

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Papers and Publications in Animal
Science

Animal Science Department

January 1992

Multiple-Trait Restricted Maximum Likelihood for Simulated Measures of Ovulation Rate with Underlying Multivariate Normal Distributions

L. Dale Van Vleck

University of Nebraska-Lincoln, dvan-vleck1@unl.edu

K. E. Gregory

Roman L. Hruska U.S. Meat Animal Research Center

Follow this and additional works at: <https://digitalcommons.unl.edu/animalscifacpub>



Part of the [Animal Sciences Commons](#)

Van Vleck, L. Dale and Gregory, K. E., "Multiple-Trait Restricted Maximum Likelihood for Simulated Measures of Ovulation Rate with Underlying Multivariate Normal Distributions" (1992). *Faculty Papers and Publications in Animal Science*. 247.

<https://digitalcommons.unl.edu/animalscifacpub/247>

This Article is brought to you for free and open access by the Animal Science Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Papers and Publications in Animal Science by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Multiple-Trait Restricted Maximum Likelihood for Simulated Measures of Ovulation Rate with Underlying Multivariate Normal Distributions¹

L. D. Van Vleck* and K. E. Gregory†

Roman L. Hruska U.S. Meat Animal Research Center,
U.S. Department of Agriculture, ARS,

*Lincoln, NE 68583-0908 and †Clay Center, NE 68933

ABSTRACT: A data set that was used to estimate covariance components with REML for an animal model with eight measures of ovulation rate treated as separate traits was used as a template to simulate data sets of eight multivariate normal traits that were then truncated to binomial traits. The model for simulation included eight measures on 610 animals with 1,071 animals in the numerator relationship matrix. Heritabilities were equal for the eight measures, and both genetic and phenotypic correlations among the measures were equal. Ten replications for each combination of heritability (.15, .25, and .35) and genetic correlation (.50, .66, and .90) were simulated on the normal scale. For each replicate, estimates of the eight heritabilities and 28 genetic correlations were obtained by multiple-trait

REML. The usual transformation of heritability estimated on the binomial scale overestimated heritability on the normal scale. Genetic correlations on the binomial scale seriously underestimated the correlations on the normal scale. Standard errors of the estimates obtained by replication were somewhat larger than the approximate SE from REMLPK (the multitrait REML program of K. Meyer). A final set of 10 simulated replications with heritability of .25 and genetic correlation of 1.00 resulted in average estimates of .18 for heritability and of .66 for genetic correlation that agree closely with those from the analysis of measures of ovulation at eight estrous cycles used as a template; averages for heritability of .16 and for genetic correlation of .66 were obtained.

Key Words: Ovulation Rate, Monte Carlo Method, Variance Estimation

J. Anim. Sci. 1992. 70:57-61

Introduction

Analysis (Van Vleck et al., 1991) of measures of number of ovulations at each of eight consecutive estrous cycles of 610 puberal heifers in the Twinning Project of the U.S. Meat Animal Research Center resulted in estimates of heritability that averaged .16 and ranged from .09 to .22 for each cycle treated as a separate trait. The 28 genetic correlations among the eight cycles (traits) averaged .66, with a range of .08 to .96. The phenotypic correlation averaged .12, with a range

of -.05 to .24. The analysis was with a multiple-trait animal model using REML. The data included 610 heifers with ovulation rates and 461 ancestors that contributed to relationships among the heifers and their parents. The model included seven birth groups (year and calving season of birth), six calendar months of first measurement, and three ages in months at first measurement.

Restricted maximum likelihood is based on a multivariate normal distribution for the data. Ovulation rates, however, are measured as single or double ovulations (i.e., follow a joint binomial distribution). Robertson and Lerner (1949) have provided a transformation to estimate heritability on an underlying normal scale from an estimate on the binomial scale. Such transformations are relatively good for paternal sib analyses (Van Vleck, 1972; Ollausson and Ronningen, 1975), but they are somewhat less good for analyses based

¹Published as paper no. 9507, Journal Ser., Nebraska Agric. Res. Div., Univ. of Nebraska, Lincoln 68583-0908.

Received February 25, 1991.

Accepted July 23, 1991.

on progeny on parent regression provided that the underlying distribution is normal (see review by McGuirk, 1989). Ollausson and Ronningen (1975) have reasoned that genetic correlations would be the same on both scales, although the phenotypic correlations should be transformed by essentially the same method as heritabilities. Those simulation results and later ones shown by McGuirk (1989, B. J. McGuirk and R. Thompson, unpublished data) were based on half-sib and parent-progeny analyses.

The purposes of doing the simulations reported here were to examine for a multiple-trait animal model as follows: 1) estimates of heritability and genetic and phenotypic correlations from multivariate binomial data with an underlying known multivariate normal distribution and 2) empirical estimates of sampling variances to compare with those approximated by the multiple trait REML program (REMLPK) of K. Meyer (1985, 1986, and personal communication).

The scope of the project was limited to the design matrix of the heifer ovulation data and a few combinations of underlying parameters. Of particular interest was why from actual data the genetic correlations among estrous cycles from 3 to 21 wk apart were only .66, whereas intuitively a genetic correlation of near unity might be expected.

Materials and Methods

The incidence matrices and numerator relationship matrix corresponding to the analysis of number of ovulations for eight consecutive estrous cycles described in the introduction (Van Vleck et al., 1991) were used as the template for the simulation. Levels of the three fixed factors were assigned to each record but effects were not added (i.e., were set equal to zero) and were simply "noise" because REML essentially works on the likelihood after fixed effects are eliminated. The inverse of the numerator relationship matrix, A^{-1} , for the 1,071 animals of which 610 had records was computed using the NRMTRI program of K. Meyer (personal communication) according to the procedure developed by Quaas (1976), which accounts for inbreeding. The inverse of A^{-1} was obtained to provide A . Included in the 451 animals without records were 40 sires. Non-zero elements of the A and A^{-1} matrices were 13,959 (average of 12 in each off-diagonal row) and 3,510 (average of 3.3 per off-diagonal row), respectively. Additive genetic values for the 1,071 animals for the eight traits (each estrous cycle corresponded to a different trait) were generated (Van Vleck, 1991) as $\mathbf{a} = \mathbf{L}_A * \mathbf{L}_G \mathbf{v}$, where \mathbf{a} is the vector of additive genetic values ordered by trait within animal

(length $8 \times 1,071$), \mathbf{L}_A is the lower Cholesky factor of A such that $\mathbf{L}_A \mathbf{L}_A^T = A$, $*$ indicates the direct product operator, \mathbf{L}_G is the lower Cholesky factor of G , G is the genetic variance-covariance matrix for the eight traits, and \mathbf{v} is a vector of pseudorandom, standard normal deviates $N(0,1)$ of length $8 \times 1,071$.

As described by Van Vleck (1991), \mathbf{a} can be accumulated one animal at a time from subvectors of \mathbf{v} of order 8 premultiplied by \mathbf{L}_G and then accumulated as multiples of elements in the column of \mathbf{L}_A corresponding to the animal.

Environmental effects were simulated for the eight traits for each of the 610 animals with records as $\mathbf{e}_i = \mathbf{L}_E \mathbf{v}_i$, where \mathbf{L}_E is the lower Cholesky factor of E , the variance-covariance matrix of the eight environmental effects on the same animal, and \mathbf{v}_i is a vector of pseudorandom, standard normal deviates of length eight.

For the multivariate normal simulation, E was chosen such that $G + E = P$ for given G and P where the diagonals of P were ones and off-diagonals were the phenotypic correlation.

For animals with records, the corresponding elements of \mathbf{a}_i and \mathbf{e}_i were added together to obtain the simulated multivariate normal records for animal i . Each record was compared to the standard normal truncation point, 1.0365, for which 85% of the records would be below and 15% above that point. A 1 corresponding to a single ovulation was assigned to records below the truncation point, and a 2 was assigned to records above the truncation point; thus, the mean on the binomial scale was 1.15.

Nine combinations of parameters were used in an attempt to bracket the solutions from the original data. Combinations were for heritability of .15, .25, and .35 and genetic correlations of .50, .66, and .90. The correlations among the eight traits were equal. The intent was for a phenotypic correlation of .20, which was implemented for six combinations with heritability of .15 and .25. When heritability was .35, a phenotypic correlation of .20 and a genetic correlation of .90 made the variance-covariance matrix of the eight environmental effects negative definite. Thus, the phenotypic correlation was set to .25 for all combinations with heritability of .35.

Each combination was replicated 10 times. Data simulation was rapid but estimation of the covariance matrices for each replicate required approximately 6 h of computer time for the 5,000 rounds of iteration using the REMLPK program of K. Meyer (personal communication) modified to use the Henderson (1984) quadratics, which result in slower convergence. The other quadratics in previous analyses often resulted in failure to converge (Van Vleck et al., 1991).

Table 1. Summary of estimates for 10 replications of simulations of binomial data for 10 combinations of heritability and genetic correlations for underlying multiple-trait normal distributions

Heritability: Genetic correlation:	Parameters on normal scale									
	.15 ^a			.25 ^a			.35 ^b			.25 ^b
	.50	.66	.90	.50	.66	.90	.50	.66	.90	1.00
Estimates from binomial scale										
	Heritabilities ^c									
Mean	.11	.12	.11	.16	.13	.14	.18	.20	.21	.18
Maximum	.33	.37	.34	.46	.32	.32	.47	.60	.37	.51
Minimum	.01	.02	.02	.03	.02	.02	.04	.03	.06	.04
SD	.06	.06	.06	.08	.06	.06	.09	.10	.07	.07
	Genetic correlations ^d									
Mean	.27	.36	.43	.36	.51	.58	.36	.51	.65	.66
Maximum	.97	.97	.96	.94	.99	.98	.95	.96	.98	.96
Minimum	-.75	-.62	-.72	-.80	-.67	-.43	-.53	-.77	-.29	-.30
SD	.38	.37	.36	.33	.36	.28	.30	.31	.22	.23
	Phenotypic correlations ^d									
Mean	.09	.10	.09	.10	.09	.09	.11	.12	.12	.12
Maximum	.22	.24	.22	.28	.27	.28	.23	.26	.26	.26
Minimum	-.03	-.02	-.05	-.02	-.01	-.04	-.03	-.01	.00	-.02
SD	.05	.05	.05	.04	.05	.05	.05	.05	.05	.05

^aOn normal scale; phenotypic correlation = .20.^bOn normal scale; phenotypic correlation = .25.^c80 estimates from 10 replicates of eight traits.^d280 estimates from 10 replicates of 28 correlations among eight traits.

Results from the first nine combinations of parameters suggested simulation of a 10th unusual parameter set. In that set, the genetic correlations among the eight traits on the normal scale were set to unity. The simulation procedure was modified to produce only one genetic value per animal, which was then in all eight of its records. Heritability of .25 was chosen, which does not allow a phenotypic correlation of .20 but does allow .25 (environmental correlation of zero), which was used.

Results and Discussion

The results for the 10 replicates and 10 combinations of underlying parameters are summarized in Table 1.

Heritabilities. The average heritabilities for all combinations of parameters are much greater than would have been predicted by the usually cited transformation from normal to binomial data (Robertson and Lerner, 1949), especially with underlying heritability of .15. This overestimation is even greater than anticipated from previous simulations in which transformation of half-sib estimates performed better than transformations of regressions of progeny on parent (McGuirk, 1989). Mäntysaari et al. (1991), using a simulated sire model (no relationships) with REML, found close agreement between heritability transformed

from binomial scale and heritability on the normal scale for two binomially distributed traits derived from a bivariate normal distribution. The overestimates from multiplying the binomial estimate by $(1 - p)/(i^2 p) = 2.35$ are substantial (30 to 70% depending on heritability and genetic correlation on the normal scale) for this situation: animal model and binomial threshold frequency of 15%.

Genetic Correlations. In contrast to the expectation of no difference in genetic correlations on the binomial and normal scales summarized in the review of McGuirk (1989) for half-sib analyses, with the animal model the genetic correlations are markedly reduced on the binomial scale compared with the normal scale. The genetic correlations from analyses on the binomial scale averaged from 48 to 76% of the parameters on the normal scale. The reduction was greatest for the smallest heritability simulated. Within a heritability, the reduction was generally greatest for the largest genetic correlation, .90. Mäntysaari et al. (1991), using a sire model with two binary variables, found reductions, although smaller than these, in the estimate of genetic correlation for incidences of .15 and .05 but no reduction for incidence of .25.

Phenotypic Correlations. The average phenotypic correlations on the binomial scale agree with those predicted with the Ollausson and Ronning (1975) and Vinson et al. (1976) transformation, which is the same as the Robertson and Lerner

(1949) transformation for heritability when the means are the same for both traits. Mäntysaari et al. (1991), however, found that the transformation with a sire model overestimated the underlying residual correlation, particularly for binomial traits with low incidence (.05 and .15).

Simulation with Genetic Correlations of Unity. From the view of the ovulation rate analysis that was the stimulus for this study, the simulation results summarized in the last column of Table 1 are of the greatest interest. Before the first simulations with correlations of .50, .66, and .90, there was a fear that using the same genetic effect for all eight measures on the underlying scale would result in singular matrices of estimated genetic (co)variances. After the first simulations showed a reduction of the genetic correlation on the binomial scale, simulations with genetic correlations of unity on the normal scale were also done. An underlying heritability of .25 was chosen, based on the previous simulations, to try to match the .16 found for the ovulation rate analysis (Van Vleck et al., 1991). The average estimates for the 10 replicates are in close agreement with those from the ovulation analysis: heritability, .18 vs .16; genetic correlation, .66 vs .66; and phenotypic correlation, .12 vs .12.

The simulation results suggest that in the population studied, if ovulation rate measured as a binomial variable has an underlying normal distribution, then the underlying genetic correlation is very large (.90 to 1.00) among measures of ovulation rate at various estrous cycles of heifers. The empirical estimates of SD suggest that the SE for the mean of 28 estimates is approximately $.23/\sqrt{28} = .043$. That SE about the binomial scale average of .66 for the genetic correlation suggests that an underlying genetic correlation would be .90 or larger. The SE of the average of eight heritability estimates would be approximately $.07/\sqrt{8} = .025$. That SE combined with a mean of .16 from the ovulation rate data suggests that heritability on the binomial scale is in the range of .11 to .21 and that if the underlying scale is normally distributed heritability on the normal scale would be approximately .25.

Standard Errors. The empirical SD do not vary much by parameter combination, although the SD of estimates of genetic correlations are somewhat smaller for the largest genetic correlation and the SD of estimates of heritability are slightly larger for larger heritability. These empirical SE are generally slightly larger than those provided by the algorithm in REMLPK. For the ovulation rate analysis that was the template for the simulation, the reported SE of individual SE over the eight

heritabilities and 28 correlations were as follows: .03 to .08 for heritability; .03 to .30 for genetic correlations; and .02 to .04 for phenotypic correlations. The REMLPK procedure seems to provide somewhat smaller SE than those found by this simulation.

Conclusions

The conclusions are within the limited scope of this simulation of binomial variables with an underlying multivariate normal distribution. The usual transformation of heritability on the binomial to normal scale results in larger overestimation than previously reported for half-sib and progeny-parent models. In contrast to previous studies with half-sib models, the animal model yields estimates of genetic correlations that are considerably reduced on the binomial scale from the underlying normal scale.

As suggested long ago (e.g., C. R. Henderson, personal communication, 1962), simulation can be used to establish approximate SE for estimates from methods that do not have exact computing forms for SE. For this design and binomial variables, the algorithm in K. Meyer's REMLPK program was shown to slightly underestimate the SE found from simulation. Computing time to simulate data to estimate standard errors might still be a limiting factor for REML with an animal model when the number of equations becomes large. In this simulation with equal heritabilities and genetic and phenotypic correlations and less than 1,100 equations after canonical transformation, each replication took approximately 6 h on a microcomputer with an 80386-20 processor and a math coprocessor with all of the efficiencies of REMLPK (linked list storage, canonical transformation, transformation of the relationship matrix to an identity [Quaas, 1989], and tridiagonalization). Therefore, the number of replicates and parameter combinations was limited and would be limited, although less so, on large mainframe computers. Nevertheless, for specific designs and a limited range of parameters, simulation provides a way to compute empirical SE or to check the accuracy of approximate methods.

A conclusion specific to the template for this simulation is that if the underlying distribution is multivariate normal, then the genetic correlation among measures of ovulation rate for a series of estrous cycles is likely to be near unity rather than the .66 estimated from measures on the binomial scale. Similarly, heritability is likely to be considerably greater than the .16 calculated on the binomial scale, perhaps as large as .30 to .35.

Implications

The usual transformation of heritabilities estimated on a binomial scale will overestimate heritability on an underlying normal scale when the data are analyzed with an animal model. This result suggests less optimism than in the past for the amount of progress to be expected from selection on traits of low heritability measured on a yes or no basis (e.g., single vs twin ovulations). On the other hand, measurement on the binomial scale seems to reduce the apparent genetic correlation for traits that have a correlation of near unity on the underlying normal scale. Efficiency of selection based on means of such a correlated trait would be greater than that expected based on the genetic correlation associated with the binomial measurements.

Literature Cited

- Henderson, C. R. 1984. Applications of Linear Models in Animal Breeding. University of Guelph, Ontario, Canada.
- Mäntysaari, E. A., R. L. Quaas, and Y. T. Gröhn. 1991. Simulation study on covariance component estimation for two binary traits in an underlying continuous scale. *J. Dairy Sci.* 74:580.
- McGuirk, B. J. 1989. The estimation of genetic parameters for all-or-none and categorical traits. In: W. G. Hill and T.F.C. Mackay (Ed.) *Evolution and Animal Breeding. Reviews on Molecular and Quantitative Approaches in Honour of Alan Robertson.* p 175. C.A.B. International, Wallingford, UK.
- Meyer, K. 1985. Maximum likelihood estimation of variance components for a multivariate mixed model with equal design matrices. *Biometrics* 41:153.
- Meyer, K. 1986. Restricted maximum likelihood to estimate genetic parameters - in practice. *Proc. 3rd World Cong. Genet. Applied Livest. Prod. Lincoln, NE. Vol. XII:454.*
- Ollausson, A., and K. Ronningen. 1975. Estimation of genetic parameters for threshold characters. *Acta Agric. Scand.* 25:201.
- Quaas, R. L. 1978. Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics* 32:949.
- Quaas, R. L. 1989. Transformed mixed model equations: A recursive algorithm to eliminate A^{-1} . *J. Dairy Sci.* 72:1937.
- Robertson, A., and I. M. Lerner. 1949. The heritability of all-or-none traits: viability in poultry. *Genetics* 34:395.
- Van Vleck, L. D. 1972. Estimation of heritability of threshold characters. *J. Dairy Sci.* 55:218.
- Van Vleck, L. D. 1991. Simulation of animal models with multiple traits and non-additive genetic effects. *J. Anim. Sci.* 69(Suppl. 1):74 (Abstr.).
- Van Vleck, L. D., K. E. Gregory, and S. E. Echternkamp. 1991. Ovulation rate and twinning rate in cattle: heritabilities and genetic correlation. *J. Anim. Sci.* 69:3213.
- Vinson, W. E., J. M. White, and R. H. Kliever. 1976. Overall classification as a selection criterion for improving categorically scored components of type in Holsteins. *J. Dairy Sci.* 59:2104.