

Prediction of genetic merit from data on binary and quantitative variates with an application to calving difficulty, birth weight and pelvic opening

J.L. FOULLEY, D. GIANOLA* and R. THOMPSON**

*I.N.R.A., Station de Génétique quantitative et appliquée,
Centre de Recherches zootechniques, F 78350 Jouy-en-Josas*

**Department of Animal Science, University of Illinois, Urbana, Illinois 61801, U.S.A.,*

***A.R.C. Unit of Statistics, University of Edinburgh, Mayfield Road, Edinburgh, EH9 3JZ,
Scotland*

Summary

A method of prediction of genetic merit from jointly distributed quantal and quantitative responses is described. The probability of response in one of two mutually exclusive and exhaustive categories is modeled as a non-linear function of classification and « risk » variables. Inferences are made from the mode of a posterior distribution resulting from the combination of a multivariate normal density, a priori, and a product binomial likelihood function. Parameter estimates are obtained with the Newton-Raphson algorithm, which yields a system similar to the mixed model equations. « Nested » Gauss-Seidel and conjugate gradient procedures are suggested to proceed from one iterate to the next in large problems. A possible method for estimating multivariate variance (covariance) components involving, jointly, the categorical and quantitative variates is presented. The method was applied to prediction of calving difficulty as a binary variable with birth weight and pelvic opening as « risk » variables in a *Blonde d'Aquitaine* population.

Key-words : sire evaluation, categorical data, non-linear models, prediction, Bayesian methods.

Résumé

Prédiction génétique à partir de données binaires et continues : application aux difficultés de vêlage, poids à la naissance et ouverture pelvienne.

Cet article présente une méthode de prédiction de la valeur génétique à partir d'observations quantitatives et qualitatives. La probabilité de réponse selon l'une des deux modalités exclusives et exhaustives envisagées est exprimée comme une fonction non linéaire d'effets de facteurs d'incidence et de variables de risque. L'inférence statistique repose sur le mode de la distribution a posteriori qui combine une densité multinormale a priori et une fonction de vraisemblance produit de binomiales. Les estimations sont calculées à partir de l'algorithme de Newton-Raphson qui conduit à un système d'équations similaires à celles du modèle mixte. Pour les gros fichiers, on suggère des méthodes itératives de résolution telles que celles de Gauss-Seidel et du gradient conjugué. On propose également une méthode d'estimation des composantes de variances et covariances relatives aux variables discrètes et continues. Enfin, la méthodologie présentée est illustrée par une application numérique qui a trait à la prédiction des difficultés de vêlage en race bovine *Blonde d'Aquitaine* utilisant d'une part, l'appréciation tout-ou-rien du caractère, et d'autre part, le poids à la naissance du veau et l'ouverture pelvienne de la mère comme des variables de risque.

Mots-clés : Évaluation des reproducteurs, données discrètes, modèle non linéaire, prédiction, méthode bayésienne.

I. Introduction

In many animal breeding applications, the data comprise observations on one or more quantitative variates and on categorical responses. The probability of «successful» outcome of the discrete variate, e.g., survival, may be a non-linear function of genetic and non-genetic variables (sire, breed, herd-year) and may also depend on quantitative response variates. A possible course of action in the analysis of this type of data might be to carry out a multiple-trait evaluation regarding the discrete trait as if it were continuous, and then utilizing available linear methodology (HENDERSON, 1973). Further, the model for the discrete trait should allow for the effects of the quantitative variates. In addition to the problems of describing discrete variation with linear models (COX, 1970; THOMPSON, 1979; GIANOLA, 1980), the presence of stochastic «regressors» in the model introduces a complexity which animal breeding theory has not addressed.

This paper describes a method of analysis for this type of data based on a Bayesian approach; hence, the distinction between «fixed» and «random» variables is circumvented. General aspects of the method of inference are described in detail to facilitate comprehension of subsequent developments. An estimation algorithm is developed, and we consider some approximations for posterior inference and fit of the model. A method is proposed to estimate jointly the components of variance and covariance involving the quantitative and the categorical variates. Finally, procedures are illustrated with a data set pertaining to calving difficulty (categorical), birth weight and pelvic opening.

II. Method of inference : general aspects

Suppose the available data pertain to three random variables: two quantitative (e.g., calf's birth weight and dam's pelvic opening) and one binary (e.g., easy vs. difficult calving). Let the data for birth weight and dam's pelvic opening be represented by the vectors y_1 and y_2 , respectively. Those for calving difficulty are represented by a set Y of indicator variables describing the configuration of the following $s \times 2$ contingency table:

Row	Category of response	
	Easy calving	Difficult calving
1	n_{11}	$n_{1.} - n_{11}$
2	n_{21}	$n_{2.} - n_{21}$
:	:	:
i	n_{i1}	$n_{i.} - n_{i1}$
:	:	:
s	n_{s1}	$n_{s.} - n_{s1}$

where the s rows indicate conditions affecting individual or grouped records. The two categories of response are mutually exclusive and exhaustive, and the number of observations in each row, $n_{i.} \neq 0$, is assumed fixed. The random quantity n_{i1} (or, conversely, $n_{i.} - n_{i1}$) can be null, so contingency tables where $n_{i.} = 1$, for $i = 1, \dots, s$, are allowed. The data can be represented symbolically by the vector $Y' = (Y_1, Y_2, \dots, Y_s)$, where $Y_i = \sum_{r=1}^{n_i} Y_{ir}$ with Y_{ir} being an indicator variable equal to 1 if a response occurs and zero otherwise.

The data Y, y_1 and y_2 , and a parameter vector θ are assumed to have a joint density $f(Y, y_1, y_2, \theta)$ written as

$$f(Y, y_1, y_2, \theta) = f_2(Y, y_1, y_2 | \theta) \cdot f_1(\theta). \tag{1}$$

where $f_1(\theta)$ is the marginal or a priori density of θ . From (1)

$$f_4(\theta | Y, y_1, y_2) = f_2(Y, y_1, y_2 | \theta) \cdot f_1(\theta) / f_3(Y, y_1, y_2) \tag{2}$$

where $f_3(Y, y_1, y_2)$ is the marginal density of the data, i.e., with θ integrated out, and $f_4(\theta | Y, y_1, y_2)$ is the a posteriori density of θ . As $f_3(Y, y_1, y_2)$ does not depend on θ , one can write (2) as

$$f_4(\theta | Y, y_1, y_2) \propto f_2(Y, y_1, y_2 | \theta) \cdot f_1(\theta) \tag{3}$$

which is Bayes theorem in the context of our setting. Equation (3) states that inferences can be made a posteriori by combining prior information with data translated to the posterior density via the likelihood function $f_2(Y, y_1, y_2 | \theta)$. The dispersion of θ reflects the a priori relative uncertainty about θ , this based on the results of previous data or experiments. If a new experiment is conducted, new data are combined with the prior density to yield the posterior. In turn, this becomes the a priori density for further experiments. In this form, continued iteration with (3) illustrates the process of knowledge accumulation (CORNFIELD, 1969). Comprehensive discussions of the merits, philosophy and limitations of Bayesian inference have been presented by CORNFIELD (1969), and LINDLEY & SMITH (1972). The latter argued in the context of linear models that (3) leads to estimates which may be substantially improved from those arising in the method of least-squares. Equation (3) is taken in this paper as a point of departure for a method of estimation similar to the one used in early developments of mixed model prediction (HENDERSON *et al.*, 1959). Best linear unbiased predictors could also be derived following Bayesian considerations (RÖNNINGEN, 1971; DEMPFILE, 1977).

The Bayes estimator of θ is the vector $\hat{\theta}$ minimizing the expected a posteriori risk

$$R(\hat{\theta}; Y, y_1, y_2) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} l(\hat{\theta}, \theta) f_4(\theta | Y, y_1, y_2) d(\theta) \tag{4}$$

where $l(\hat{\theta}, \theta)$ is a loss function (MOOD & GRAYBILL, 1963). If the loss is quadratic

$$l(\hat{\theta}, \theta) = \sum_{i=1}^k (\hat{\theta}_i - \theta_i)^2 = (\hat{\theta} - \theta)'(\hat{\theta} - \theta) \tag{5}$$

then

$$\begin{aligned} \frac{\partial R(\hat{\theta}; Y, y_1, y_2)}{\partial \hat{\theta}} &= \frac{\partial}{\partial \hat{\theta}} \left\{ \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} [\hat{\theta}'\hat{\theta} - \hat{\theta}'\theta - \theta'\hat{\theta} + \theta'\theta] f_4(\theta | Y, y_1, y_2) d\theta \right\} \\ &= 2[\hat{\theta} - E(\theta | Y, y_1, y_2)] \end{aligned} \tag{6}$$

Equating (6) to zero, yields $\hat{\theta} = E(\theta | Y, y_1, y_2)$. Note that differentiating (6) with respect to $\hat{\theta}$ yields a positive number, i.e., $\hat{\theta}$ minimizes the expected posterior risk, and $\hat{\theta}$ is identical to the best predictor of θ in the squared-error sense of HENDERSON (1973). Unfortunately, calculating $\hat{\theta}$ requires deriving the conditional density of θ given Y, y_1 and y_2 , and then computing the conditional expectation. In practice, this is difficult or impossible to execute as discussed by HENDERSON (1973). In view of these difficulties, LINDLEY & SMITH (1972) have suggested to approximate the posterior mean by the mode of the posterior density; if the posterior is unimodal and approximately symmetric,

its mode will be close to the mean. HARVILLE (1977) has pointed out, that if an improper prior is used in place of the « true » prior, the posterior mode has the advantage over the posterior mean, of being less sensitive to the tails of the posterior density.

In (3), it is convenient to write

$$f_2(\mathbf{Y}, y_1, y_2 | \boldsymbol{\theta}) = f_6(\mathbf{Y} | y_1, y_2, \boldsymbol{\theta}) \cdot f_5(y_1, y_2 | \boldsymbol{\theta}) \tag{7}$$

so the log of the posterior density can be written as

$$\ln[f_4(\boldsymbol{\theta} | \mathbf{Y}, y_1, y_2)] = \ln[f_6(\mathbf{Y} | y_1, y_2, \boldsymbol{\theta})] + \ln[f_5(y_1, y_2 | \boldsymbol{\theta})] + \ln[f_1(\boldsymbol{\theta})] + \text{const.} \tag{8}$$

III. Model

A. Categorical variate

The probability of response (e.g., easy calving) for the i^{th} row of the contingency table can be written as some cumulative distribution function with an argument peculiar to this row. Possibilities (GIANOLA & FOULLEY, 1983) are the standard normal and logistic distribution functions. In the first case, the probability of response is

$$P_{i1} = \Phi(\mu_i) = \int_{-\infty}^{\mu_i} \phi(x) dx \tag{9}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and distribution functions of a standard normal variate, respectively, and μ_i is a location variable. In the logistic case,

$$P_{i1} = [1 + e^{-\mu_i^*}]^{-1} \tag{10}$$

The justification of (9) and (10) is that they provide a liaison with the classical threshold model (DEMPSTER & LERNER, 1950; GIANOLA, 1982). If an easy calving occurs whenever the realized value of an underlying normal variable, $z \sim N(\delta_i, 1)$, is less than a fixed threshold value t , we can write for the i^{th} row

$$P_{i1} = \text{Prob} \{z < t\} = \int_{-\infty}^{t - \delta_i} \phi(v) dv = \Phi(t - \delta_i) \tag{11a}$$

Letting $\mu_i = t - \delta_i$, $\mu_i + 5$ is the probit transformation used in dose-response relationships (FINNEY, 1952) ; defining $\mu_i^* = \mu_i \pi / \sqrt{3}$, then

$$[1 + e^{-\mu_i^*}]^{-1} = \Phi(\mu_i^* \sqrt{3} / \pi) \tag{11b}$$

For $-5 < \mu_i < 5$, the difference between the left and right hand sides of (11b) does not exceed .022, being negligible from a practical point of view.

Suppose that a normal function is chosen to describe the probability of response. Let y_{i3} be the underlying variable, which under the conditions of the i^{th} row of the contingency table, is modeled as

$$y_{i3} = \mathbf{x}'_{i3} \boldsymbol{\beta}_3 + \mathbf{z}'_{i3} \mathbf{u}_3 + e_{i3} \tag{12a}$$

where \mathbf{x}'_{i3} and \mathbf{z}'_{i3} are known row vectors, $\boldsymbol{\beta}_3$ and \mathbf{u}_3 are unknown vectors, and e_{i3} is a residual. Likewise, the models for birth weight and pelvic opening are

$$y_{i1} = \mathbf{x}'_{i1} \boldsymbol{\beta}_1 + \mathbf{z}'_{i1} \mathbf{u}_1 + e_{i1} \tag{12b}$$

and

$$y_{i2} = \mathbf{x}'_{i2}\boldsymbol{\beta}_2 + \mathbf{z}'_{i2}\mathbf{u}_2 + e_{i2} \tag{12c}$$

Define μ_i in (9) as

$$\begin{aligned} \mu_i &= E(y_{i3} | \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \mathbf{u}_1, \mathbf{u}_2, \boldsymbol{\beta}_3, \mathbf{u}_3, y_{i1}, y_{i2}) \\ &= \mathbf{x}'_{i3}\boldsymbol{\beta}_3 + \mathbf{z}'_{i3}\mathbf{u}_3 + E(e_{i3} | e_{i1}, e_{i2}) \end{aligned} \tag{13}$$

which holds if e_{i3} is correlated only with e_{i1} and e_{i2} . In a multivariate normal setting

$$E(e_{i3} | e_{i1}, e_{i2}) = [\rho_{31} \ \rho_{32}] \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \sigma_{e_3}/\sigma_{e_1} & 0 \\ 0 & \sigma_{e_3}/\sigma_{e_2} \end{bmatrix} \begin{bmatrix} e_{i1} \\ e_{i2} \end{bmatrix} \tag{14}$$

where the ρ_{ij} 's and the σ_{e_i} 's are residual correlations and residual standard deviations, respectively. Similarly

$$\begin{aligned} \text{Var}(y_{i3} | \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \mathbf{u}_1, \mathbf{u}_2, \boldsymbol{\beta}_3, \mathbf{u}_3, y_{i1}, y_{i2}) &= \text{Var}(e_{i3} | e_{i1}, e_{i2}) \\ &= \sigma_e^2 (1 - \rho_{3.12}^2) \end{aligned} \tag{15}$$

where $\rho_{3.12}^2$ is the fraction of the residual variance of the underlying variable explained by a linear relationship with e_{i1} and e_{i2} . Since the unit of measurement in the conditional distribution of the underlying variate given $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \mathbf{u}_1, \mathbf{u}_2, \boldsymbol{\beta}_3, \mathbf{u}_3, y_{i1}$ and y_{i2} is the standard deviation, then (14) can be written as

$$\begin{aligned} E(e_{i3} | e_{i1}, e_{i2}) &= [\rho_{31}\rho_{32}] \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1/\sigma_{e_1} & 0 \\ 0 & 1/\sigma_{e_2} \end{bmatrix} \frac{1}{\sqrt{1 - \rho_{3.12}^2}} \begin{bmatrix} e_{i1} \\ e_{i2} \end{bmatrix} \\ &= b_1 e_{i1} + b_2 e_{i2} \end{aligned} \tag{16}$$

Hence, (13) can be written in matrix notation as

$$\boldsymbol{\mu} = \mathbf{X}_3\boldsymbol{\beta}_3 + \mathbf{Z}_3\mathbf{u}_3 + b_1\mathbf{e}_1 + b_2\mathbf{e}_2 \tag{18}$$

$$= \mathbf{X}_3\boldsymbol{\beta}_3 + \mathbf{Z}_3\mathbf{u}_3 + \sum_{j=1}^2 b_j (y_j - \mathbf{X}_j\boldsymbol{\beta}_j - \mathbf{Z}_j\mathbf{u}_j) \tag{19}$$

where $\mathbf{X}_1, \mathbf{X}_2, \mathbf{Z}_1$ and \mathbf{Z}_2 are known matrices arising from writing (12b) and (12c) as vectors. Now, suppose for simplicity that \mathbf{X}_3 is a matrix such that all factors and levels in \mathbf{X}_1 and \mathbf{X}_2 are represented in \mathbf{X}_3 and let $\mathbf{Z}_1 \equiv \mathbf{Z}_2 \equiv \mathbf{Z}_3$. Write

$$\mathbf{X}_1 = \mathbf{X}_3\mathbf{Q}_1; \quad \mathbf{X}_2 = \mathbf{X}_3\mathbf{Q}_2$$

where \mathbf{Q}_1 and \mathbf{Q}_2 are matrices of operators obtained by deleting columns of identity matrices of appropriate order. Thus, (19) can be written as

$$\begin{aligned} \boldsymbol{\mu} &= \mathbf{X}_3(\boldsymbol{\beta}_3 - b_1\mathbf{Q}_1\boldsymbol{\beta}_1 - b_2\mathbf{Q}_2\boldsymbol{\beta}_2) + \mathbf{Z}_3(\mathbf{u}_3 - b_1\mathbf{u}_1 - b_2\mathbf{u}_2) \\ &\quad + b_1\mathbf{y}_1 + b_2\mathbf{y}_2 \end{aligned} \tag{20}$$

Letting $\boldsymbol{\tau} = \boldsymbol{\beta}_3 - \sum_{j=1}^2 b_j\mathbf{Q}_j\boldsymbol{\beta}_j$ and $\boldsymbol{\nu} = \mathbf{u}_3 - \sum_{j=1}^2 b_j\mathbf{u}_j$, (20) can be expressed as

$$\boldsymbol{\mu} = \mathbf{X}_3\boldsymbol{\tau} + \mathbf{Z}_3\boldsymbol{\nu} + b_1\mathbf{y}_1 + b_2\mathbf{y}_2 \tag{21}$$

Note that if $b_1 = b_2 = 0$, then $\boldsymbol{\tau} = \boldsymbol{\beta}_3$, $\boldsymbol{\nu} = \mathbf{u}_3$, and (21) is equal to the expectation of (12a).

Given $\boldsymbol{\mu}$, the indicator variables \mathbf{Y} are assumed to be conditionally independent, and the likelihood function is taken as product binomial so

$$\ln[f(\mathbf{Y} | \boldsymbol{\mu})] = \sum_{i=1}^s [n_{i1}\ln(P_{i1}) + (n_i - n_{i1})\ln(1 - P_{i1})] \tag{22}$$

Now

$$f(\mathbf{Y}|\boldsymbol{\mu}) \propto f(\mathbf{Y}|y_1, y_2, \boldsymbol{\theta}^*) \tag{23}$$

where $\boldsymbol{\theta}^* = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{b}_1, \mathbf{b}_2]$. Also

$$f(\mathbf{Y}|y_1, y_2, \boldsymbol{\theta}^*) \propto f(\mathbf{Y}|y_1, y_2, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\tau}, \mathbf{u}_1, \mathbf{u}_2, \boldsymbol{\nu}, \mathbf{b}_1, \mathbf{b}_2) \tag{24}$$

Letting $\boldsymbol{\theta}' = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\tau}, \mathbf{u}_1, \mathbf{u}_2, \boldsymbol{\nu}, \mathbf{b}_1, \mathbf{b}_2]$, then from (23) and (24)

$$f(\mathbf{Y}|\boldsymbol{\mu}) \propto f(\mathbf{Y}|y_1, y_2, \boldsymbol{\theta}^*) \propto f_6(\mathbf{Y}|y_1, y_2, \boldsymbol{\theta}) \tag{25}$$

B. Conditional density of « risk » variables.

The conditional density of y_1 and y_2 given $\boldsymbol{\theta}$ is assumed to be multivariate normal with location and dispersion following from (12b) and (12c)

$$E \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} | \boldsymbol{\theta} = \begin{bmatrix} \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{Z}_1\mathbf{u}_1 \\ \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{Z}_2\mathbf{u}_2 \end{bmatrix} \tag{26}$$

$$\text{Var} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} | \boldsymbol{\theta} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{bmatrix} \tag{27}$$

where (27) is a non-singular known covariance matrix. Letting \mathbf{R}^{11} , \mathbf{R}^{12} , \mathbf{R}^{21} and \mathbf{R}^{22} be respective partitions of the inverse of (27), one can write

$$\ln[f_S(y_1, y_2|\boldsymbol{\theta})] = -\frac{1}{2} \sum_{i=1}^2 \sum_{j=1}^2 (y_i - \mathbf{X}_i\boldsymbol{\beta}_i - \mathbf{Z}_i\mathbf{u}_i) \mathbf{R}^{ij} (y_j - \mathbf{X}_j\boldsymbol{\beta}_j - \mathbf{Z}_j\mathbf{u}_j) + \text{const.} \tag{28}$$

C. Prior density.

In this paper we assume that the residual covariance matrix

$$\text{Var} \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \mathbf{R}_{13} \\ \mathbf{R}_{21} & \mathbf{R}_{22} & \mathbf{R}_{23} \\ \mathbf{R}_{31} & \mathbf{R}_{32} & \mathbf{R}_{33} \end{bmatrix}$$

is known. From (16) and (17), this implies that \mathbf{b}_1 and \mathbf{b}_2 are also known. Therefore,

$$f(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\tau}, \mathbf{u}_1, \mathbf{u}_2, \boldsymbol{\nu}, \mathbf{b}_1, \mathbf{b}_2) \propto f(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\tau}, \mathbf{u}_1, \mathbf{u}_2, \boldsymbol{\nu}) \tag{28}$$

and the vector of unknowns becomes $\boldsymbol{\theta}' = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\tau}, \mathbf{u}_1, \mathbf{u}_2, \boldsymbol{\nu}]$

Let $\boldsymbol{\theta}' = [\boldsymbol{\beta}', \mathbf{u}']$ where $\boldsymbol{\beta}' = [\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2, \boldsymbol{\tau}']$ and $\mathbf{u}' = [\mathbf{u}'_1, \mathbf{u}'_2, \boldsymbol{\nu}']$. A priori $\boldsymbol{\theta}$ follows multivariate normal distribution

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\beta}' \\ \mathbf{u}' \end{bmatrix} \sim N \left(\begin{bmatrix} \boldsymbol{\alpha}' \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_c \end{bmatrix} \right) \tag{29}$$

where

$$\text{Var}(\mathbf{u}) = \text{Var} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 - \mathbf{b}_1\mathbf{u}_1 - \mathbf{b}_2\mathbf{u}_2 \end{bmatrix} = \mathbf{G}_c \tag{30}$$

Further

$$G_c = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -b_1 I & -b_2 I & I \end{bmatrix} \begin{bmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{bmatrix} \begin{bmatrix} I & 0 & -b_1 I \\ 0 & I & -b_2 I \\ 0 & 0 & I \end{bmatrix} \quad (31)$$

with $Cov(u_i, u'_j) = G_{ij}$ ($i, j = 1, \dots, 3$). Note that G_c depends on b_1 and b_2 ; when $b_1 = b_2 = 0$, it follows from (30) that $G_c = \{G_{ij}\}$. Now

$$\ln[f_1(\theta)] = -\frac{1}{2} \{(\beta - \alpha)' \Gamma^{-1} (\beta - \alpha) + u' G_c^{-1} u\} + \text{const.} \quad (32)$$

where $G_c^{-1} = \{G_c^{ij}\}$ ($i, j = 1, \dots, 3$). Prior knowledge about β is assumed to be vague so $\Gamma \rightarrow \infty$ and $\Gamma^{-1} \rightarrow 0$. Therefore

$$\ln[f_1(\theta)] = -\frac{1}{2} \{u' G_c^{-1} u\} + \text{const.} \quad (33).$$

IV. Estimation

The terms of the log-posterior density in (8) are given in equations (22), (28) and (33). To obtain the mode of the posterior density, the derivatives of (8) with respect to θ are equated to zero. The resulting system of equations is not linear in θ and an iterative solution is required. Letting $L(\theta)$ be the log of the posterior density, the Newton-Raphson algorithm (DAHLQUIST & BJÖRCK, 1974) consists of iterating with

$$\hat{\theta}^{(i)} = \hat{\theta}^{(i-1)} - \left[\left[\frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} \right] \right]_{\theta = \hat{\theta}^{(i-1)}}^{-1} \left[\left[\frac{\partial L(\theta)}{\partial \theta} \right] \right]_{\theta = \hat{\theta}^{(i-1)}} \quad (34).$$

Note that the inverse of the matrix of second partial derivatives exists as β can be uniquely defined, e.g., with X_i having full-column rank, $i = 1, \dots, 3$. It is convenient to write (34) as

$$\left[\left[-\frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'} \right] \right]_{\theta = \hat{\theta}^{(i-1)}} (\hat{\theta}^{(i)} - \hat{\theta}^{(i-1)}) = \left[\left[\frac{\partial L(\theta)}{\partial \theta} \right] \right]_{\theta = \hat{\theta}^{(i-1)}} \quad (35).$$

A. First derivatives.

Differentiating (8) with respect to the elements of θ yields

$$\frac{\partial L(\theta)}{\partial \beta_i} = X_i' \sum_{j=1}^2 R^{ij} (y_j - X_j \beta_j - Z_j u_j), \quad i = 1, 2 \quad (36)$$

and

$$\frac{\partial L(\theta)}{\partial u_i} = Z_i' \sum_{j=1}^2 R^{ij} (y_j - X_j \beta_j - Z_j u_j) - \sum_{j=1}^2 G_c^{ij} u_j - G_c^{i3} v \quad (37).$$

The derivatives of $L(\theta)$ with respect to τ and v are slightly different

$$\frac{\partial L(\theta)}{\partial \tau} = \sum_{i=1}^s \left[n_{ii} \frac{\phi(\mu_i)}{P_{i1}} - (n_i - n_{i1}) \frac{\phi(\mu_i)}{1 - P_{i1}} \right] x_{i3} \quad (38)$$

where \mathbf{x}'_{i3} is the i^{th} row of \mathbf{X}_3 , and

$$\frac{\partial \mathbf{L}(\boldsymbol{\theta})}{\partial \mathbf{v}} = \sum_{i=1}^s \left[n_{i1} \frac{\phi(\mu_i)}{P_{i1}} - (n_{i.} - n_{i1}) \frac{\phi(\mu_i)}{1 - P_{i1}} \right] \mathbf{z}_{i3} - \sum_{j=1}^2 \mathbf{G}_c^{3j} \mathbf{u}_j - \mathbf{G}_c^{33} \mathbf{v}. \quad (39)$$

Now, let \mathbf{v} be a $s \times 1$ vector with elements

$$v_j = -\{n_{j1}i_{j1} + (n_j - n_{j1})i_{j2}\}, \quad j = 1, \dots, s$$

where $i_{j1} = -\phi(\mu_j)/P_{j1}$ and $i_{j2} = \phi(\mu_j)/(1 - P_{j1})$, and note that v_j is the opposite of the sum of normal scores for the j^{th} row. Then

$$\frac{\partial \mathbf{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\tau}} = \mathbf{X}'_3 \mathbf{v} \quad (40)$$

and

$$\frac{\partial \mathbf{L}(\boldsymbol{\theta})}{\partial \mathbf{v}} = \mathbf{Z}'_3 \mathbf{v} - \sum_{j=1}^2 \mathbf{G}_c^{3j} \mathbf{u}_j - \mathbf{G}_c^{33} \mathbf{v}. \quad (41)$$

B. Second derivatives

The symmetric matrix of second partial derivatives can be deduced from equations (36) through (41). Explicitly

$$\frac{\partial^2 \mathbf{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_i \partial \boldsymbol{\beta}'_i} = -\mathbf{X}'_i \mathbf{R}^u \mathbf{X}_i; \quad i = 1, 2 \quad (42 a)$$

$$\frac{\partial^2 \mathbf{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_1 \partial \boldsymbol{\beta}'_2} = -\mathbf{X}'_1 \mathbf{R}^{12} \mathbf{X}_2 \quad (42 b)$$

$$\frac{\partial^2 \mathbf{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_i \partial \mathbf{u}'_j} = -\mathbf{X}'_i \mathbf{R}^{ij} \mathbf{Z}_j; \quad i = 1, 2, \quad j = 1, 2 \quad (42 c)$$

$$\frac{\partial^2 \mathbf{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_i \partial \boldsymbol{\tau}'} = \mathbf{0}; \quad i = 1, 2 \quad (42 d)$$

$$\frac{\partial^2 \mathbf{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_i \partial \mathbf{v}'} = \mathbf{0}; \quad i = 1, 2 \quad (42 e)$$

$$\frac{\partial^2 \mathbf{L}(\boldsymbol{\theta})}{\partial \mathbf{u}_i \partial \mathbf{u}'_j} = -\mathbf{Z}'_i \mathbf{R}^{ij} \mathbf{Z}_j - \mathbf{G}_c^{ij}; \quad i = 1, 2, \quad j = 1, 2 \quad (42 f)$$

$$\frac{\partial^2 \mathbf{L}(\boldsymbol{\theta})}{\partial \mathbf{u}_i \partial \boldsymbol{\tau}'} = \mathbf{0}; \quad i = 1, 2 \quad (42 g)$$

$$\frac{\partial^2 \mathbf{L}(\boldsymbol{\theta})}{\partial \mathbf{u}_i \partial \mathbf{v}'} = -\mathbf{G}^{i3}; \quad i = 1, 2 \quad (42 h)$$

$$\frac{\partial^2 \mathbf{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}'} = -\mathbf{X}'_3 \mathbf{W} \mathbf{X}_3 \quad (42 i)$$

$$\frac{\partial^2 \mathbf{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\tau} \partial \mathbf{v}'} = -\mathbf{X}'_3 \mathbf{W} \mathbf{Z}_3; \quad (42 j)$$

$$\frac{\partial^2 \mathbf{L}(\boldsymbol{\theta})}{\partial \mathbf{v} \partial \mathbf{v}'} = -\mathbf{Z}'_3 \mathbf{W} \mathbf{Z}_3 - \mathbf{G}_c^{33}. \quad (42 k)$$

In (42 *i*) through (42 *k*), **W** is an *sxs* diagonal matrix with elements

$$\begin{aligned} w_{jj} &= n_{j1}i_{j1}(i_{j1} - \mu_j) + (n_j - n_{j1})i_{j2}(i_{j2} - \mu_j) \\ &= \mu_j v_j + n_{j1}i_{j1}^2 + (n_j - n_{j1})i_{j2}^2, \quad j = 1, \dots, s. \end{aligned} \tag{43}$$

Note that
$$E(w_{jj} | \mu_j) = n_{j1}i_{j1}^2 + (n_j - n_{j1})i_{j2}^2 \tag{44}$$

indicating that calculations are somewhat simpler if «scoring» is used instead of Newton-Raphson.

C. Equations

Using the first and second derivatives in (36-41) and (42 *a*-42 *k*), respectively, equations (35) can be written after algebra as (45).

In (45), $\hat{\beta}_1^{[i]}$, $\hat{\beta}_2^{[i]}$, $\hat{u}_1^{[i]}$ and $\hat{u}_2^{[i]}$ are solutions at the [*i*th] iterate while the Δ 's are corrections at the [*i*th] iterate pertaining to the parameters affecting the probability of response, e.g., $\Delta_\tau^{[i]} = \hat{\tau}^{[i]} - \hat{\tau}^{[i-1]}$. Iteration proceeds by first taking a guess for τ and v , calculating $W^{[0]}$ and $v^{[0]}$, amending the right hand-sides and then solving for the unknowns. The cycle is repeated until the solutions stabilize. Equations (45) can also be written as in (46). The similarity between (46) and the «mixed model equations» (HENDERSON, 1973) should be noted. The coefficient matrix and the «working» vector $y_3^{[i-1]}$ change in every iteration; note that $y_3^{[i-1]} = X_3 \tau^{[i-1]} + Z_3 v^{[i-1]} + (W^{[i-1]})^{-1} v^{[i-1]}$.

D. Solving the equations

In animal breeding practice, solving (45) or (46) poses a formidable numerical problem. The order of the coefficient matrix can be in the tens of thousands, and this difficulty arises in every iterate. As β_1 , β_2 , u_1 and u_2 are «nuisance» variables in this problem, the first step is to eliminate them from the system, if this is feasible. The order of the remaining equations is still very large in most animal breeding problems so direct inversion is not possible. At the *i*th iterate, the remaining equations can be written as

$$P^{[i-1]} \gamma^{[i]} = I^{[i-1]} \tag{47}$$

Next, decompose $P^{[i-1]}$ as the sum of three matrices $L^{[i-1]}$, $D^{[i-1]}$, $U^{[i-1]}$, which are lower triangular, diagonal and upper triangular, respectively. Therefore

$$\gamma^{[i]} = \{D^{[i-1]}\}^{-1} \{I^{[i-1]} - L^{[i-1]} \gamma^{[i]} - U^{[i-1]} \gamma^{[i]}\}.$$

Now, for each iterate *i*, sub-iterate with

$$\gamma^{[i, j+1]} = \{D^{[i-1]}\}^{-1} \{I^{[i-1]} - L^{[i-1]} \gamma^{[i, j+1]} - U^{[i-1]} \gamma^{[i, j]}\} \tag{48}$$

for $j=0, 1, \dots$; iteration can start with $\gamma^{[i, 0]} = 0$. As this is a «nested» Gauss-Seidel iteration, with $P^{[i-1]}$ symmetric and positive definite

$$\lim_{j \rightarrow \infty} \gamma^{[i, j]} = \gamma^{[i]} \tag{49}$$

(VAN NORTON, 1960). Then, one needs to return to (47) and to the back solution, and work with (48). The cycle finishes when the solutions γ stabilize.

$$\begin{bmatrix} X_1'R^{11}X_1 \\ X_2'R^{21}X_1 \\ Z_1'R^{11}X_1 \\ Z_2'R^{21}X_1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} X_1'R^{12}X_2 \\ X_2'R^{22}X_2 \\ Z_1'R^{12}X_2 \\ Z_2'R^{22}X_2 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} X_1'R^{11}Z_1 \\ X_2'R^{21}Z_1 \\ Z_1'R^{11}Z_1 + G_c^{11} \\ Z_2'R^{21}Z_1 + G_c^{21} \\ 0 \\ G_c^{31} \end{bmatrix} \begin{bmatrix} X_1'R^{12}Z_2 \\ X_2'R^{22}Z_2 \\ Z_1'R^{12}Z_2 + G_c^{12} \\ Z_2'R^{22}Z_2 + G_c^{22} \\ 0 \\ G_c^{32} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ X_3'W^{(i-1)}X_3 \\ Z_3'W^{(i-1)}X_3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ G_c^{13} \\ G_c^{23} \\ X_3'W^{(i-1)}Z_3 \\ Z_3'W^{(i-1)}Z_3 + G_c^{33} \end{bmatrix} \\
 = \begin{bmatrix} X_1'(R^{11}y_1 + R^{12}y_2) \\ X_2'(R^{21}y_1 + R^{22}y_2) \\ Z_1'(R^{11}y_1 + R^{12}y_2) \\ Z_2'(R^{21}y_1 + R^{22}y_2) \\ X_3'y^{(i-1)} \\ Z_3'y^{(i-1)} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ G_c^{13}y^{(i-1)} \\ G_c^{23}y^{(i-1)} \\ 0 \\ G_c^{33}y^{(i-1)} \end{bmatrix} \tag{45}$$

$$\begin{bmatrix} X_1'R^{11}X_1 \\ X_2'R^{21}X_1 \\ Z_1'R^{11}X_1 \\ Z_2'R^{21}X_1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} X_1'R^{12}X_2 \\ X_2'R^{22}X_2 \\ Z_1'R^{12}X_2 \\ Z_2'R^{22}X_2 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} X_1'R^{11}Z_1 \\ X_2'R^{21}Z_1 \\ Z_1'R^{11}Z_1 + G_c^{11} \\ Z_2'R^{21}Z_1 + G_c^{21} \\ 0 \\ G_c^{31} \end{bmatrix} \begin{bmatrix} X_1'R^{12}Z_2 \\ X_2'R^{22}Z_2 \\ Z_1'R^{12}Z_2 + G_c^{12} \\ Z_2'R^{22}Z_2 + G_c^{22} \\ 0 \\ G_c^{32} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ X_3'W^{(i-1)}X_3 \\ Z_3'W^{(i-1)}X_3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ G_c^{13} \\ G_c^{23} \\ X_3'W^{(i-1)}Z_3 \\ Z_3'W^{(i-1)}Z_3 + G_c^{33} \end{bmatrix} \\
 = \begin{bmatrix} X_1'(R^{11}y_1 + R^{12}y_2) \\ X_2'(R^{21}y_1 + R^{22}y_2) \\ Z_1'(R^{11}y_1 + R^{12}y_2) \\ Z_2'(R^{21}y_1 + R^{22}y_2) \\ X_3'W^{(i-1)}y_3^{(i-1)} \\ Z_3'W^{(i-1)}y_3^{(i-1)} \end{bmatrix} \tag{46}$$

Another possibility would be to carry out nested iterations with the conjugate gradient method (BECKMAN, 1960). In the context of (47) the method involves :

a) Set

$$\mathbf{P}^{[0]} = \mathbf{r}^{[0]} = \mathbf{I}^{[i-1]} - \mathbf{P}^{[i-1]}\boldsymbol{\gamma}^{[i, 0]}, \quad i = 1, 2, \dots$$

where $\boldsymbol{\gamma}^{[i, 0]}$ is a guess, e.g., $\boldsymbol{\gamma}^{[i, 0]} = \mathbf{0}$.

b) Calculate successively

$$\begin{aligned} \text{[b.1]} \quad \alpha^{[j]} &= \mathbf{p}^{[j]}\mathbf{r}^{[j]} / \{\mathbf{p}^{[j]}\mathbf{P}^{[i-1]}\mathbf{p}^{[j]}\} \\ \text{[b.2]} \quad \boldsymbol{\gamma}^{[i, j+1]} &= \boldsymbol{\gamma}^{[i, j]} + \alpha^{[j]}\mathbf{p}^{[j]} \\ \text{[b.3]} \quad \mathbf{r}^{[j+1]} &= \mathbf{r}^{[j]} - \alpha^{[j]}\mathbf{p}^{[j]} \\ \text{[b.4]} \quad \lambda^{[j]} &= -\{\mathbf{r}^{[j+1]}\mathbf{P}^{[i-1]}\mathbf{p}^{[j]}\} / \{\mathbf{p}^{[j]}\mathbf{P}^{[i-1]}\mathbf{p}^{[j]}\} \\ \text{[b.5]} \quad \mathbf{p}^{[j+1]} &= \mathbf{r}^{[j+1]} + \lambda^{[j]}\mathbf{p}^{[j]} \end{aligned}$$

for $j=0, 1, \dots$, until $\boldsymbol{\gamma}^{[i, j]}$ stabilizes. When this occurs, $\mathbf{P}^{[i-1]}$ and $\mathbf{I}^{[i-1]}$ in (47) are amended, and the cycle with a new index for i is started from (a). The whole process stops when $\boldsymbol{\gamma}^{[i]}$ does not change between the $[i]$ and $[i+1]$ « main » rounds. While the number of operations per iterate is higher than with Gauss-Seidel (BECKMAN, 1960), the method is known to converge faster when $\mathbf{P}^{[i-1]}$ in (47) is symmetric and positive definite (personal communication, SAMEH, 1981).

V. Approximate posterior inference and model fit

As discussed by LINDLEY & SMITH (1972) in the context of linear models, the procedure does not provide standard errors a posteriori. LEONARD (1972), however, has pointed out that an approximation of the posterior density by a multivariate normal is « fairly accurate » in most regions of the space of $\boldsymbol{\theta}$, provided that none of the n_{i1} or $n_{i.} - n_{i1}$ are small. If this approximation can be justified, given any linear function of $\boldsymbol{\theta}$, say $\mathbf{t}'\boldsymbol{\theta}$, one can write, given the model

$$\begin{aligned} E[\mathbf{t}'\boldsymbol{\theta} | \mathbf{Y}, y_1, y_2] &= \mathbf{t}'\hat{\boldsymbol{\theta}} & (50 a) \\ \text{Var}[\mathbf{t}'\boldsymbol{\theta} | \mathbf{Y}, y_1, y_2] &= \mathbf{t}'\mathbf{C}\mathbf{t} & (50 b) \end{aligned}$$

where $\hat{\boldsymbol{\theta}}$ is the posterior mode and \mathbf{C} is the inverse of the coefficient matrix in (46); note that \mathbf{C} depends on the data through the matrix \mathbf{W} . Further

$$[\mathbf{t}'\boldsymbol{\theta} - \mathbf{t}'\hat{\boldsymbol{\theta}}] / (\mathbf{t}'\mathbf{C}\mathbf{t})^{1/2} \sim N(0, 1) \quad (50 c)$$

thus permitting probability statements about $\mathbf{t}'\boldsymbol{\theta}$. In many instances it will be impossible to calculate \mathbf{C} on computational grounds.

The probability of response for each of the rows in the contingency table can be estimated from (9) with μ evaluated at $\hat{\mu}$. Approximate standard errors of the estimates of response probabilities can be obtained from large sample theory. However, caution should be exercised as an approximation to an approximation is involved.

When cell counts are large, e.g., n_{i1} and $n_{i.} - n_{i1} > 5$, the statistic

$$\chi^2 = \sum_{i=1}^s \frac{(n_{i1} - n_i \hat{P}_{i1})^2}{n_i \hat{P}_{i1} (1 - \hat{P}_{i1})} \quad (51)$$

can be referred to a chi-square distribution with s -rank (X_3) degrees of freedom. Lack of fit may result from inadequate model specification in which case alternative models should be entertained.

VI. Unknown variance-covariance structure

The matrices $R_{ij}(i, j = 1, \dots, 3)$ and G_c are assumed known so that they are treated as nuisance arrays in (8) and (46). In animal breeding practice there are generally «good» estimates of these matrices so they could be used in (45) or (46) to proceed with the method, in the same way as in linear methodology (HENDERSON, 1973). The effect of replacing R and G_c matrices by estimates on the posterior distribution of θ is not known, and should be studied by Monte-Carlo methods.

If the analysis were to proceed in an entirely Bayesian context, prior distributions would need to be specified for the elements of these matrices. This is not addressed in the present paper as it does not appear clear what densities should be considered for the distribution of covariance components. For a discussion of Bayes estimation of variance components, see HILL (1965), TIAO & TAN (1965), TIAO & BOX (1967), LINDLEY & SMITH (1972) and HARVILLE (1977). LEONARD (1972) considered estimation of variance components with binomial data for a one-way model.

Equations (46) suggest methods for estimating variance and covariance components in this quantitative-categorical setting. Write

- a) $W^{(i)} = (A^{(i)})^2 = (\{\sqrt{w_{jj}^{(i)}}\})^2, j = 1, \dots, s$
- b) $X_3^{(i)} = A^{(i)}X_3$
- c) $Z_3^{(i)} = A^{(i)}Z_3$

and

d) $q_3^{(i)} = A^{(i)}y_3^{(i)}$

Equations (46) can then be written as (52) below.

$$\begin{bmatrix}
 X_1'R^{11}X_1 & X_1'R^{12}X_2 & X_1'R^{11}Z_1 & X_1'R^{12}Z_2 & 0 & 0 \\
 X_2'R^{21}X_1 & X_2'R^{22}X_2 & X_2'R^{21}Z_1 & X_2'R^{22}Z_2 & 0 & 0 \\
 Z_1'R^{11}X_1 & Z_1'R^{12}X_2 & Z_1'R^{11}Z_1 + G_c^{11} & Z_1'R^{12}Z_2 + G_c^{12} & 0 & G_c^{13} \\
 Z_2'R^{21}X_1 & Z_2'R^{22}X_2 & Z_2'R^{22}Z_1 + G_c^{21} & Z_2'R^{22}Z_2 + G_c^{22} & 0 & G_c^{23} \\
 0 & 0 & 0 & 0 & X_3^{(i)}X_3^{(i)'} & X_3^{(i)}Z_3^{(i)'} \\
 0 & 0 & G_c^{31} & G_c^{32} & Z_3^{(i)}X_3^{(i)'} & Z_3^{(i)}Z_3^{(i)'} + G_c^{33}
 \end{bmatrix}
 \begin{bmatrix}
 \beta_1^{(i+1)} \\
 \beta_2^{(i+1)} \\
 u_1^{(i+1)} \\
 u_2^{(i+1)} \\
 \tau^{(i+1)} \\
 v^{(i+1)}
 \end{bmatrix}
 =
 \begin{bmatrix}
 X_1'(R^{11}y_1 + R^{12}y_2) \\
 X_2'(R^{21}y_1 + R^{22}y_2) \\
 Z_1'(R^{11}y_1 + R^{12}y_2) \\
 Z_2'(R^{21}y_1 + R^{22}y_2) \\
 X_3^{(i)}q_3^{(i)} \\
 Z_3^{(i)}q_3^{(i)}
 \end{bmatrix}
 \tag{52}$$

The above equations suggest at each iterate the multivariate linear model

$$\begin{bmatrix}
 y_1 \\
 y_2 \\
 q^{(i)}
 \end{bmatrix}
 =
 \begin{bmatrix}
 X_1 & 0 & 0 \\
 0 & X_2 & 0 \\
 0 & 0 & X_3^{(i)}
 \end{bmatrix}
 \begin{bmatrix}
 \beta_1^{(i+1)} \\
 \beta_2^{(i+1)} \\
 \tau^{(i+1)}
 \end{bmatrix}
 +
 \begin{bmatrix}
 Z_1 & 0 & 0 \\
 0 & Z_2 & 0 \\
 0 & 0 & Z_3^{(i)}
 \end{bmatrix}
 \begin{bmatrix}
 u^{(i+1)} \\
 u_2^{(i+1)} \\
 v^{(i+1)}
 \end{bmatrix}
 +
 \begin{bmatrix}
 E_1^{(i)} \\
 E_2^{(i)} \\
 E_3^{(i)}
 \end{bmatrix}$$

$i = 0, 1 \dots$ (53)

with $\beta_1^{(i+1)}, \beta_2^{(i+1)}$ and $\tau^{(i+1)}$ «fixed» and $u_1^{(i+1)}, u_2^{(i+1)}, v^{(i+1)}$ and the E 's random, with covariance matrix

$$\text{Var} \begin{bmatrix} \mathbf{u}_1^{(i+1)} \\ \mathbf{u}_2^{(i+1)} \\ \mathbf{v}^{(i+1)} \\ \mathbf{E}_1^{(i)} \\ \mathbf{E}_2^{(i)} \\ \mathbf{E}_3^{(i)} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{11,c} & \mathbf{G}_{12,c} & \mathbf{G}_{13,c} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{G}_{21,c} & \mathbf{G}_{22,c} & \mathbf{G}_{23,c} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{G}_{31,c} & \mathbf{G}_{32,c} & \mathbf{G}_{33,c} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R}_{11} & \mathbf{R}_{12} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R}_{21} & \mathbf{R}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \quad i=0, 1 \dots \quad (54)$$

holding at every iterate. Note that the residual variance of $\mathbf{q}^{(i)}$ is unity so this part of the covariance structure does not need to be estimated. Provided that ρ_{31} and ρ_{32} are known, the method can be used to estimate the additive genetic covariance matrix between the quantitative traits and the hypothetical underlying variate with binary expression.

Expressions in (53) and (54) suggest that some of the methods for estimating variance and covariance components in linear models could be used to estimate the covariance structure in (54). One possibility would be to mimic the computations used in estimation via restricted maximum likelihood (SCHAEFFER *et al.*, 1978) for multivariate normal data. As computational feasibility is of paramount importance, a multivariate extension of Henderson's «simple» method (HENDERSON, 1980) could be useful here. However, this method does not preclude negative estimates of variance components. Estimation of genetic parameters in non-linear models is an open area of potential importance.

VII. Numerical application

Data were obtained from 47 *Blonde d'Aquitaine* heifers mated to the same bull and assembled to calve in the Casteljaloux Station, France. Each calving record included information on the following: region of origin and sire of the heifer, pelvic opening and season of calving, sex and birth weight of the calf, and calving difficulty score (1: normal birth, 2: slight assistance, 3: assisted, 4: mechanical aid, and 5: cesarean). For the purpose of the analysis, twin calves were excluded and calving difficulty was recoded as: a) «Easy» (scores 1, 2 and 3) or b) «Difficult» (scores 4 and 5). The data are presented in Table 1. As shown in Table 2, 23.4 % of the calvings were «difficult» and there were marked differences in the incidence of difficult calvings between sexes and maternal grandsires.

A. Models

Birth weight was modeled as

$$Y_{ijklm} = D_i + T_j + L_k + S_l + e_{ijklm} \quad (55 a)$$

where D_i is the effect of the i^{th} region of origin of the heifer ($i=1,2$), T_j is the effect of the j^{th} season of calving ($j=1,2$), L_k is the effect of the k^{th} sex of calf ($k=1$: male, 2 =female), S_l is the effect of the l^{th} sire of the heifer ($l=1, \dots, 6$), and e_{ijklm} is a residual. The vectors β_1 and \mathbf{u}_1 were defined as

$$\beta_1' = [D_1 + T_2 + L_2, D_2 + T_2 + L_2, T_1 - T_2, L_1 - L_2] \quad (55 b)$$

and

$$\mathbf{u}_1' = [S_1, S_2, S_3, S_4, S_5, S_6]. \quad (55 c)$$

TABLE 1
Calving difficulty records for Blonde d'Aquitaine heifers at Casteljalous France.
Données de vêlage des génisses de race Blonde d'Aquitaine contrôlées à Casteljalous, France.

Heifer Origin	Sire of heifer	Calving season	Pelvic opening (cm ²)	Sex ^(a) of calf	Calf's birth weight (kg)	Calving ^(b) difficulty	Heifer Origin	Sire of heifer	Calving season	Pelvic opening (cm ²)	Sex ^(a) of calf	Calf's birth weight (kg)	Calving ^(b) difficulty
1	1	1	328.0	M	41.0	E	1	4	2	328.0	M	47.0	D
1	1	1	304.5	M	37.5	E	1	4	2	344.0	F	51.0	D
1	1	1	354.8	F	41.5	E	1	4	2	243.0	F	39.0	E
1	1	2	374.0	F	40.0	E	2	4	1	333.3	M	44.5	E
1	1	2	285.0	F	43.0	E	1	5	1	346.5	M	40.5	E
1	1	2	310.0	F	42.0	E	1	5	1	304.5	F	43.5	E
1	1	2	270.0	F	35.0	E	1	5	2	396.0	M	42.5	E
2	1	1	336.0	F	46.0	E	1	5	2	285.0	M	48.8	D
2	1	1	374.0	F	40.5	E	1	5	2	328.0	M	38.5	D
2	1	2	346.5	F	39.0	E	1	5	2	346.5	M	52.0	E
1	2	1	346.5	M	41.4	E	1	5	2	344.0	F	48.0	E
1	2	1	333.3	M	43.0	E	2	5	1	357.0	F	41.0	E
1	2	2	346.5	F	34.0	D	2	5	1	357.0	M	50.5	D
1	2	2	307.5	M	47.0	D	2	5	2	317.8	M	43.7	D
1	2	2	315.0	M	42.0	E	2	5	2	346.5	M	51.0	D
2	2	2	300.0	M	44.5	E	1	6	1	290.0	F	51.6	D
2	2	2	273.0	M	49.0	E	1	6	1	260.0	M	45.3	D
1	3	1	341.0	M	41.6	E	1	6	1	357.0	F	36.5	E
2	3	1	300.0	M	36.0	E	1	6	2	354.8	M	50.5	E
2	3	1	328.0	F	42.7	E	1	6	2	273.0	M	46.0	D
2	3	2	266.0	F	32.5	E	1	6	2	315.0	M	45.0	E
2	3	2	302.3	F	44.4	E	1	6	2	256.5	F	36.0	E
2	3	2	320.0	M	46.0	E	2	6	1	317.8	F	43.5	E
							2	6	1	290.0	F	36.5	E

a) M: male; F: female.

b) D: difficult; E: easy.

TABLE 2
 Marginal mean (or frequency) of calving variables by level of factors considered.
 Moyenne (ou fréquence) marginale des variables de vêlage par niveau de facteur.

Factor		N°	Variable ^(a)		
			BW	PO	CD
Heifer origin	1	30	43.02	319.60	0.267
	2	17	43.02	321.48	0.177
Calving season	1	20	42.23	327.96	0.200
	2	27	43.61	314.59	0.259
Sex of calf	M	25	44.59	-	0.360
	F	22	41.24	-	0.091
Sire of heifer	1	10	40.55	328.28	0.000
	2	7	42.99	317.40	0.286
	3	6	40.53	309.55	0.000
	4	4	45.38	312.08	0.500
	5	11	45.46	338.98	0.364
	6	9	43.43	301.57	0.333
Total		47	43.02	320.28	0.234

(a) BW: birth weight in kg.; PO: pelvic opening in cm²; CD: frequency of difficult calvings.

The model for pelvic opening was

$$Z_{ijkl} = D'_i + T'_j + S'_k + e'_{ijkl} \tag{56 a}$$

where D'_i is the effect of the i^{th} department of origin of the heifer ($i=1,2$), T'_j is the effect of the j^{th} season of calving ($j=1,2$), S'_k is the effect of the k^{th} sire of heifer ($k=1, \dots, 6$) and e'_{ijkl} is a residual. The vectors β'_2 and u'_2 were defined as

$$\beta'_2 = [D'_1 + T'_2, D'_2 + T'_2, T'_1 - T'_2] \tag{56 b}$$

$$u'_2 = [S'_1, S'_2, S'_3, S'_4, S'_5, S'_6]. \tag{56 c}$$

The data in Table 1 can be regarded as a 47×2 contingency table, with rows corresponding to each record, and columns being « DIFFICULT » and « EASY » calvings. Hence, $n_i = 1$ for $i = 1, \dots, 47$, and $Y' = [Y_1, \dots, Y_{47}]$, with Y_i being a scalar variable with realized value 1 if a difficult calving occurs, or 0 otherwise. The probability of difficult calving for the i^{th} row was assumed a normal integral with argument modeled as

$$\mu_{i(jklm)} = D''_j + T''_k + L''_l + S''_m + b_1(Y_{ijklm} - 43.02) + b_2(Z_{ijklm} - 320.28) \tag{57 a}$$

where D''_j is the effect of the j^{th} department of origin ($j=1,2$), T''_k is the effect of the k^{th} season of calving ($k=1,2$), L''_l is the effect of the l^{th} sex ($l=1$: male, 2 =female), and S''_m is the effect of the m^{th} sire of the heifer; b_1 and b_2 are partial « regression » coefficients of the underlying variate on birth weight of the calf and pelvic opening of the heifer, respectively. These coefficients were assumed known with $b_1 = .1643$ and $b_2 = -.0184$; the logic for the choice of these values is presented in the following section. Note that as $\mu_{i(jklm)}$ increases, so does the probability of difficult calving; also, $\mu_{i(jklm)}$

increases with increased birth weight and decreases with increased pelvic opening. The vector τ and ν were then

$$\tau = [D'_1 + T'_2 + L'_2, D'_2 + T'_2 + L'_2, T'_1 - T'_2, L'_1 - L'_2] \quad (57 b)$$

$$\nu = [S'_1, S'_2, S'_3, S'_4, S'_5, S'_6]. \quad (57 c)$$

B. Conditional covariance

Given θ , the variance-covariance matrix of birth weight and pelvic opening is

$$\text{Var} \begin{bmatrix} y_1 | \theta \\ y_2 | \theta \end{bmatrix} = \begin{bmatrix} \sigma_{e_1}^2 & \sigma_{e_{12}} \\ \sigma_{e_{12}} & \sigma_{e_2}^2 \end{bmatrix} \otimes \mathbf{I}_{47 \times 47}$$

where \otimes is the Kronecker product. The values used for the residual covariance matrix were (MENISSIER & SAPA, personal communication): $\sigma_{e_1}^2 = 25$, $\sigma_{e_2}^2 = 1089$ and $\sigma_{e_{12}} = 41.25$.

The coefficients b_1 and b_2 were calculated as in (16) and (17) from $\rho_{12} = .25$, $\rho_{13} = .50$ and $\rho_{23} = -.30$; the residual variance in the underlying scale, which was set equal to 1, corresponds to (15). These values yielded $b_1 = .1643$ and $b_2 = -.0184$.

C. Prior distribution

The parameter vector for this problem was

$$\theta' = [\beta'_1 \beta'_2 \tau' u'_1 u'_2 \nu']. \quad (58)$$

Prior knowledge about β_1 , β_2 and τ was assumed to be vague. The covariance matrix of u_1 , u_2 , and ν was

$$\text{Var} \begin{bmatrix} u_1 \\ u_2 \\ \nu \end{bmatrix} = G_c \otimes \mathbf{I}_{6 \times 6} \quad (59)$$

where G_c is a 3×3 matrix calculated as in (31). The unconditional prior covariance matrix was taken as

$$\mathbf{G} = \{ \rho_{G_{ij}} \sigma_{u_i} \sigma_{u_j} \} \quad \begin{matrix} i = 1, \dots, 3 \\ j = 1, \dots, 3 \end{matrix} \quad (60)$$

where $\rho_{G_{ij}}$ is the genetic correlation between traits i and j in the underlying scale. The genetic correlations used were (MENISSIER & S.A.P.A, personal communication) : $\rho_{G_{13}} = .70$ and $\rho_{G_{23}} = -.50$. The standard deviations were calculated as

$$\sigma_{u_i} = \sigma_{e_i} / \sqrt{\lambda_i}, \quad i = 1, \dots, 3 \quad (61)$$

with $\lambda_i = (4 - h_i^2) / h_i^2$, and $h_1^2 = .15$, $h_2^2 = .40$ and $h_3^2 = .30$. Further

$$\sigma_{e_3} = 1 / \sqrt{1 - \rho_{3,12}^2} \quad (62)$$

with $\rho_{3,12}^2 = .4427$. We obtained

$$\mathbf{G} = \begin{bmatrix} .9740 & 3.7997 & .2635 \\ 3.7997 & 121.0000 & -2.0978 \\ .2635 & -2.0978 & .1455 \end{bmatrix}; \quad \mathbf{G}_c = \begin{bmatrix} .9740 & 3.7997 & .1734 \\ 3.7997 & 121.0000 & -.4956 \\ .1734 & -.4956 & .0260 \end{bmatrix}$$

Computations were also carried under the hypothesis of no «risk» relationship, i. e., $b_1 = b_2 = 0$. In this case, a different prior covariance matrix was used

$$G_0 = \begin{bmatrix} .9740 & 3.7997 & .1967 \\ 3.7997 & 121.0000 & -1.5661 \\ .1967 & -1.5661 & .0811 \end{bmatrix}$$

obtained from G by appropriate rescaling of elements. For example, and taking into account that $1/\sqrt{1 - \rho_{2,12}^2} = 1.3395$

$$.0811 = .1455/(1.3395)^2.$$

Note that $h_3^2 = 4 \times .0811/(1 + .0811) = .30$, $\rho_{G_{13}} = .70$ and $\rho_{G_{23}} = -.50$, as it should be. In this instance, the μ_i 's are expressed in standard deviation units of the underlying variate for calving difficulty «unadjusted» for residual variation in birth weight and pelvic opening. In order to compare estimates obtained under $b_1 \neq 0$ and $b_2 \neq 0$ with those calculated with $b_1 = b_2 = 0$, the latter were multiplied by 1.3395 to express them in the same scale.

D. Logistic approximation

In each of the two cases ($b_1 \neq 0$ and $b_2 \neq 0$, and $b_1 = b_2 = 0$) computations were also conducted using the logistic approximation in (11 b). Since the residual variance in the logistic scale is $\pi^2/3$, the prior covariance matrices G_c and G_0 discussed in the previous section were rescaled as

$$G_c^* = LG_cL'; \quad G_0^* = LG_0L'$$

where L is a 3×3 diagonal matrix with elements 1, 1 and $\pi/\sqrt{3}$. Solutions to (45) and (46) obtained with the logistic approximation were then divided by $\pi/\sqrt{3}$ to make them comparable to those obtained with the normal scale.

E. Iteration

Starting values for τ and ν are needed to iterate with (45) or (46). Two different sets of starting values were used. The first was the τ and ν roots of (45) with $W^{(i-1)} = I$, $\nu^{(i-1)} = t$ being a vector of (0,1) variables (1: difficult calving; 0: otherwise) and $\nu^{(i-1)} = 0$. These roots yielded $\tau^{(0)}$ and $\nu^{(0)}$ which were used to compute $\mu_{i(jklm)}^{(0)}$ in (57 a); in turn, these values permitted calculation of $W^{(0)}$ in both the normal and logistic cases. The second starting set was the solution to (45) with $W^{(i-1)} = I$, $\nu^{(i-1)} = t^*$ being a vector of empirical logits $\left(\ln \left[\frac{1 + .5}{0 + .5} \right] = 1.099 \right)$ if a difficult calving occurred and -1.099 otherwise) and $\nu^{(i-1)} = 0$.

Iteration stopped when $\sqrt{\Delta'/\Delta}/29 < 10^{-10}$, where $\Delta = \widehat{\theta}^{(i)} - \widehat{\theta}^{(i-1)}$. In each of the four cases resulting from the combination of normal or logistic functions with hypotheses about residual correlation ($b_1 \neq 0$ and $b_2 \neq 0$ vs. $b_1 = b_2 = 0$), convergence to the same solution occurred irrespective of the starting set used. Six rounds of iteration were required for the starting set using $\nu^{(i-1)} = t^*$; seven rounds were required when $\nu^{(i-1)} = t$ was used. From a practical point of view, however, iteration could have stopped at the third round. Results of iteration using a normal integral, $b_1 \neq 0$ and $b_2 \neq 0$, and $\nu^{(i-1)} = t$ as a trial vector are shown in Table 3.

TABLE 3

Solutions by round of iteration in the case of a normal function, $b_1 \neq 0$ and $b_2 \neq 0$, and a trial vector of (0, 1) variables.

Solution à différentes itérations dans le cas d'une fonction normale avec $b_1 \neq 0$, $b_2 \neq 0$ et un vecteur initial de variables (0, 1).

Component ^(a) of θ	Iteration ^(b)		
	0	1	7
1. $D_1 + T_2 + L_2$	41.664	41.671	41.688
2. $D_2 + T_2 + L_2$	42.307	42.282	42.315
3. $T_1 - T_2$	-1.221	-1.208	-1.194
4. $L_1 - L_2$	3.028	3.017	2.978
5. S_1	-0.336	-0.305	-0.426
6. S_2	-0.210	-0.243	-0.270
7. S_3	-0.633	-0.582	-0.664
8. S_4	0.372	0.352	0.491
9. S_5	0.730	0.821	0.885
10. S_6	0.077	-0.043	-0.017
11. $D'_1 + T'_2$	313.138	313.134	313.143
12. $D'_2 + T'_2$	312.590	312.644	312.610
13. $T'_1 - T'_2$	14.899	14.876	14.848
14. S'_1	4.401	4.349	4.548
15. S'_2	0.168	0.228	0.284
16. S'_3	-3.800	-3.907	-3.738
17. S'_4	-1.855	-1.812	-2.106
18. S'_5	11.113	10.968	10.876
19. S'_6	-10.027	-9.826	-9.864
20. $D''_1 + T''_2 + L''_2$	0.135	-1.193	-1.772
21. $D''_2 + T''_2 + L''_2$	0.107	-1.390	-2.134
22. $T''_1 - T''_2$	-0.041	0.405	0.432
23. $L''_1 - L''_2$	0.244	0.420	1.022
24. S''_1	-0.116	-0.113	-0.126
25. S''_2	-0.050	-0.053	-0.056
26. S''_3	-0.084	-0.078	-0.088
27. S''_4	0.091	0.088	0.106
28. S''_5	0.030	0.039	0.045
29. S''_6	0.129	0.116	0.119

(a) Components 1-10: birth weight, 11-19: pelvic opening, 20-29: underlying variate for calving difficulty.

(b) Convergence attained at the seventh round of iteration.

F. Model fit, estimates and their posterior precision

The models were evaluated for fit by referring the statistic in (51) to a chi-square distribution with $47 - 4 = 43$ degrees of freedom. None of the chi-square values

$$\chi^2(b_1 \neq 0, b_2 \neq 0, \text{normal}) = 26.19; \quad \chi^2(b_1 = b_2 = 0, \text{normal}) = 37.56$$

$$\chi^2(b_1 \neq 0, b_2 \neq 0, \text{logistic}) = 27.12; \quad \chi^2(b_1 = b_2 = 0, \text{logistic}) = 37.35$$

could be considered significant so there was no evidence to reject the model. However,

given the sparsity of the contingency table analyzed in this example, the approximation of (51) to a chi-square statistic may be poor.

Differences between final round estimates of θ obtained with the normal ($\hat{\theta}_N$) and the logistic ($\hat{\theta}_L$) functions were small so the latter will not be presented here. In fact,

$$[(\hat{\theta}_L - \hat{\theta}_N)/(\hat{\theta}_L - \hat{\theta}_N)/29]^{\frac{1}{2}} = 2.94 \times 10^{-2}.$$

Estimates of components of θ obtained using the normal distribution, and their estimated posterior precision (square root of estimated posterior variance) are shown in Table 4. The contrast $L_1'' - L_2''$ was estimated at 1.022 and 1.315 for the cases ($b_1 \neq 0, b_2 \neq 0$) and ($b_1 = b_2 = 0$), respectively. These indicate that if a male calf is born, the probability of a «difficult» calving would be larger than if a female calf is born,

TABLE 4

Estimates of components of θ and their posterior precision in the case of a normal function.
Estimations des composantes de θ et de leur précision a posteriori dans le cas d'une fonction normale.

Component ^(a) of θ	« Risk » relationship	
	$b_1 \neq 0, b_2 \neq 0$	$b_1 = 0, b_2 = 0$
1. $D_1 + T_2 + L_2$	41.688 ± 1.425	41.697 ± 1.424
2. $D_2 + T_2 + L_2$	42.315 ± 1.669	42.378 ± 1.669
3. $T_1 - T_2$	-1.194 ± 1.502	-1.206 ± 1.502
4. $L_1 - L_2$	2.978 ± 1.457	2.937 ± 1.455
5. S_1	-0.426 ± 0.862	-0.571 ± 0.847
6. S_2	-0.270 ± 0.867	-0.240 ± 0.854
7. S_3	-0.664 ± 0.894	-0.805 ± 0.884
8. S_4	0.491 ± 0.899	0.544 ± 0.891
9. S_5	0.885 ± 0.831	0.878 ± 0.815
10. S_6	-0.017 ± 0.852	0.194 ± 0.834
11. $D_1' + T_2'$	313.143 ± 8.468	313.287 ± 8.465
12. $D_2' + T_2'$	312.610 ± 10.710	312.362 ± 10.707
13. $T_1' - T_2'$	14.848 ± 9.970	14.756 ± 9.968
14. S_1'	4.548 ± 8.361	5.605 ± 8.306
15. S_2'	0.284 ± 8.763	0.314 ± 8.672
16. S_3'	-3.738 ± 9.125	-2.422 ± 9.059
17. S_4'	-2.106 ± 9.387	-3.268 ± 9.297
18. S_5'	10.876 ± 8.235	10.313 ± 8.146
19. S_6'	-9.864 ± 8.525	-10.543 ± 8.433
20. $D_1'' + T_2'' + L_2''$	-1.772 ± 0.563	-1.666 ± 0.634
21. $D_2'' + T_2'' + L_2''$	-2.134 ± 0.692	-1.957 ± 0.777
22. $T_1'' - T_2''$	0.432 ± 0.522	-0.150 ± 0.599
23. $L_1'' - L_2''$	1.022 ± 0.588	1.315 ± 0.643
24. S_1''	-0.126 ± 0.129	-0.361 ± 0.310
25. S_2''	-0.056 ± 0.133	-0.100 ± 0.311
26. S_3''	-0.088 ± 0.138	-0.217 ± 0.326
27. S_4''	0.106 ± 0.141	0.285 ± 0.328
28. S_5''	0.045 ± 0.124	0.016 ± 0.294
29. S_6''	0.119 ± 0.128	0.376 ± 0.302

(a) Components 1-10: birth weight; 11-19: pelvic opening; 20-29: underlying variate for calving difficulty.

irrespective of whether the effects of birth weight and pelvic opening are removed. This is consistent with the findings of BELIC & MENISSIER (1968). However, the difference in the underlying scale between male and female calves was smaller when birth weight was included as a «risk» variable. If this result were true, it would suggest that part of the difference between sexes in liability for calving difficult is not associated with differences in birth weight. The effect of including «risk» variables in the model was clear in relation to differences between seasons. Season 1 was more favourable in the ($b_1=0, b_2=0$) model perhaps because of calves with lighter birth weight and dams with larger pelvic opening; when these differences were taken into account ($b_1 \neq 0, b_2 \neq 0$), season 2 turned out to be more favourable.

G. Sire evaluation

As pointed out before, $v = u_3 - b_1 u_1 - b_2 u_2$, so sire solutions presented in Table 4 for the two different models are not comparable. Sires can be ranked for calving difficulty in the full model by using the statistic

$$\hat{u}_3 = \hat{v} + b_1 \hat{u}_1 + b_2 \hat{u}_2 \quad (63)$$

where \hat{v} , \hat{u}_1 and \hat{u}_2 are the sire components of $\hat{\theta}$ associated with the underlying variate, birth weight and pelvic opening, respectively. From a practical point of view, one may be interested in ranking sires in terms of probability of difficult calving rather than in a hypothetical underlying scale. For example, breeders may wish to know the probability that a heifer sired by the m^{th} bull, born in region 1, calving a male calf in season 1 will experience a difficult calving. An estimate of this probability can be calculated as $\hat{\Pi}_{111,m} = \Phi[\hat{D}'_1 + \hat{T}'_1 + \hat{L}'_1 + S'_m + b_1(\hat{D}'_1 + \hat{T}'_1 + \hat{L}'_1 - 43.02) + b_2(\hat{D}'_1 + \hat{T}'_1 - 320.28)]$. (64)

Using (64) for sires 1 to 6 yields

Sire 1: .253 Sire 4: .436

Sire 2: .312 Sire 5: .347

Sire 3: .304 Sire 6: .463.

In more general situations, e.g., artificial insemination, the probability of difficult calving associated with using the m^{th} sire in a given distribution of regions, calving seasons and sexes of calf may be of interest. This probability could be estimated as

$$\hat{\Pi}_{\dots m} = \sum_{jkl} \delta_{jkl} \hat{\Pi}_{jklm} \quad (65)$$

with $\hat{\Pi}_{jklm}$ as in (64) and δ_{jkl} being an arbitrary weight such that $\sum_{jkl} \delta_{jkl} = 1$. For the example considered in this paper, we took $\delta = 1/8$ because there were 8 region \times season \times sex subclasses, and ranked sires using (63) and (65). Results are shown in Table 5 for the normal and logistic distributions. As already indicated, differences between the normal and logistic models were negligible, and the estimated probability of difficult calving ranged between .116 and .239. Note that evaluations based on raw frequencies (Table 2) gave the probability rankings :

Sire 1 = Sire 3 < Sire 2 < Sire 6 < Sire 5 < Sire 4.

However, the ranking in Table 5 was

Sire 1 < Sire 3 < Sire 2 < Sire 5 < Sire 4 < Sire 6.

This indicates that evaluation based on raw frequencies can be seriously misleading. However, the progeny group sizes were small (Table 2) and none of the evaluations calculated with (63) could be considered different from zero (Table 5).

TABLE 5

Sire evaluation and estimated probability of difficulty calving^(a).
Évaluation des pères et estimation de la probabilité d'apparition d'un vêlage difficile.

Sire	Evaluation \pm Posterior precision		Probability of Difficult calving	
	Normal	Logistic	Normal	Logistic
1	-0.280 \pm 0.321	-0.279 \pm 0.321	0.117	0.116
2	-0.106 \pm 0.326	-0.103 \pm 0.326	0.147	0.148
3	-0.128 \pm 0.337	-0.127 \pm 0.336	0.143	0.144
4	0.225 \pm 0.341	0.232 \pm 0.340	0.217	0.226
5	-0.009 \pm 0.311	-0.007 \pm 0.310	0.166	0.169
6	0.298 \pm 0.319	0.284 \pm 0.318	0.235	0.239

(a) See equations (63) and (65).

VII. Conclusions

This paper presents a solution to the problem of estimating the genetic merit of candidates for selection when both quantal and continuous information is available in a set of individuals. The proposed method was adapted to the situation where the probability of «response» is a function of continuous «risk» variables. Also, consideration is given to the assumption that candidates for selection are sampled from a distribution with second moments known, a priori. The method can be extended to multiple ordered or unordered categories of response along the lines presented by GIANOLA & FOULLEY (1983).

The method is non-linear and approximates the best predictor in a squared error sense. Theoretical objections arising in analysis of categorical data with linear models (e.g., GIANOLA, 1982) are eliminated. For example, when calving difficulty is measured as an «all or none» trait, sire \times sex of calf interactions are usually found to be «significant». This may be associated with a scaling problem. Suppose we wish to compare two sires and that the values in the underlying scales are μ_{1M} , μ_{1F} , μ_{2M} and μ_{2F} ; the subscripts indicate the sire and the sex of the calf. Further, suppose that there is no interaction between sex and sire in the underlying scale, i.e.,

$$\mu_{1M} - \mu_{2M} = \mu_{1F} - \mu_{2F}.$$

However, $\Phi(\mu_{1M}) - \Phi(\mu_{2M})$ may be different from $\Phi(\mu_{1F}) - \Phi(\mu_{2F})$ because $\Phi(x)$ does not vary linearly with x .

The method of estimation is based on Bayes theorem, but is not completely Bayesian in the sense that the variance-covariance structure is regarded as representing a set of «nuisance» parameters. In principle, prior knowledge (or lack of) about variances and covariances could be represented via a prior distribution (LINDLEY & SMITH, 1972) and modal estimates obtained from the posterior density. HARVILLE (1977) has indicated that estimators of variances obtained from the joint posterior mode can be degenerate if uninformative priors are used. This author qualified the modes of the marginal posterior density of the variance components as «seemingly superior» estimators.

Important numerical problems arise when the procedure is applied to the estimation of vectors with thousands of elements, the usual situation in applied animal breeding. Nevertheless, the order of the computations is comparable to that arising in multi-dimensional BLUP multiplied by the number of « main » iterates needed to achieve convergence. When the « risk » variables are considered in the model, the method requires that every experimental unit with a categorical response includes information on the quantitative variates.

Acknowledgements

Daniel GIANOLA wishes to acknowledge I.N.R.A., France, for support during his stay at Jouy-en-Josas, and the Holstein Association, Brattleboro, Vermont, U.S.A., for supporting his work in categorical data. Dr. Stephen P. SMITH, Cornell University, U.S.A., is thanked for useful comments.

Received February 2, 1983.

Accepted April 29, 1983.

References

- BECKMAN F.S., 1960. *The solution of linear equations by the conjugate-gradient method*. In : Ralston A., Wilf H.S., *Mathematical methods for digital computers*, Wiley, New York.
- BELIC M., MENISSIER F., 1968. Étude de quelques facteurs influençant les difficultés du vêlage en croisement industriel. *Ann. Zootech.*, **17**, 107-142.
- CORNFIELD J., 1969. The Bayesian outlook and its application. *Biometrics*, **25**, 617-657.
- COX D.R., 1970. *The Analysis of Binary Data*. Chapman and Hall, London.
- DAHLQUIST G., BJÖRCK A., 1974. *Numerical Methods*. Prentice Hall, Englewood Cliffs.
- DEMPFLE L., 1977. Relation entre BLUP (Best Linear Unbiased Prediction) et estimateurs Bayésiens. *Ann. Genet. Sel. Anim.*, **9**, 27-32.
- DEMPSTER E.R., LERNER I.M., 1950. Heritability of threshold characters. *Genetics*, **35**, 212-235.
- FINNEY D.J., 1952. *Probit analysis : a statistical treatment of the sigmoid response curve*. 2nd ed., University Press, Cambridge.
- GIANOLA D., 1980. A method of sire evaluation for dichotomies. *J. Anim. Sci.*, **51**, 1266-1271.
- GIANOLA D., 1982. Theory and analysis of threshold characters. *J. Anim. Sci.*, **56**, 1079-1096.
- GIANOLA D., FOULLEY J.L., 1983. Sire evaluation for ordered categorical data with a threshold model. *Génét. Sél. Evol.*, **15**, 201-223.
- HARVILLE D.A., 1977. Maximum likelihood approaches to variance component estimation and to related problems. *J. Am. Stat. Assoc.*, **72**, 320-338.
- HENDERSON C.R., 1973. Sire evaluation and genetic trends. *Proceedings of the Animal Breeding and Genetics Symposium in honor of Dr Jay L. Lush, Blacksburg, Virginia, July 29, 1972*, 10-41, ASAS-ADSA, Champaign, Illinois.
- HENDERSON C.R., 1980. A simple method for unbiased estimation of variance components in the mixed model. *J. Anim. Sci.*, **51** (Suppl. 1), 119.
- HENDERSON C.R., KEMPTHORNE O., SEARLE S.R., VON KROSIGK C.M., 1959. The estimation of environmental and genetic trends from records subject to culling. *Biometrics*, **15**, 192-218.
- HILL B.M., 1965. Inference about variance components in the one-way model. *J. Am. Stat. Assoc.*, **60**, 806-825.
- LEONARD T., 1972. Bayesian methods for binomial data. *Biometrika*, **59**, 581-589.
- LINDLEY D.V., SMITH A.F.M., 1972. Bayes estimates for the Linear model. *J.R. Stat. Soc. B.*, **24**, 1-41.
- MOOD A.M., GRAYBILL F.A., 1963. *Introduction to the Theory of Statistics*. McGraw-Hill Book Co., New York.

- RÖNNINGEN K., 1971. Some properties of the selection index derived by «Henderson's Mixed Model Method». *Z. Tierz. Züchtbiol.*, **88**, 186-193.
- SCHAEFFER L.R., WILTON J.W., THOMPSON R., 1978. Simultaneous estimation of variance and covariance components from multitrait mixed model equations. *Biometrics*, **34**, 199-208.
- THOMPSON R., 1979. Sire evaluation. *Biometrics*, **35**, 339-353.
- TIAO G.C., TAN W.Y., 1965. Bayesian analysis of random effects models in the analysis of variance. I. Posterior distribution of variance components. *Biometrika*, **52**, 37-53.
- TIAO G.C., BOX G.E.P., 1967. Bayesian analysis of a three-component hierarchical design model. *Biometrika*, **54**, 109-125.
- VAN NORTON R., 1960. The solution of linear equations by the Gauss-Seidel Method. In: Ralston A. Wilf H.S., *Mathematical methods for digital computers*, Wiley, New York.