Properties of Threshold Model Predictions\textsuperscript{1,2}

Gamal A. Abdel-Azim and P. J. Berger\textsuperscript{3}

Department of Animal Science, Iowa State University, Ames 50011

\textbf{ABSTRACT:} Estimation of genetic parameters and accuracy of threshold model genetic predictions were investigated. Data were simulated for different population structures by using Monte Carlo techniques. Variance components were estimated by using threshold models and linear sire models applied to untransformed data, logarithmically transformed data, and transformation to Snell scores. Effects of number of categories (2, 5, and 10), incidence of categories (extreme, moderate, and normal), heritability in the underlying scale (.04, .20, and .50), and data structure (unbalanced and balanced) on accuracy of genetic prediction were investigated. The real importance of using a threshold model was to estimate genetic parameters. An expected heritability of .20 was estimated to be .22 and .10 by a threshold model and a linear model, respectively. Accuracy increased significantly with a larger number of categories, a more normal distribution of incidences, increased heritability, and more balanced data. Even threshold models were shown to be more efficient with more than two categories (e.g., binomial). Transformation of scale did not accomplish the purpose intended.

Key Words: Accuracy, Prediction, Threshold Models, Genetics

\textcopyright 1999 American Society of Animal Science. All rights reserved.


\section*{Introduction}

Two general classes of phenotypes are measured in animal breeding data, continuous and discrete. Many traits of importance, such as calving ease, disease resistance, and livability are measured on a discrete scale that is categorical. Genetic evaluation for categorical traits is achieved through different methodology than evaluation of continuous traits. Best linear unbiased prediction (BLUP) is the best method for prediction of random effects if the response variable is continuous (Henderson, 1973). Categorical variables violate many assumptions for mixed linear models; therefore, BLUP is not appropriate (Thompson, 1979; Gianola, 1982).

Based on the threshold concept, nonlinear methods have been described for sire evaluation for categorical traits (Gianola, 1982; Gianola and Foulley, 1983; Harville and Mee, 1984; Gilmour et al., 1985).

\section*{Materials and Methods}

\subsection*{Population Structures}

Four properties of the population were varied. While one property was changed, the remaining
properties were kept constant. The four characteristics varied were heritability (.05, .20, and .50), number of categories (2, 5, and 10), frequency of observed categories (nearly normally distributed, moderately skewed, and extremely skewed), and balancedness (balanced and unbalanced). Table 1 describes the populations simulated. There were 40 sires in each data set. Each sire had 180 progeny in the balanced data, making a total of 7,200 records per data set. The unbalanced data structure was made up of 13 sires with 360 progeny each, 13 sires with 180 progeny each, and 14 sires with 12 progeny each, making 7,188 records. For each scenario evaluated, 20 replicate populations were simulated and analyzed. Average variance ratios, mean squared error (MSE), and average bias of heritability were calculated for each population structure.

The MSE and bias were reported as ratios of true heritability used in simulation. The MSE was computed as

\[ \left( \frac{\sum_{i=1}^{20} (\hat{h}_i^2 - h^2)^2}{20} \right) h^2 \]

where \( \hat{h}_i^2 \) is the heritability estimate of the \( i \)th replicate, \( i = 1, \ldots, 20 \); and \( h^2 \) is the true heritability. Bias was computed as

\[ \left( \frac{\sum_{i=1}^{20} (\hat{h}_i^2 - h^2)/20}{h^2} \right) h^2 \]

### Simulation of Data

To simulate categorical data, an underlying continuous response variable was first simulated and then categorized at fixed boundary points to obtain a categorical response variable. The underlying response variable was simulated for a LM with two fixed factors and one random factor in addition to the random error. The model was

\[ y_{ijkl} = \mu + SX_i + PTY_j + S_k + e_{ijkl} \]  \[ \text{[1]} \]

where \( y_{ijkl} \) is a response in a continuous response variable; \( \mu \) is the overall mean; \( SX_i \) is a fixed factor representing sex with \( i = 1, 2 \); \( PTY_j \) is another fixed factor representing parity with \( j = 1, 2, 3 \); \( S_k \) is a random sire factor, \( k = 1, 2, \ldots, 40 \); and \( e_{ijkl} \) is a random residual. Sire effects were generated from a normal distribution with mean zero and variance \( \sigma^2_s = .25h^2; S_k \sim \text{IID } N(0, \sigma^2_s) \), \( k = 1, \ldots, 40 \), where IID means independently and identically distributed. Elements of the vector of random residual effects, \( e \), were each generated from a normal distribution with mean zero and variance \( \sigma^2_e = 1 - .25h^2; e_{ijkl} \sim \text{IID } N(0, \sigma^2_e) \).

Both sides of Eq. [1] were divided by the standard deviation of the error to make an error variance equal to 1.0. The underlying variable was categorized, after being computed, at fixed thresholds or boundary points determined according to some frequency distribution.

### Prediction Procedure

Threshold model equations, as developed by Harville and Mee (1984), were approached in a procedure similar to Henderson's mixed linear model equations (Henderson, 1973). Assuming known variance components, mixed linear model equations can be developed by maximizing the joint probability density function (pdf) of \( y \) and \( \beta \) of the following model:

\[ y = X\alpha + Z\beta + \epsilon \]  \[ \text{[2]} \]

where \( X \) and \( Z \) are \( N \times p \) and \( N \times q \) incidence matrices for fixed and random effects, respectively; \( \alpha \) and \( \beta \) are \( p \times 1 \) and \( q \times 1 \) vectors of unknown fixed and random

<table>
<thead>
<tr>
<th>Case</th>
<th>( h^2 )</th>
<th>Sires</th>
<th>Progeny/ sire</th>
<th>Categories</th>
<th>Structure</th>
<th>Incidence, %a</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.04</td>
<td>40</td>
<td>180</td>
<td>5</td>
<td>balanced</td>
<td>75, 12, 7, 4, 2</td>
</tr>
<tr>
<td>2</td>
<td>.20</td>
<td>40</td>
<td>180</td>
<td>5</td>
<td>balanced</td>
<td>75, 12, 7, 4, 2</td>
</tr>
<tr>
<td>3</td>
<td>.50</td>
<td>40</td>
<td>180</td>
<td>5</td>
<td>balanced</td>
<td>75, 12, 7, 4, 2</td>
</tr>
<tr>
<td>4</td>
<td>.20</td>
<td>40</td>
<td>180</td>
<td>2</td>
<td>balanced</td>
<td>95, 5</td>
</tr>
<tr>
<td>5</td>
<td>.20</td>
<td>40</td>
<td>180</td>
<td>5</td>
<td>balanced</td>
<td>75, 12, 7, 4, 2</td>
</tr>
<tr>
<td>6</td>
<td>.20</td>
<td>40</td>
<td>180</td>
<td>10</td>
<td>balanced</td>
<td>30, 20, 14, 8, 10, 7, 5, 3, 2, 1</td>
</tr>
<tr>
<td>7</td>
<td>.20</td>
<td>40</td>
<td>180</td>
<td>5</td>
<td>balanced</td>
<td>90, 7, 15, 1, 5</td>
</tr>
<tr>
<td>8</td>
<td>.20</td>
<td>40</td>
<td>180</td>
<td>5</td>
<td>balanced</td>
<td>75, 12, 7, 4, 2</td>
</tr>
<tr>
<td>9</td>
<td>.20</td>
<td>40</td>
<td>180</td>
<td>5</td>
<td>balanced</td>
<td>17, 20, 25, 35, 3</td>
</tr>
<tr>
<td>10</td>
<td>.20</td>
<td>40</td>
<td>180</td>
<td>5</td>
<td>balanced</td>
<td>75, 12, 7, 4, 2</td>
</tr>
<tr>
<td>11</td>
<td>.20</td>
<td>40</td>
<td>360, 180, 12</td>
<td>5</td>
<td>unbalanced</td>
<td>75, 12, 7, 4, 2</td>
</tr>
</tbody>
</table>

*aFrequency distribution of two, five, or 10 categories.*
parameters, respectively; \( \epsilon \) is an \( N \times 1 \) vector of continuous residuals; \( y \) is an \( N \times 1 \) vector of continuous response variables, and \( N \) is the total number of observations.

Maximizing \( f(y, \beta) = g(y|\beta) \times h(\beta) \) produces the mixed linear model equations (Harville and Mee, 1984). A similar approach was followed to derive the TM equations after expressing the outward categorical scale, say, \( w \), in terms of the underlying continuous scale, \( y \). The relationship between \( Y_i \) and \( W_i \), where \( Y_i \) belongs to the \( k \)th interval, we observe that \( W_i = k \).

Note that the first part of the previous joint pdf produces the probability that \( W_i = k \) is

\[
p(W_i) = \int_{\xi_{k-1}}^{\xi_k} g(Y_i) \, dY_i
\]

and the pdf of \( w \) is

\[
p(w) = \prod_{i=1}^{N} \int_{\xi_{k-1}}^{\xi_k} g(Y_i) \, dY_i
\]

where \( k = 1, \ldots, M \) and the threshold model equations are obtained by maximizing

\[
f(w, \beta) = \prod_{i=1}^{N} \int_{\xi_{k-1}}^{\xi_k} g(Y_i - x_i \cdot \alpha - z_i \cdot \beta) \, dY_i
\]

\(
\times h(\beta)
\)

Note that the first part of the previous joint pdf represents \( p(w|\beta) \).

The threshold model equations are written as follows:

\[
\begin{bmatrix}
Q^k \\
L^k X \\
L^k Z \\
XL^k X R^k X R^k Z \\
ZL^k Z R^k X R^k Z + D^{-1} \\
t(\xi^k, \sigma^k, \beta^k; w) \\
X' \epsilon(\xi^k, \sigma^k, \beta^k; w) \\
Z' \epsilon(\xi^k, \sigma^k, \beta^k; w) - D^{-1} \beta
\end{bmatrix}
\]

Explicit expressions for the elements of \( Q, L, R, t, \) and \( \epsilon \) are given in Harville and Mee (1984). The \( \xi^k, \sigma^k, \beta^k \) are vectors of estimates obtained after building the equations \( k \) times for \( (M-2) \) boundary points (\( M \) is the number of response categories); \( p \) fixed parameters; and \( q \) random parameters, respectively. Equations [3] were solved for the adjustments \( \xi^k + 1 - \xi^k, \sigma^k + 1 - \sigma^k, \) and \( \beta^k + 1 - \beta^k, \) and then solutions were added to the previous iterate. This process was repeated until the corrections became zero. Initial guesses for \( \xi^0, \sigma^0, \) and \( \beta^0 \) were computed to build the system of equations for the first time. New values of the variance components were computed between sets of converged solutions until the variance components converged (Harville and Mee, 1984). Solutions and the variance ratio, \( e^2 / s^2 \), had converged when the largest difference between successive iterates was <10^{-7}.

Solutions were obtained by using the algorithm described by Harville and Mee (1984). Threshold Model Analysis program, TMA, was used in computing the threshold model predictions. Further details can be found in Abdel-Azim (1996).

By definition, accuracy of prediction is the correlation between true and predicted breeding values (Falconer, 1989; Mrode, 1996). Product moment correlations were, therefore, computed between true and predicted sire transmitting abilities (TA) for each replicate. Other reasons for choosing correlation were its indication of correctness of sire rankings and also its contribution to the magnitude of expected response to selection.

Computing correlation coefficients was followed by testing the differences within each factor (e.g., correlations of true and predicted transmitting abilities were compared for 2 vs 5, 2 vs 10, and 5 vs 10 categories). Similar comparisons were made for the other factors in the simulation.

Sampling distributions of correlations greater than zero are skewed. Therefore, it was necessary to transform correlations to make a test of hypothesis for pairwise differences among correlations. We used the transformation procedure devised by Fisher (1921), who suggested the following formula to transform \( r \) to a normal deviate, \( z \):

\[
z = 0.5 \ln \frac{1 + r}{1 - r}
\]

and the formula

\[
r = \frac{e^{2z} - 1}{e^{2z} + 1}
\]

to transform \( z \) back to \( r \).

A usual t-test was then used to test the differences between every two means within each factor; transformed \( r \) (\( z \)) were used in the test. A pooled value of the correlation for 20 correlations in each case was found after subtracting a bias from each \( z_i, i = 1, \ldots, 20 \) that accumulates if many correlations are averaged, then computing averages of corrected \( z_i \), and
transforming the averages back to \( r \) by using Eq. [5]. The bias as reported by Fisher is \( \rho / (2(n - 1)) \), where \( n = 40, \) number of sires, and \( r \) of the uncorrected average of \( z_i \) was used in place of \( \rho \) (Snedecor and Cochran, 1967; Steel and Torrie, 1960).

The 95% confidence limits were found for each pooled correlation, computed for the average of the corrected \( z_i \) as

\[
\overline{z_{corr}} \pm 1.96 \left( \sqrt{\frac{1}{n} \sum_{i=1}^{20} (n_i - 3)} \right)
\]

where the term multiplied by 1.96 is the approximate standard error of \( \overline{z_{corr}} \). Note that the normalized correlation, \( z \), has \( 1/(n - 3) \) as its sampling variance. In addition, the confidence limits of \( \overline{z_{corr}} \) were transformed back to the observed scale of \( r \) by using Eq. [5] to represent confidence limits for the pooled correlation coefficient of each scenario.

**Transforming and Scaling the Outward Scale**

The categorical response variable was first transformed to logarithms. Transformation to logarithms was an attempt to normalize the categorical scale before estimating variance components based on LM. Coefficient of variation of the simulated categorical data was always more than 20%, indicating that transformation was expected to have an appreciable effect (Falconer, 1989).

Snell (1964) suggested a scoring procedure to be used with subjective measurements to make residual deviations normally distributed and residual variance homogeneous. An underlying continuous variable that follows a logistic distribution that closely approximates the normal distribution was assumed by the scoring procedure. The assumption of a logistic distribution instead of a normal distribution allows considerable simplification. Snell’s method described both exact and approximate solutions, but the approximate solution was easier to obtain and was adequate for most practical purposes. In this study, the approximate solution for the boundary points was obtained and then midpoints were taken as scores. Computed scores were not equally spaced, and they were used instead of the original subjective scores to obtain REML estimates of the variance components.

**Variance Component Estimation**

Variance components were estimated using linear sire models applied to untransformed data (LM-RAW), logarithmically transformed data (LM-LOG), and transformation to Snell (1964) scores (LM-SNELL), based on the following LM:

\[ w = X\beta + Zu + \epsilon \]  \[ \text{[6]} \]

where \( w \) is an \( N \times 1 \) vector of categorical data; \( X \) is an \( N \times p \) known model matrix for the fixed effects (\( p = 6 \)); \( Z \) is an \( N \times q \) known model matrix for the random effect (\( q = \text{total number of sires} \)); \( \beta \) and \( u \) are \( p \times 1 \) and \( q \times 1 \) vectors of unknown fixed and random parameters, respectively; and \( \epsilon \) is an \( N \times 1 \) vector of random residuals. Assumptions were

\[
\begin{align*}
E \left[ \frac{w}{\epsilon} \right] &= X\beta \\
\text{var} \left[ \frac{u}{\epsilon} \right] &= \begin{bmatrix} \sigma^2_1 & 0 \\ 0 & \sigma^2_a \end{bmatrix}
\end{align*}
\]

The REML estimates of the variance components were obtained by using SAS PROC MIXED (Littell et al., 1996), assuming that the raw and transformed data were the observable response variables.

Variance components were also estimated based on the TM described by Eq. [6]. Assumptions for the standardized TM with an error variance 1 and first boundary point 0 were written as follows:

\[
\begin{align*}
E \left[ \frac{w}{\epsilon} \right] &= X \left[ \begin{array}{c} \alpha_1 - \hat{z}_1 \\
\alpha_2
\end{array} \right] g_{e - 1} \\
\text{var} \left[ \frac{\beta}{\epsilon} \right] &= \gamma^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}
\end{align*}
\]

where \( \alpha = [\alpha_1, \alpha_2] \) and \( \hat{z}_1 \) is the first boundary point, which is subtracted to be equal to 0. Solutions for effects in [3] and estimates of variance components followed the algorithms given in Harville and Mee (1984). Further details of specific computational strategies were given in Abdel-Azim (1996). The computational strategy made use of specific “tools” available in Animal Breeder’s Tool Kit (Golden et al., 1997).

Variance components were estimated by iteration in a way similar to LM. The TM equations were nonlinear and needed to be solved by iteration. After solving the equations, estimates of the variance ratio, \( \hat{\gamma} \), were obtained by the iterative algorithm, Eq. [7]. The whole process of solving the equations and estimating the variance was repeated several times until the estimates converged (or until the difference between two successive ratios was less than \( 10^{-7} \)):

\[
\hat{\gamma}^{k+1} = q \left[ \hat{\beta} \left( \hat{\gamma}^k \right) \right]' \times \hat{\beta} (\hat{\gamma}^k) + \text{tr} C_{q \times q}
\]

where \( \hat{\gamma}^k \) is the ratio obtained from the \( k \)th iterate; \( C_{q \times q} \) is the \( q \times q \) lower right corner of a generalized inverse of the coefficient matrix in Eq. [3]; and \( \hat{\beta} (\hat{\gamma}^k) \) is a \( q \times 1 \) vector of random solutions expressed as a function of the variance ratio. Note that \( \beta \) was functionally dependent on the variance ratio used in TM equations.

Additive genetic variance (AGV) on the outward scale was expected to increase by increasing the number of categories. The AGV, as derived for two categories by Dempster and Lerner (1950) and for
any number of categories by Gianola (1982), is written as

\[
AGV = h^2 \sum_{j=1}^{m-1} z_j (\eta_{j+1} - \eta_j)^2
\]

where \( h^2 \) is heritability on the underlying scale; \( z_j \) is the ordinate of a standard normal density corresponding to the boundary point between categories \( j \) and \( j + 1 \); and \( \eta_j \) (i.e., score assigned to a response category) is a weight given to category \( j \). By using Eq. [9] and assuming an incremental increase of one unit for weights assigned to sequential categories, amounts of the AGV for 2, 5, and 10 categories were explicitly computed as

\[
AGV(2) = 0.0021
\]
\[
AGV(5) = 0.097
\]
\[
AGV(10) = 0.9003
\]

The ratio of AGV with five categories to two categories was 46 (= 0.0970/0.0021) and from 5 to 10 categories the ratio was 9.3. Of course, the magnitude of these ratios depended on the distribution of observations in categories given in Table 1.

From the fundamentals of quantitative genetics, accuracy is expected to increase as AGV increases (Falconer, 1989). This can be shown for mass selection, where accuracy equals \( s_A / s_P \), which increases with greater values of \( s_A \). In general, the correlation between predicted and true TA decreases with greater amounts of nonadditive genetic variance associated with the outward scale. This happens because the effects predicted in \( \beta \) were only due to the additive genetic part; hence, \( \text{cor}(\beta, \beta) \) increases with larger amounts of AGV in \( \beta \).

### Results

#### Number of Categories

Table 2 gives pooled correlations between predicted and true TA along with 95% confidence limits. The association between true and predicted sire TA increased as the number of categories increased. Although the difference between two and five categories was greater than the difference between five and ten categories, paired comparisons of two vs five, two vs 10, and five vs 10 were significantly different (Pr > |T| was always < .001). Figure 1 gives the 20 correlations computed for cases 4, 5, and 6, representing two, four, and 10 categories. More variability was associated with two categories than five or 10, also differences on average between two and five were greater than the differences between five and 10.

#### Incidence of Categories

Correlations for extreme, moderate, and normal incidence are summarized in Figure 2. Accuracy tended to decrease as categories became more extreme. Pooled correlations and their corresponding 95% confidence limits for the three levels of incidence are given in Table 2. The pooled correlation increased as the incidence became more normal. The contrasts of extreme vs moderate, extreme vs normal, and moderate vs extreme were all different (P < .001).

The decrease in accuracy observed with more extreme incidence was associated with a decrease in AGV computed by using Eq. [9] for normal, moderate, and extreme incidence of categories as .2330, .0971, and .0175, respectively. The decrease in AGV explains the decrease in accuracy.

#### Heritability in the Underlying Scale and Data Structure

Accuracy increased as heritability in the underlying scale increased (Table 2). This can also be seen from Eq. [9]. Greater values of \( h^2 \) result in greater amounts

### Table 2. Accuracy of genetic evaluation expressed as correlations between true and predicted transmitting ability, and corresponding 95% confidence limits for three levels of heritability, three levels of categorization, three frequency distributions of five categories, and two data structures

<table>
<thead>
<tr>
<th>Case</th>
<th>Lower limit (pooled)</th>
<th>Upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.679</td>
<td>.716</td>
</tr>
<tr>
<td>2</td>
<td>.898</td>
<td>.911</td>
</tr>
<tr>
<td>3</td>
<td>.956</td>
<td>.962</td>
</tr>
<tr>
<td>4</td>
<td>.755</td>
<td>.784</td>
</tr>
<tr>
<td>5</td>
<td>.898</td>
<td>.911</td>
</tr>
<tr>
<td>6</td>
<td>.944</td>
<td>.951</td>
</tr>
<tr>
<td>7</td>
<td>.846</td>
<td>.865</td>
</tr>
<tr>
<td>8</td>
<td>.898</td>
<td>.911</td>
</tr>
<tr>
<td>9</td>
<td>.935</td>
<td>.944</td>
</tr>
<tr>
<td>10</td>
<td>.749</td>
<td>.779</td>
</tr>
<tr>
<td>11</td>
<td>.898</td>
<td>.911</td>
</tr>
</tbody>
</table>

*Factors influencing the correlations in each case are defined in Table 1.*
of AGV and in more accuracy. The contrasts of .04 vs .20, .04 vs .50, and .20 vs .50 were all significant (P < .001).

Accuracy was also affected by the structure of the data (Table 2). Accuracy decreased about 15% in unbalanced data compared with balanced data. Differences between accuracy of genetic predictions using balanced and unbalanced data were significant (P < .001).

Estimates of Genetic Parameters

Estimates of heritability and variance ratio by TM, LM-RAW, LM-LOG, and LM-SNELL are summarized in Table 3. Threshold model estimates of heritability and variance ratio were relatively similar to their expected values for each of the factors of heritability, number of categories, frequency of categories, and data structure. However, LM estimates of the same parameters on the outward scale were different from their expected values. Heritability was underestimated by approximately one-half for categories of two and five. Variance ratio was overestimated by LM. Comparison between TM and LM estimates shows the real importance of using a TM to estimate genetic parameters for categorical data. The decrease in accuracy when using the outward scale of measurement in LM cannot be sacrificed. Estimates obtained by LM-RAW, LM-LOG, and LM-SNELL were similar to estimates obtained by TM when the number of categories was at 10. Transformations of LM-LOG and LM-SNELL made little improvement over LM-RAW.

Table 4 gives MSE and biases of heritability estimates for the 11 scenarios. In most of the scenarios, heritability estimates based on TM had smaller MSE and less bias than estimates based on LM. Larger MSE for two vs five categories is visible in Figure 1. The decrease in MSE with change in the frequency distribution is visible in Figure 2.

In general, the superiority of estimates of variance components based on TM was remarkable over all the different scenarios studied. Estimates based on LM were considerably improved in two scenarios (i.e., 10 response categories, case 6, and normal frequency distribution with five response categories, case 9) but did not outperform the TM estimates. Improved accuracy with 10 categories was more obvious than with five or two categories because 10 categories makes the categorical response variable closer to the continuous variable on the underlying scale.

In addition, we found that transforming data or creating a new set of scores had minimal or no effect on improving LM estimates of variance components. Therefore, transformation is not recommended, especially if it requires a considerable amount of computations as in the scenario of Snell’s method.

Discussion

We have stressed the ability of the threshold model to fit variables on an underlying scale of continuous variation. This required the use of simulation to establish the properties of threshold models under different experimental conditions. Simulation was necessary to enable comparison between true and
predicted transmitting ability for the same model (e.g., TM). Henderson (1975) stressed the need to use simulation to compare alternative sire evaluation methods.

Hoeschele (1988) compared predicted TA by using TM and LM. No attention was given to properties of estimated variance components. Our choice of parameters was similar to the parameters used by Hoeschele (1988); however, we included a wider range of parameters defined by heritability, number of categories, frequency distribution, and population structure. The use of five and 10 categories was new. This may well be an important message conveyed by this research. That is, even threshold models can be more efficient when the number of categories was increased beyond two (i.e., binomial case). We com-

Table 3. Comparison of heritability and variance ratio estimates computed by four methods: threshold model (TM); linear model for raw data (LM-RAW); linear model for log-transformed data (LM-LOG); and linear model for Snell scores (LM-SNELL).

<table>
<thead>
<tr>
<th>Case</th>
<th>Heritability</th>
<th>Estimated Variance Components</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>TM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( h^2 )</td>
</tr>
<tr>
<td>1</td>
<td>.0500</td>
<td>96.524</td>
</tr>
<tr>
<td>2</td>
<td>.2150</td>
<td>18.139</td>
</tr>
<tr>
<td>5</td>
<td>.2150</td>
<td>18.139</td>
</tr>
<tr>
<td>6</td>
<td>.2129</td>
<td>18.496</td>
</tr>
<tr>
<td>7</td>
<td>.1787</td>
<td>22.759</td>
</tr>
<tr>
<td>8</td>
<td>.2150</td>
<td>18.139</td>
</tr>
<tr>
<td>10</td>
<td>.2150</td>
<td>18.139</td>
</tr>
<tr>
<td>11</td>
<td>.2150</td>
<td>19.041</td>
</tr>
</tbody>
</table>

*aEach value is the average of 20 replications.
*bFactors influencing the estimates for each case are defined in Table 1.
*cTrue heritability and variance ratio used in the simulation are defined in Table 1.
Table 4. Comparison of mean square error and bias of heritability estimates among four methods: threshold model (TM); linear model for raw data (LM-RAW); linear model for log-transformed data (LM-LOG); and linear model for Snell scores (LM-SNELL)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.165</td>
<td>0.2507</td>
<td>0.111</td>
<td>0.4544</td>
<td>0.102</td>
<td>0.4055</td>
<td>0.0999</td>
<td>0.3984</td>
</tr>
<tr>
<td>2</td>
<td>0.082</td>
<td>0.0751</td>
<td>0.0486</td>
<td>0.4836</td>
<td>0.0430</td>
<td>0.4533</td>
<td>0.0419</td>
<td>0.4472</td>
</tr>
<tr>
<td>3</td>
<td>0.0426</td>
<td>0.0177</td>
<td>0.1258</td>
<td>0.4763</td>
<td>0.1143</td>
<td>0.4479</td>
<td>0.1137</td>
<td>0.4464</td>
</tr>
<tr>
<td>4</td>
<td>0.0130</td>
<td>0.0214</td>
<td>0.1300</td>
<td>0.8041</td>
<td>0.1300</td>
<td>0.8040</td>
<td>0.1300</td>
<td>0.8040</td>
</tr>
<tr>
<td>5</td>
<td>0.0082</td>
<td>0.0751</td>
<td>0.0486</td>
<td>0.4836</td>
<td>0.0430</td>
<td>0.4533</td>
<td>0.0419</td>
<td>0.4472</td>
</tr>
<tr>
<td>6</td>
<td>0.0085</td>
<td>0.0643</td>
<td>0.0080</td>
<td>0.1153</td>
<td>0.0069</td>
<td>0.0659</td>
<td>0.0014</td>
<td>0.0499</td>
</tr>
<tr>
<td>7</td>
<td>0.0114</td>
<td>0.1067</td>
<td>0.1068</td>
<td>0.7256</td>
<td>0.0979</td>
<td>0.6941</td>
<td>0.0966</td>
<td>0.6896</td>
</tr>
<tr>
<td>8</td>
<td>0.0082</td>
<td>0.0751</td>
<td>0.0486</td>
<td>0.4836</td>
<td>0.0430</td>
<td>0.4533</td>
<td>0.0419</td>
<td>0.4472</td>
</tr>
<tr>
<td>9</td>
<td>0.0063</td>
<td>0.0141</td>
<td>0.0082</td>
<td>0.1277</td>
<td>0.0107</td>
<td>0.1847</td>
<td>0.0075</td>
<td>0.0015</td>
</tr>
<tr>
<td>10</td>
<td>0.0082</td>
<td>0.0751</td>
<td>0.0486</td>
<td>0.4836</td>
<td>0.0430</td>
<td>0.4533</td>
<td>0.0419</td>
<td>0.4472</td>
</tr>
<tr>
<td>11</td>
<td>0.0171</td>
<td>0.0733</td>
<td>0.0523</td>
<td>0.4897</td>
<td>0.0436</td>
<td>0.4450</td>
<td>0.0439</td>
<td>0.4466</td>
</tr>
</tbody>
</table>

*Average of 20 replications.
*Factors influencing the estimates for each case are defined in Table 1.

Implications

This research demonstrated that the real importance of using a threshold model was to estimate variance components for categorical data. The loss in accuracy of genetic parameter estimates obtained by using the outward scale in linear models, compared with threshold models on the underlying scale of continuous variation, cannot be sacrificed. Accuracy increased as number of categories for a categorical response variable increased, frequency distribution of the categories became more normal, heritability in the underlying scale increased, and with balanced data. Larger amounts of the additive genetic variance in the observed scale can be used as an indication of prediction accuracy.

Literature Cited


