

A NOTE ON THE USE OF CONDITIONAL MODELS
TO ESTIMATE ADDITIVE GENETIC VARIANCE IN SELECTED POPULATIONS

Julius H.J. van der Werf
Dept of Animal Breeding, Wageningen Agricultural University
P.O. Box 338, 6700 AH Wageningen, The Netherlands

SUMMARY

A conditional model to estimate genetic parameters in data from a selected base population was investigated. The method assumes base parents as fixed and a conditional variance is based upon the Mendelian sampling of gametes from the base parents. In a simulation study, only sires were selected and mated to all females to create the next generation (2 progeny per female). Selection was for five generations but only animals of generations 4 and 5 were assumed to have performance records and parents known. Due to selection, initial genetic variance was underestimated when base animals from generation 3 were assumed random. However, when all base animals were assumed fixed, underestimation of genetic variance was even larger whereas the residual variance was overestimated. When males of generation 4 were not selected to have progeny, estimated variances were empirically unbiased. It was concluded that estimates for genetic parameters with the conditional model were not biased by selection of base animals, but a new bias was introduced when descendants of fixed base animals were selected to have progeny.

INTRODUCTION

A mixed model accounts for the change of variance when additive genetic relationships are known, tying animals back to a certain base population that consists of unrelated, unselected and noninbred animals (Sorensen and Kennedy, 1984). However, data used for estimation of genetic parameters arise generally from recording during a limited time period and a group of animals with unknown parents is treated as the base population. Base population animals may then be selected and estimates of genetic variance, may be biased by selection of ancestors (Van der Werf and De Boer, 1990).

Procedures were proposed to estimate additive genetic values (Henderson, 1985, 1988) or variance (Graser et al., 1987) for the case that base animals were selected. In essence it was suggested to treat base animals as fixed and to estimate genetic variance independent from the variance among selected base animals. The method uses the fact that variance from Mendelian Sampling is assumed to be not affected by prior selection (Bulmer, 1971).

Statistical and genetic properties of the conditional model have not been well tested. The aim of this paper is to show some results of testing the conditional model by simulation.

MATERIAL AND METHODS

A Monte Carlo simulation study was carried out as described by Van der Werf and De Boer (1990), considering a normally distributed trait regulated by a large number of loci. A base generation (generation 0) of 20 males and 200 females, all assumed unrelated and unselected, was mated at random to produce

200 males and 200 females. In each of the subsequent generations, the 20 phenotypically best males were selected and each of them was mated to 10 females. Animals were only mated within generation and the only fixed effect simulated was the mean. Two random effects were simulated; an additive genetic and a residual effect, both distributed as $N(0,10)$. Additive genetic values for animals not belonging to the base generation were

$$a_i = \frac{1}{2}a_{s_i} + \frac{1}{2}a_{d_i} + \phi_i \quad [1]$$

where s_i is the sire, d_i is the dam, and ϕ_i can be seen as a random variable representing Mendelian sampling, with $\text{var}(\phi_i) = \frac{1}{4}[1 - \frac{1}{2}(F_s + F_d)]\sigma_a^2$, and F_s and F_d being inbreeding coefficients of sire and dam, respectively. The total number of generations simulated was 10.

An animal model was used for data analysis and a derivative-free algorithm supplied Restricted Maximum Likelihood estimates of the variance components (Graser et al., 1987; Meyer, 1989). An analysis with all breeding values assumed random was compared with an analysis treating genetic effects of the base animals as fixed effects, as suggested by Graser et al. (1987).

RESULTS

After five generations of selection, additive genetic variance decreased due to inbreeding, gametic disequilibrium and covariance among animals. Mean additive genetic values in generations 3, 4 and 5 were 3.72, 5.47 and 7.14, respectively, and average inbreeding for generations 1 to 3 was .013, .020, and .027, respectively. Additive genetic variance averaged over replicates was 8.55, 8.19, and 8.12, in generation 3, 4 and 5, respectively.

Estimates of genetic variance are in Table 1 for the case that only records from generations 4 and 5 are used and generation 3 is assumed to be the base population. Estimates were expected to be affected by the average inbreeding coefficient of the base animals and by selection of base animals, since a random model assumes non-inbred and unselected base animals. Average inbreeding coefficient at generation 3 was .012. Genetic variance among animals of generation 3 was 8.53. The estimate from a model with base animals random was 8.58, and estimating genetic variance treating only sires as fixed resulted in 9.14. The estimate from a model treating both base-sires and -dams as fixed was much lower; 6.03, and the estimate of residual variance were higher than the simulated value of 10.

TABLE 1. Estimated genetic variance ($\hat{\sigma}_a^2$) using simulated data from selected generations of different populations and using a model treating base animals as random or as fixed ^{a b}.

base generation	data (gen)	n	$\hat{\sigma}_a^2$	$\hat{\sigma}_e^2$
3 random	4-5	20	8.58 (1.86) ^c	10.44
3 sires fixed	4-5	20	9.14 (2.26)	10.19
3 sires and dams fixed	4-5	18	6.03 (2.14)	11.77

^a initial genetic and residual variance was 10

^b twenty out of 200 males were selected each generation

^c empirical standard deviations from n replicates

To investigate further the large bias observed, a data set with records only from the first two generations was analysed. Note that base animals from generation 0 were noninbred and unselected. Results are shown in Table 2, indicating that estimates for genetic variance were also biased with a conditional model using data from generations 1 and 2. Apparently, having a selected base population was not the cause of bias from the conditional model.

We also analysed data for the case of no selection of parents to have progeny. Results from no selection, however, were not biased, indicating that selection of animals with records had caused biased estimates. This was even more clearly demonstrated in a data set, in which animals with records were randomly chosen to have progeny, but those from previous generations had been selected. Analyzing records from generations 4 and 5, and selecting parents, except in generation 4, gave an unbiased estimate of genetic and residual variance (Table 2). Hence, estimates obtained with the conditional model were not biased by selection of base animals, but the procedure yields biased estimates when descendants of fixed base animals were selected to have progeny. Notice that all data on which these selection decisions were based were included in the model.

TABLE 2. Estimated genetic variance ($\hat{\sigma}_a^2$) using simulated data from selected generations with varying heritability and a model treating base animals as fixed ^{a b}.

base gen.	data (gen)	$\hat{\sigma}_a^2$	$\hat{\sigma}_e^2$
<i>selection</i>			
0 random	1-2	9.97 (1.80) ^c	10.17
0 fixed	1-2	6.51 (3.00)	11.79
3 random	4-5	8.58 (1.87)	10.44
3 fixed	4-5	6.03 (2.14)	11.77
<i>no selection</i>			
0 random	1-2	9.92 (1.76)	10.35
0 fixed	1-2	9.95 (4.16)	10.26
3 random	4-5	9.79 (2.18)	10.01
3 fixed	4-5	10.08 (3.81)	9.83
<i>selection, except in generation 4</i>			
3 fixed	4-5	9.91 (3.77)	9.93

^a initial genetic variance was 10

^b twenty out of 200 males were selected each generation

^c empirical standard deviations from n replicates

DISCUSSION

The conditional model assuming base animals fixed is an easy method to obtain estimates of genetic variance that are not biased by selection on information recorded prior to the data set. Estimates of effects of base dams, when fixed and nested within sires, can be seen as estimates of family means. The variance of the random genetic component is then derived from deviations from family means, and can only be estimated when data is available on at least two generations. In a hierarchical structure with two sexes known, the variances are estimated from contrasts between families, and contrasts within families. If base animals were random, information of unselected dams about the selected dam mean could also be used to account for this selection (Thompson, 1973; 1976). When base animals are fixed, some of this information is not used. Curnow (1961) showed how the variance of offspring records (y_p) conditional to parental records (y_b), is estimated using $z_p = y_p - [\text{cov}(y_b, y_p) / \text{var}(y_b)] y_b$. When base animals are taken as fixed, the deviations considered are $z_p = (y_p - \bar{y}_b)$ and therefore parent offspring regressions may not always be appropriate.

The conclusion that selection causes biased estimates, although all records on which selection was based were used, seems contradictory to previous findings (e.g. Henderson, 1975). However, selection was accounted for using the appropriate covariance structure among random variables. This property apparently does not hold when some variables are treated as fixed. Selection on records not in the data is often accommodated by including groups in the model. In fact, the 'Westell grouping' strategy (Quaas, 1988) with one base animal per group is equivalent to a model with fixed base parents. It is not clear yet whether such models could lead to bias in some cases of prediction of breeding values.

ACKNOWLEDGEMENTS

The author is very grateful to Robin Thompson for valuable suggestions, and to Karin Meyer for providing her DFREML programs.

REFERENCES

- BULMER, M.G. 1971. *Ameri. Nat.* 105:201.
CURNOW, R.N. 1961. *Biometrics* 17:553.
GRASER, H.U., S.P. SMITH and B. TIER. 1987. *J. Anim. Sci.* 64:1362.
HENDERSON, C.R. 1975. *Biometrics* 31:423.
HENDERSON, C.R. 1985. *J. Dairy Sci.* 68:443.
HENDERSON, C.R. 1988. *J. Dairy Sci.* 71:3399.
MEYER, K. 1986. *Proc. 3rd World Congr. Genet. Appl. Livest. Prod.* 12:454.
MEYER, K. 1989. Submitted to *Genet. Sél. Evol.*
QUAAS, R.L. 1976. *Biometrics* 32:949.
QUAAS, R.L. 1984. *BLUP School Handbook, Anim. Genet. Breed. Unit, Univ. of New England, N.S.W. 2351, Australia.* p. 21-23.
QUAAS, R.L. 1988. *J. Dairy Sci.* 71:1338.
SORENSEN, D.A., and B.W. KENNEDY. 1984. *J. Anim. Sci.* 59:1213.
THOMPSON, R. 1973. *Biometrics* 29: 527.
THOMPSON, R. 1976. *Biometrics* 32:283.
VAN DER WERF, J.H.J., and I.J.M. DE BOER. 1990. *J. Anim. Sci.* (submitted).