DETECTION OF GENOTYPE BY ENVIRONMENT INTERACTION USING MERINO SIRE EVALUATION DATA

J.W.James and D.J.Cottle

Department of Wool and Animal Science, The University of New South Wales, Sydney, NSW 2052

SUMMARY

A method is presented for determining the power of a test for genotype by environment interaction using data on progeny tests in different environments when some sires are tested in more than one environment. The method is applied to Merino sire evaluation and shows that the available data are sufficient to justify use of the data to test for such an interaction.

Keywords: Genotype-environment interaction, sire evaluation, Merino.

INTRODUCTION

The great value of sire evaluation progeny tests using link rams is that rams tested in different trials can be indirectly compared through comparison with the rams linking the various trials. However, many breeders are concerned that such comparisons may be biased by the occurrence of genotype by environment interaction (GEI). These interactions are known to be important for some traits such as fleece rot, but are generally regarded as less important for traits such as wool weight and fibre diameter, though reliable estimates are scarce (Woolaston 1987). Studies on genotype by environment interactions in Merinos have been relatively infrequent, and further analyses are desirable. The availability of data on link rams which have been tested in two or more locations or years, including both central test station and on-farm trials, provides an opportunity for further studies. This opportunity is a byproduct of the sire evaluation program and only a small portion of the total data is relevant to testing for genotype by environment interaction. In assessing the value of such an analysis it is important to calculate the power of the statistical test, that is, the probability of obtaining a statistically significant result if GEI is in fact present. If the power is low, the absence of a significant effect of GEI may be attributable either to the poor discrimination of the data or to the negligible influence of GEI.

In this paper two main issues are examined. With the size and structure of the available data set, what is the chance of detecting an interaction if one exists? How would the power be increased if some previously tested sires were compared in a special trial? A subsidiary question is : How does the answer to the second question depend on the choice of sires for the special trial?

METHODS

An ideal data set with which to test the importance of GEI involving rams and trials would be completely balanced. There would be s sires, each having an equal number n of progeny recorded in each of the t trials. The data from such an arrangement would be analysed to give an analysis of variance in the form shown in Table 1, in which V_R is the variance among progeny of a given sire in a particular trial, while V_I is the variance of sire by trial interaction effects.

The significance of GEI effects would be tested by F = I/R, which would under the null hypothesis of no GEI effects, be $F[f_I,f_R]$, a standard F ratio with f_I and f_R degrees of freedom. If the null hypothesis is false, on the other hand, the ratio F would be distributed as (1 + nQ) times an $F[f_I,f_R]$ where Q equals V_I/V_R, and the power of the statistical test can be computed from this.
The test is significant if the observed F value > $F[f_I,f_R;\alpha]$ which is the critical value for a test at the α significance level. When the null hypothesis is false the F value calculated is 1 + nQ times as large as when the null hypothesis is true, so the chance of getting a significant result is

$$\text{Prob}\{ F[f_I,f_R] > F[f_I,f_R;\alpha]/(1 + nQ) \}.$$

This can be evaluated using computer programs for the F distribution.

Table 1. Form of analysis of variance for balanced data

SOURCE	D.F.		Mean Square	Expected M.S.
Sires	s-1	$= f_S$	S	.....
Trials	t-1	$= f_T$	T	.....
Interaction	(s-1)(t-1)	$= f_I$	I	$V_R + nV_I$
Residual	st(n-1)	$= f_R$	R	V_R

The available data from 46 central test and on-farm sire evaluations considered here are very unbalanced, with a majority of the sire by trial subclasses empty. However, provided that some sires are compared in at least two trials a test for interaction can be made. This can be done by fitting to the data a model which includes only sire and trial effects, and comparing this with a model which fits effects for all non-empty sire by trial subclasses. The difference between the sums of squares accounted for by the two models is due to interaction, and can be tested against the residual mean square from fitting the subclasses model. Because of the unbalanced structure, the distribution of the ratio I/R is no longer that of (1 + nQ) times a standard F variate when the null hypothesis is false. However, to approximate the power of the test it is reasonable to assume that the observed F will have a distribution which is approximately (1 + kQ) times $F[f_I,f_R]$ if k is chosen to give the correct mean value of F.

The appropriate value of k can be found using results which are well known (Henderson 1984, p.130). Suppose **D** is the coefficient matrix for fitting the additive model with no interaction, where the first t equations are for trial effects and the last s equations are for sire effects. Then **D** has the following form:

$$\begin{bmatrix}
n_{1+} & 0 & 0 & \cdots & n_{11} & n_{12} & n_{13} & \cdots \\
0 & n_{2+} & 0 & \cdots & n_{21} & n_{22} & n_{23} & \cdots \\
.. & .. & .. & \cdots & .. & .. & .. & \cdots \\
.. & .. & .. & \cdots & .. & .. & .. & \cdots \\
n_{11} & n_{21} & n_{31} & \cdots & n_{+1} & 0 & 0 & \cdots \\
n_{12} & n_{22} & n_{32} & \cdots & 0 & n_{+2} & 0 & \cdots \\
.. & .. & .. & \cdots & .. & .. & .. & \cdots \\
.. & .. & .. & \cdots & .. & .. & .. & \cdots
\end{bmatrix}$$

where n_{ij} is the number of progeny of sire j in trial i, and a + subscript signifies summation over the

corresponding subscript. The matrix is not of full rank but solutions can be obtained either by using a generalised inverse or by imposing restrictions on the solutions. In the present work a generalised inverse denoted $\mathbf{D}^-$ was used.

A second matrix, denoted **C**, is also required. The elements of **C** are n_{ij}^2 where n_{ij} occurs in **D** while the diagonal elements are the sums of the n_{ij}^2 terms in the corresponding rows. Other terms are zero. The trace of the matrix product $\mathbf{D}^-\mathbf{C}$ is then readily computed.

If c denotes the number of filled subclasses, the number of degrees of freedom for interaction is (c-1)-(s-1)-(t-1). If N denotes the total number of progeny in all c subclasses, the value of k, the coefficient of Q, can be found as

$$k = (N - \text{trace}(\mathbf{D}^-\mathbf{C})) / f_I.$$

The value of f_R is N - c. In the present problem this was very large (>7500) and was taken to be effectively infinity, so that V_R can be taken as known, in which case the problem reduces to dealing with a chi-squared distribution. Thus the power of the test was approximated by

$$\text{Prob}\{ X^2[f_I] > X^2[f_I;\alpha] / (1 + kQ) \}$$

where $X^2[f_I]$ is a chi-squared variable with f_I degrees of freedom and $X^2[f_I;\alpha]$ is the critical value at the α significance level.

A file was constructed in which the subclass numbers of progeny for 63 sires in 46 trials were recorded. These were sires which had been progeny tested in more than one trial. Not all sires would contribute to a test of interaction, but it was not necessary to identify and remove such sires since the results are unaffected by their presence or absence assuming effectively infinite error degrees of freedom.

Three sets of 12 rams were considered for the special trial. Set A included sires with the fewest progeny in previous trials, set B the sires with the most progeny, and set C was chosen with the aim of better linking fine and medium wool trials. It was assumed that the new trial used 30 progeny per ram. In addition, three other sets were constructed in which 60 progeny from 6 of the rams in sets A, B and C were used to form sets AA, BB and CC.

RESULTS

There were 180 filled cells in the 63 by 46 matrix of progeny numbers. The value of f_I for the existing data set was 72, and for the augmented data sets A, B and C was 83, corresponding to the degrees of freedom for 12 more classes with one extra trial. For sets AA, BB and CC the number of degrees of freedom was 77. The values of k for the 7 designs were (0 denotes no extra data):

0	A	B	C	AA	BB	CC
35.9335	34.5495	35.9763	35.5322	35.6744	37.4590	36.7217

Clearly there were very minor differences between designs in this factor, though the designs with 60 progeny per sire had larger k values than the designs with 30. It is thus expected that data sets with the same number of degrees of freedom would have similar power.

To calculate power, it is necessary to assume a value for Q. For a trait with a heritability of 0.4, the sire variance component is $0.1V_P$ and the within sire variance component V_R is $0.9V_P$, where V_P is the

phenotypic variance. If the variance due to sire by environment interaction is 20% of the genetic variance between sires it would be $0.02V_P$ and thus Q = 1/45. A level of interaction of this magnitude, implying that only 80% of a sire's genetic superiority in one trial could be expected to be repeated in a second trial, would certainly be important. It might also be argued that a level of interaction of only half this magnitude would also be important, and this would correspond to a value of Q = 1/90. Calculations have been made for both values of Q, and at significance levels of 0.05 and 0.01, the levels conventionally used. The power values are shown in Table 2.

DISCUSSION

The conclusions are clear. If the magnitude of interaction is as large as Q = 1/45, there is a very high chance of obtaining significant evidence of its existence with the existing data set. The power is slightly improved by all the augmented designs, primarily because more degrees of freedom are available.

Table 2. Power of tests for interaction for different amounts of interaction and data structures.

Data Structure	Q = 1/45		Q = 1/90	
	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
0	0.9668	0.8991	0.6664	0.4294
A	0.9758	0.9208	0.6921	0.4569
B	0.9811	0.9352	0.7166	0.4860
C	0.9796	0.9310	0.7091	0.4770
AA	0.9731	0.9143	0.6854	0.4501
BB	0.9798	0.9322	0.7148	0.4849
CC	0.9772	0.9253	0.7029	0.4705

If interaction is only half as strong, the probability of success in testing for interaction is not nearly so great, but is still more than 50%. The augmented designs increase the power rather more in this case, but the increase is still not large.

Of the augmented designs, Set B gives the highest power, in agreement with previous results in balanced designs that large family size is advantageous (James 1979). Power is always higher for 12 rather than for 6 sires used in the extra trial, showing that the extra degrees of freedom are more valuable than the increase in k obtained in the designs with 6 sires in the extra trial. It would clearly be worthwhile to undertake an analysis of such a data set to test for interaction, but the addition of data from a special trial would have only a small effect on the power.

ACKNOWLEDGEMENTS This work was supported by a grant (UNS179) from IWS. We are grateful to Allan Casey, Kathy Coelli and Kevin Atkins for their help with the available data structure.

REFERENCES

Henderson, C.R. (1984) "Applications of linear models in animal breeding". University of Guelph, Guelph.

James, J.W. (1979) in "Introduction of New Dairy Cattle Genotypes into Australia",p.20, SCA-Australia Technical Report Series No.2.

Woolaston, R.R (1987) in "Merino Improvement Programs in Australia",p.421, editor B.J.McGuirk, Australian Wool Corporation, Melbourne.