appl. Genet.,* 43(2):209-216.

Misztal, I., Tsuruta, S., Strabel, T. *et al.* (2002). *Proc. 7th WCGALP*. CD-ROM Communication No. 28-07.

Meijering, A., and Gianola, D. (1985). *Gen. Sel. Evol.,* 17:115-131.

Ramirez-Valverde, R., Misztal, I., and Bertrand, J.K. (2001). *J. Anim. Sci.*, 79:333-338.

Steinbock, L., Näsholm, A., Berglund, B. *et al*. (2003). *J. Dairy Sci.,* 86:2228-2235.

Varona, L., Misztal, I., and Bertrand, J.K. (1999). *J. Anim. Sci.,* 77:2003-2007.