

## ESTIMATION OF HERITABILITY IN SELECTION EXPERIMENTS

Robin Thompson and Kevin D. Atkins

AFRC Institute of Animal Physiology and Genetics Research, Roslin, Midlothian, EH25 9PS, U.K. and Agricultural Research and Veterinary Centre, Forest Road, Orange, NSW 2800, Australia.

### SUMMARY

Two alternative ways of analysing selection experiments are to compare response with selection differential and to fit an individual animal model by maximum likelihood. A technique is suggested to enable the information from between lines on the heritability estimate to be quantified and an example is given.

### INTRODUCTION

This paper arose out of estimating heritability from sheep selection experiments (Atkins and Thompson, 1986) and more general considerations of designing selection experiments with overlapping generations. There are at least two approaches that might be followed. Firstly, response could be compared with selection differential. The regression of response, in terms of difference between divergent lines or one line with a control, with selection differential is a natural way of considering selection experiments and gives a graphical way of presenting the results. Hill (1971) and Johnson (1977) has shown how to construct appropriate variances for responses. James (1986) and Thompson and Juga (1989) have discussed problems of defining selection differentials in overlapping generations.

A second approach would be to use maximum likelihood estimation with a individual animal model. This is becoming computationally easier especially using algorithms from Meyer (1986, 1989) and can give estimates of trend and genetic parameters not biased by selection. However, these procedures, especially the derivative-free ones are something of a black box and it is not obvious whether the information mainly comes from within or between line comparisons, parent-offspring or collateral information.

The purpose of this paper is to suggest an analysis that allows assessment of the amount of information from between lines comparisons in a selection experiment and apply it to a experiment.

### MATERIALS AND METHODS

It is not clear how an individual animal model could be reparameterized when we have two lines and wish to find estimates of heritability both within and between lines. A sequence of analyses however can be constructed that does allow the question to be answered. The within lines likelihood can be accumulated from analyses within each line in turn and this can give estimates of some fixed effects in each of the lines. It is not altogether obvious how to do a direct between line analysis finding a pooled estimate for each effect because of the complicated error structure depending on the drift variances and the variance parameters to be estimated. However, a combined analysis on all the data can be carried out and the difference in log-likelihood,  $L_D$  between the combined analysis and within line analysis can be constructed. By minimizing  $L_D$ , as a function of heritability, one can in effect get a between line estimate of heritability. This heritability will make the estimates of effects between lines as consistent as possible taking account of the variance structure.

Informally if we have generation means of  $y_1$  and  $y_2$  in two lines then approximately

$$y_1 = a + s_1(h^2)h^2 \text{ with } \text{var}(y_1) = V_1 \quad (1)$$

$$y_2 = a + s_2(h^2)h^2 \text{ with } \text{var}(y_2) = V_2 \quad (2)$$

where  $a$  is the environmental effect associated with each generation and  $s_1(h^2)h^2$  measures the expected response in each generation and is to a first approximation linear in  $h^2$ . The selection differentials  $s_i(h^2)$  can be approximated by  $[(a_i(h^2 + \Delta h^2) - a_i(h^2))/\Delta h^2]$  where  $a_i(h^2)$  is the estimate of effects in line  $i$  using heritability  $h^2$ . An estimate of  $h^2$  from (1) and (2) is  $rs^{-1}$ , with  $s = [V_1^{-1}s_1(h^2) - V_2^{-1}s_2(h^2)]' (V_1^{-1} + V_2^{-1})^{-1} [V_1^{-1}s_1(h^2) - V_2^{-1}s_2(h^2)]$  and  $r = [V_1^{-1}s_1(h^2) - V_2^{-1}s_2(h^2)]' (V_1^{-1} + V_2^{-1})^{-1} [V_1^{-1}y_1 - V_2^{-1}y_2]$  and if  $V_1^{-1} = V_2^{-1}$ ,  $s = [s_1(h^2) - s_2(h^2)]' 2V_1^{-1} [s_1(h^2) - s_2(h^2)]$  and  $r = [s_1(h^2) - s_2(h^2)]' 2V_1^{-1} [y_1 - y_2]$ . This is the form of a weighted regression estimate of generation differences ( $y_1 - y_2$ ) on differences in selection differentials.

The method is applied to a selection experiment reported by Atkins and Thompson (1986). The first two years were initial years but after that 10 rams were mated per year in each of three lines, high, control and low. 260 ewes lambed at 2-6 years and 200 lambs were reared. The generation interval was 2.5 years. The selection intensity was 1.37 in males and 0.13 in females.

### RESULTS

As a benchmark the case when a control group is kept and parents are randomly selected to produce animals in the next generation is considered. In these designs there are  $s$  individuals, each individual has  $n$  offspring and one of these offspring is chosen to have  $n$  offspring, the process being repeated for  $g$  generations to give  $(gn+1)$  recorded animals. In practice there are other branches in the pedigree and many links providing information are ignored. Female pedigrees in species such as pigs or sheep are modelled by small values of  $n$ , larger  $n$  is more appropriate to male pedigrees.

Table 1. Asymptotic variance per observation for maximum likelihood (ML) and regression (Reg) estimates of heritability ( $h^2$ ) for various  $h^2$ , family sizes ( $n$ ) and generations ( $g$ )

$g$	$h^2$	$n$	0.05		0.10		0.4		0.6	
			ML	Reg	ML	Reg	ML	Reg	ML	Reg
1	2	2	5.50	6.07	5.64	5.12	5.89	6.12	5.43	5.82
1	3	3	4.53	5.46	4.70	5.65	5.53	5.76	5.19	5.49
1	5	5	3.60	5.02	3.98	5.22	5.49	5.76	5.27	5.52
1	9	9	2.76	4.86	3.32	5.23	6.01	6.40	5.94	6.18
1	17	17	2.10	5.04	2.88	5.75	7.48	8.13	7.63	7.92
2	2	2	4.19	5.11	4.46	5.22	5.26	5.41	4.92	5.13
2	3	3	3.63	4.83	3.98	4.98	5.21	5.37	4.91	5.10
2	5	5	3.02	4.64	3.51	4.40	5.44	6.65	5.18	5.40
2	9	9	2.43	4.68	3.10	5.10	6.20	6.56	5.98	6.32
2	17	17	1.94	4.97	2.86	5.75	7.88	8.61	7.76	8.38
4	2	2	3.57	4.63	3.91	4.75	4.95	5.07	4.64	4.83
4	3	3	3.14	4.52	3.62	4.68	5.06	5.20	4.75	4.97
4	5	5	2.74	4.48	3.31	4.74	5.44	5.65	5.12	5.43
4	9	9	2.28	4.59	3.03	5.04	6.32	6.73	5.99	6.51
4	17	17	1.88	4.93	2.89	5.77	8.09	9.01	7.81	8.79

Table 1 shows asymptotic variances of ML estimates of  $h^2$  and an estimate derived from parent-offspring covariances and phenotypic variances. These results show the gain from increasing generations in reducing the variance of  $h^2$  per observation, especially at small  $n$ . In one sense there are  $s$  families of size  $gn + 1$ , but unequal covariances between members

of the family. For example, there are  $gn(n + 1)/2$  half-sibs and  $gn$  offspring-parent pairs. Qualitatively these results mimic Robertson's results for one generation. The parent-offspring information extracts a large proportion of the information especially for heritabilities near 0.5. When  $g = 1$  the information on heritability comes from the regression of offspring on parent and the derivation about the regression line which has a variance that is a function of  $h^2(1-h^2)$  (Robertson, 1977). As  $h^2(1-h^2)$  is relatively flat over the range of 0.4-0.6 it will be difficult to estimate heritability from this information.

Table 2 shows estimates of year means for different values of heritability near to the estimates from analysis of the first eleven years. As  $h^2$  increases the estimated year effects decrease in the high line and increase in the low line and are relatively constant in the control line. The terms  $s_t(h^2)$  can be found by looking at differences in columns in Table 2.

**Table 2.** Year estimates for each line for different heritability values

$h^2$	High			Low			Control		
	0.58	0.60	0.62	0.58	0.60	0.62	0.58	0.60	0.62
<u>Years</u>									
1	0.926	0.938	0.944	1.504	1.495	1.486	0.846	0.827	0.828
2	2.063	2.067	2.070	2.553	2.548	2.543	2.099	2.096	2.092
3	1.300	1.264	1.229	2.064	2.104	2.144	1.958	1.928	1.917
4	3.202	2.146	2.090	2.610	2.654	2.697	2.667	2.645	2.641
5	1.133	1.055	0.976	1.638	1.702	1.767	1.507	1.512	1.517
6	1.870	1.786	1.701	3.031	3.102	3.173	1.654	1.648	1.643
7	2.978	2.871	2.761	3.260	3.353	3.446	3.380	3.375	3.370
8	2.837	2.712	2.584	3.221	3.351	3.481	3.059	3.067	3.075
9	2.279	2.141	2.000	2.775	2.921	3.068	2.514	2.524	2.534
10	1.748	1.577	1.404	2.500	2.767	2.936	2.655	2.651	2.659
11	2.309	2.119	1.926	2.349	2.512	2.676	2.410	2.414	2.418

The terms are approximately independent of  $h^2$  and change approximately linearly with year. The terms are approximately equal in magnitude for the selected lines and larger than in the control line. The variance matrix of the year estimates ( $a_t$ ) is given in Table 3 with  $\text{var}(a_t) > \text{var}(a_s)$  ( $t > s$ ) and covariance ( $a_t, a_s$ ) ( $t > s$ ).

**Table 3.** Control variance-covariance matrix of year estimates ( $h^2 = 0.60$ ) ( $\times 10^3$ )

1	32										
2	08	38									
3	12	10	27								
4	10	10	15	29							
5	10	10	13	17	31						
6	10	10	13	15	19	31					
7	10	10	13	15	17	20	35				
8	10	11	13	16	17	18	22	38			
9	10	11	14	17	18	19	20	25	43		
10	10	10	13	16	19	19	20	30	28	45	
11	09	10	13	15	19	19	21	27	25	30	46
	1	2	3	4	5	6	7	8	9	10	11

converging to a constant value as  $t$  increases indicating the drift component. Table 4 gives estimates of heritability standard errors and their information (inverse of variance) (found by

approximating the likelihood by a quadratic in  $h^2$ ) for combinations of lines for differing number of years.

**Table 4.** Summary of heritability (s.e.) estimates [Information expressed as inverse of sampling variance]

<u>Code</u>	<u>Flock</u>	<u>Year 6</u>	<u>Year 11</u>	<u>Year 16</u>	<u>Year 21</u>
1	H	0.640 (0.060) [278]	0.618 (0.042) [568]	0.627 (0.034) [852]	0.640 (0.030) [1144]
2	L	0.578 (0.068) [213]	0.585 (0.048) [425]	0.599 (0.039) [617]	0.647 (0.032) [990]
4	C	0.731 (0.064) [248]	0.667 (0.042) [578]	0.683 (0.033) [925]	0.695 (0.028) [1249]
3	H+L	0.598 (0.037) [717]	0.596 (0.024) [1787]	0.593 (0.019) [2824]	0.610 (0.016) [3853]
5	H+C	0.649 (0.041) [589]	0.621 (0.027) [1360]	0.634 (0.022) [2074]	0.648 (0.019) [2865]
6	C+L	0.658 (0.044) [521]	0.631 (0.030) [1142]	0.636 (0.023) [1898]	0.660 (0.020) [2391]
7	H+C+L	0.630 (0.033) [942]	0.615 (0.021) [2293]	0.614 (0.017) [3651]	0.631 (0.014) [4994]

**Table 5.** Combination of between flock information

	<u>Year 6</u>	<u>Year 11</u>	<u>Year 16</u>	<u>Year 21</u>
(H+C)-H-C	63	214	297	393
(C+L)-C-L	60	139	302	352
(H+L)-H-L	226	794	1401	1714

#### DISCUSSION

Table 4 shows that the estimates of heritability were remarkably consistent both over years and over lines. Table 5 shows that the information on heritability increased approximately linearly with time. Table 5 shows that the comparison of High with Low is four times as valuable as a comparison between a line with the control as might be expected since the selection differentials in the first case are twice those in the second case. The information per animal in the whole experiment is 0.4 corresponding to a variance of 2.5 or half the values for  $h^2 = 0.6$  in Table 1. Within line information is roughly two thirds as valuable as the total response. In selection experiments with lower heritabilities and more effective selection on females the contribution of within line information will be less important.

#### REFERENCES

- ATKINS, K.D. and THOMPSON, R. (1986). *Animal Production*, 43 : 421-435.  
 HILL, W.G. (1971). *Biometrics*, 27 : 293-311.  
 JAMES, J.W. (1986). *Animal Production*, 42 : 411-415.  
 JOHNSON, D.L. (1977). In *Proceedings of the International Conference on Quantitative Genetics*. eds. Pollak, E., Kempthorne, O. and Bailey, T.B., 851-858.  
 MEYER, K. (1986). *Proc. 3rd World Congr. Genet. Appl. Livestock Prod.* 12 : 454-459.  
 MEYER, K. (1989). *Genet. Sel. Evol.* 21 : 317-340.  
 ROBERTSON, A. (1977). *J. of Anim. Breed. and Gen.*, 94 : 131-135.  
 THOMPSON, R. and JUGA, J. (1989). *Animal Production*, 49 : 203-208.