COMPARISON OF COMPUTATIONAL PERFORMANCE AND BREEDING VALUE ACCURACY FOR BINOMIAL AND TRINOMIAL THRESHOLD MODELS

C. J. Kaiser¹ and B. L. Golden²

¹Animal Genetics and Breeding Unit³, University of New England, Armidale, NSW, 2351
²Department of Animal Sciences, Colorado State University, Fort Collins, USA 82093

SUMMARY

We analyzed 100 simulated data sets. Each set consisted of 50,000 animals generated from a foundation of 500 dams and 20 sires. Our objective was to determine if subdividing undesirable categories with moderate incidence improved breeding value accuracy while remaining computationally feasible. Scores were simulated by truncating an underlying standard normal liability distribution at 0 and 1.5, creating three categories. We used a heritability of .33 and sex effects of .25 for females and .5 for males so that approximately 40.2, 37.1, and 22.7% of animals were in categories one, two and three, respectively, with category one representing unaffected animals. For the binomial analysis, we combined categories two and three. Maximum a posteriori analysis of trinomial vs binomial scores increased breeding value accuracy by 2.7% for all animals (72.4 vs 69.6%). Breeding value accuracies within the first and second quintiles were increased .23 and .53%. The nature of accuracy differences across the breeding value distribution requires further study, but the increase associated with the trinomial model suggests increased selection response should be possible if animals considered for selection are in or near the undesirable category. While the binomial model used less CPU time (360 vs 458 s), computing resources were more than adequate to analyze this size data set with the trinomial model.

Keywords: Categorical trait, threshold model, generalized linear models, mixed models.

INTRODUCTION

Generalized linear mixed model equations for threshold categorical traits such as calving difficulty or dystocia were originally developed for the case of multiple observation categories (Gianola and Foulley 1983). In addition, heritability estimates from linear analyses of categorical scores were higher when more categories were included in the model (Burfenning et al. 1978a; b; 1981; Naazie et al. 1991; Kriese et al. 1994). Despite this, most subsequent developments in categorical models considered only the binary case, because the equations are simpler and the threshold equation does not have to be built (Foulley and Gianola 1986; Höschle et al. 1986; Janss and Foulley 1993).

Meuwissen et al. (1995) found no increase in selection response from subdividing the undesirable category. However, they simulated undesirable category incidences of only 10 and 25%, while respective selection rates for males and females were 10 and 50%. Thus, selected animals were less closely related to affected animals, especially those at the lower end of the distribution. Dystocia incidence is often much higher, with reported rates of 40% or more (Laster et al. 1973; Laster

³ AGBU is a Joint Institute of NSW Agriculture and The University of New England.
Moreover, direct and maternal components of dystocia liability are negatively correlated (Burfening et al. 1981). Under these conditions, affected animals or their relatives may be selected, especially when several traits are selected on, many of which may also be negatively correlated with dystocia.

The objective of this study was to determine if, under higher incidence levels, using a trinomial model improves accuracy over a binomial model without greatly increasing computational requirements.

MATERIALS AND METHODS
We simulated 100 data sets, each with 50,000 animals. Each foundation population consisted of 500 dams and 20 sires. In each mating season, one calf per dam was generated. Following this, 10% of sires and 20% of dams were replaced by calves born that year, to maintain a constant herd size. Mating, selection, and replacement were all done randomly. Thus, just short of 99 mating seasons were required for each data set.

We generated a single categorical score per animal by truncating an underlying normal liability. Each animal's liability was the sum of a sex effect of 5 for males and .25 for females; an additive genetic component, distributed N[0, .25A], where A is Wright's numerator relationship matrix; and error, distributed N[0, I]. Upper thresholds for the first and second categories were set at 0 and 1.5, resulting in expected frequencies of 40.2, 37.1, and 22.7 for categories one, two, and three, respectively. For the binomial analysis, categories two and three were combined. For purposes of discussion, the lowest score was considered to be the desirable category, similar to BIF calving ease scores (BIF 1996). Thus animals with the lowest breeding values would be candidates for selection.

For each data set, we obtained maximum a posteriori (MAP) breeding value estimates from the vector of categorical scores, Y, by fitting a generalized linear mixed model with a probit link (Gianola and Foulley 1983) to the model

\[ y = X\beta + Zu + e \]

where y is a vector of underlying normally distributed liabilities; X and Z are incidence matrices relating fixed and random effects to liabilities; \( \beta \) is a vector of fixed effects including sex effects and one threshold for the binomial model or two thresholds for the trinomial one; u is a random vector of additive genetic components; and e is a random error vector.

Both models were analyzed with dscat (Kaiser 1996), a program designed for MAP Method R variance component analysis (Snelling et al. 1995) for one categorical trait and zero or more continuous traits. This program can optionally produce a MAP estimate of the solution vector given a set of variance components.

For each data set, we calculated correlations between predicted and true breeding values (accuracies) for all animals, under both models. To determine how well each model differentiated among potential replacement animals under extreme and more moderate selection intensities, we...
also calculated the correlation between predicted and true breeding values for animals in either the first or second quintile.

As a measure of relative computational performance, we also obtained required CPU times for each analysis on a DEC alpha Station 500, running at 333 MHz, with 128 megabytes of RAM. CPU times were required times for the program to do all necessary data preparation, perform the required number of Fisher scoring rounds to reach convergence after starting with an initial beta vector equal to zero, and write the results to a file. An analysis was considered to have converged when the average squared change to elements of the solution vector between Fisher scoring rounds was less than 10⁻⁶.

Because identical data sets were analyzed under both models, results for each set were considered pairs of observations. We therefore used a paired t-test to test whether the differences in accuracy and CPU time, obtained from the trinomial model and binomial models, were greater than zero under the model

\[ y_{ij} = \mu_j + s_{ij} + e_{ij} \]

where \( y_{ij} \) is the difference in measure \( j \) for data set \( i \), \( \mu_j \) is the mean difference for measure \( j \), \( s_{ij} \) is the effect of data set \( i \) on difference \( j \), and \( e_{ij} \) is random error.

RESULTS AND DISCUSSION

Compared to the binomial analysis, the trinomial model yielded higher breeding value accuracies for all groups of animals (\( P < .001 \); Table 1). The accuracy improvement for the first quintile was minor, suggesting the extra category provided very little information on the top animal’s genetic merit. Under single trait selection schemes, these would be the animals considered as replacements, and it could therefore be expected that there would be little or no difference in selection response, which is in agreement with the findings of Meuwissen et al. (1995), even though the incidence rates were higher than those of the previous study. The improvement in accuracy from the trinomial model would be greater with fewer years of data, as would occur in most breed society data sets. Each additional year’s data increases the information on parents’ breeding values. Therefore, the additional information gained by subdividing categories becomes less important.

Table 1. Mean breeding value accuracies from binomial and trinomial analyses

<table>
<thead>
<tr>
<th></th>
<th>Trinomial</th>
<th>Binomial</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (%)</td>
<td>SE</td>
<td>Mean (%)</td>
</tr>
<tr>
<td>All animals</td>
<td>72.4</td>
<td>.33</td>
<td>69.6</td>
</tr>
<tr>
<td>First quintile</td>
<td>40.3</td>
<td>.49</td>
<td>40.1</td>
</tr>
<tr>
<td>Second quintile</td>
<td>17.8</td>
<td>.27</td>
<td>17.2</td>
</tr>
</tbody>
</table>
However, in a multitrait selection scheme, the selection threshold would likely be in the second quintile, unless all traits are positively correlated, or selection intensity is very high. In this case, it would be more important to differentiate between those animals. The .53% accuracy gain among these animals may therefore be important if it represents reranking of animals. Further research will seek to determine the magnitude of accuracy differences under a range of incidence rates, as well as when both direct and maternal components are considered.

Mean computing time for the binomial analyses was 356.9 s compared to 455.7 s for the trinomial analysis. While the simpler model took 21.7% less time, both analyses converged in less than eight minutes. This suggests that computer resources should not be limiting in most cases. In addition, preliminary comparisons on an IBM reduced instruction computer architecture yielded virtually identical CPU times for the two models (Kaiser 1996). This may be due either to compiler differences or inherent performances of the different architectures. Again, further research will investigate this question.

These results indicate that when incidence of an undesirable character is moderate to high, subdividing categorical data into additional categories increases breeding value prediction accuracy, not only across all animals, but among animals which would be considered for selection. This suggests producers and breed associations are improving genetic evaluation by using multiple categories to describe the nature of scored traits such as dystocia. Further research should seek to describe differences between the two models under varying incidence rates and more complex scenarios, such as a direct maternal model.

REFERENCES