

# EFFECT OF SELECTIVE RECORDING ON ESTIMATES OF HERITABILITY

L. ALFONSO<sup>1</sup>  
J.L. NOGUERA  
D. BABOT  
J. ESTANY

Area of Animal Production, Centre UdL-IRTA  
25198 Lleida

## SUMMARY

In this work the effect of selective recording from older animals on estimates of heritability is analysed. Heritability of litter size was estimated in a pig population where data were selectively recorded for the older sows. Estimates were obtained in different data sets derived from all the available information. Data set were compared in terms of predictive ability. Underestimated heritability values were obtained when all available data were used. A maternal trait was also simulated and missing data generated by deleting records from females without known parents. Different animal models, usually applied when there is no selection or selective breeding is carried out, were considered in the heritability estimation of this trait. Results showed that none of these models offered unbiased estimates. Only the models that took into account the effect on the residual mean and the genetic mean and variance removed an important part of this bias. However, to be applied, selection carried out should be known. It is not obvious how to overcome a selective recording problem when the selection process followed is unknown.

**KEY WORDS:** Heritability  
Selective recording

## INTRODUCTION

Important efforts have been made recently to improve (co)variance estimation methods and algorithms. Simultaneously, significant resources have been spent to have bigger and better management information systems for animal populations. However, data could be selectively recorded, especially for the older animals, and this selection could affect the estimates of heritability ( $h^2$ ).

We can think in two different kinds of selection, distinguishing between selective breeding and selective recording. Although theoretically selective breeding and selective recording are two different kinds of selection, in real terms they are closely related.

---

Recibido: 26-11-97

Aceptado para su publicación: 1-10-98

<sup>1</sup> Present address: Area of Animal Production, Universidad Pública de Navarra, Campus Arrosadía, 31006 Pamplona.

According Henderson (1975, 1984) and in a simple model: selection can be based directly on records ( $y$ ), on genetic values ( $a$ ) or on residuals ( $e$ ). To think of a direct selection based on  $a$  or  $e$  is artificial because neither  $a$  nor  $e$  can be known exactly (R.Thompson, comment in Im *et al.*, 1989). In fact, only selection based on  $y$  could be carried out, and it will affect the distributions of  $a$  and  $e$ , basically its mean and variance values, depending on the  $h^2$  value.

The effect of selective breeding on  $h^2$  estimates has been profusely analysed. Different models have been suggested to take into account the effect of this selection calling to the properties of the Animal Model and the REML methodology. The effect of selective recording on  $h^2$  estimates has not received, however, the same deep treatment in the literature. The models suggested when selective breeding is carried out can not take into account completely the effect of selective recording on  $h^2$  estimates. Nevertheless it has not been established if they can offer less biased estimates than the use of the simplest Animal Model.

The aim of this work is first to show how  $h^2$  estimates could be affected by selective recording of data from older animals, and second to analyse, by simulation, the biases produced on these estimates under different models.

## MATERIAL AND METHODS

### Pig population

The effect of selective recording was analysed in a Landrace pig population. The trait considered was the number of piglets born alive (NBA). A total of 38047 records were known. The population was commercial and maternally specialised, organised in two nucleus and three multiplication herds. Pedigree data comprised 15208 animals with 10612 sows with litters recorded in the period 1982-1992. In this period the population was selected mainly by productive traits. Only in the last two years prolificacy was considered in the breeding goal.

Since 1991 available information was collected by a new management information system, the GTEP-IRTA System (Estany *et al.*, 1992; Noguera *et al.*, 1992). It was known that all records of litter size from non productive sows in 1988 were erased from the data base although genealogical information was maintained. This means that only a sample of all the records produced during the period 1982-1987 was available. It can be suspected that the older sows presents in 1988 were the most productive, and so, data could be partially selected for litter size. A more detailed description of data from this population can be found in Ruiz (1993).

Different data sets were used in order to assess how missing records from older animals affects the  $h^2$  estimates. A first data set comprised all available information (D1982). A second data set (D1986) was derived assuming that herd size was constant in time. Information from different years was considered for each herd depending on when the number of productive sows was stabilised: in one nucleus data were considered since 1986; in the other nucleus and in two multiplication herds since 1987; and in the last multiplication herd since 1988. The last data set considered (D1988) was derived ignoring all information produced before 1988 when some records were erased. Data sets are described in Table 1.

## Simulated Data

A three generations population was simulated assuming litter size as the trait recorded. A number of 10 males and 200 females per generation were considered. Animals were randomly selected and mated avoiding full-sibs. Genetic values of parental animals, all assumed unrelated and unselected, were generated from a normal distribution considering the infinitesimal model and Hardy-Weinberg and linkage equilibrium. Genetic values of descendants were simulated as the average genetic value of parents plus a Mendelian sampling term. Phenotypic values were computed for females adding to the genetic value a normal residual value with mean nine and variance 7.12 minus the additive genetic variance ( $\sigma_a^2$ ). Two values of  $\sigma_a^2$  were considered: 0.36 and 3.56 ( $h^2=0.051$  and 0.5 respectively). Although 0.5 is not a realistic value for the  $h^2$  of litter size, it was also considered to analyse what happens with higher values. Missing records were produced by deleting records of parental females from the simulated datasets. Records deleted were those lower than an established value ( $\omega$ ) of the phenotypic value (9 or 12). One hundred replicates were carried out to obtain an empirical standard error of estimated  $h^2$  no greater than 0.01.

The situation simulated is clearly different from the pig population previously referred. The main differences are, first, the structure of population (number of animals, number of generation, mating type,...), second, the model that explains the trait, and third, the selective recording process. The objective was to analyse a very simplified situation that helps to understand how selective recording from older animals can affect ( $h^2$ ) estimates under different statistical models of analysis.

**TABLE 1**  
**DESCRIPTION OF DATA SETS USED IN THE ESTIMATION OF**  
**HERITABILITY OF LITTER SIZE IN THE PIG POPULATION**

*Descripción del conjunto de datos utilizado para estimar el valor de heredabilidad del tamaño de camada en la población de cerdos*

Data set <sup>1</sup>	Number of animals with data	Number of litters	Mean	Phenotypic variance
D1982	10,600	38,047	8.59	6.00
D1986	9,468	29,330	8.60	6.71
D1988	8,728	26,222	8.62	6.71

<sup>1</sup> See text explanation.

<sup>1</sup> Ver explicación en el texto.

## Pig data analyses

REML heritability estimates for NBA were obtained for the three data sets derived, considering always all the pedigree information available. The model employed was the usual Animal Repeatability Model used in the genetic evaluation of the population (Estany *et al.*, 1995), that includes as fixed factor the herd-year-season of parity. The evolution of

mean and standard phenotypic deviation of NBA according to the birth year of sows was analysed using all available data.

Data sets D1982 and D1986 were compared in terms of predictive ability. The predictive ability of each data set was studied as follows. Each data set was subdivided in two parts. Each part was of the same size, with all animals represented and a similar distribution of records within levels of fixed factor. The records of each part were predicted from BLUE and BLUP values obtained previously with the other part. After predictions, correlations between observed and predicted records, and mean square errors were computed. The analyses were performed using DFREML (Meyer, 1989), PEST (Groeneveld *et al.*, 1990) and SAS (SAS Inc., 1995) softwares.

### Simulated data analyses

Different models, summarised in Table 2, were used to estimate the ( $h^2$ ) in the simulated data. Models were derived from the following expression of Mixed Model Equations of Animal Model (Alfonso and Estany, in preparation):

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A_{\alpha}^{*-1} \end{bmatrix} \cdot \begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

Assuming no inbreeding, the rules to compute  $A^{*-1}$  are: 1) let  $i$ ,  $j$ , and  $k$  represent the equation numbers of an individual, its sire or its sire's group (if base sire) and its dam or dam's group (if base dam), respectively;  $m = 0, 1$  or  $2$  = number of base parents of the individual;  $\delta_i$ ,  $\delta_j$  and  $\delta_k$  the proportion of genetic additive variance after selection, in animal, sire (sire's group) and dam (dam's group), respectively; 2) proceed through the list of animals represented in  $a$ , attributing to each:  $x_i = 2/\delta_i$ , if  $m=0$ ;  $x_i = [4/(2+\delta_j)]/\delta_i$  if  $m=1$  and the sire is the base parent;  $x_i = [4/(2+\delta_k)]/\delta_i$  if  $m=1$  and the dam is the base parent;  $x_i = [4/(2+\delta_j+\delta_k)]/\delta_i$  if  $m=2$ ; 3) then add the following contributions to  $A^{*-1}$ :  $x_i$  to elements ( $i,i$ );  $-0.5x_i$  to elements ( $i,j$ ), ( $j,i$ ), ( $i,k$ ) and ( $k,i$ ); and  $0.25x_i$  to elements ( $j,j$ ), ( $k,k$ ), ( $j,k$ ) and ( $k,j$ ).

The first model (M1) was the simplest random model. The rest of models (M2 to M7) included the generation of animals as fixed factor with two levels, the parental and the two other generations. Three of these models (M3, M5 and M7) considered also the group of parental animals as fixed factor with two levels: parental animals with records and without records. For models M6 and M7 the value of  $\delta$  was computed considering that the variance of a variable with a standard normal distribution and selected by truncation in a value, selecting a proportion  $p$ , is reduced by a proportion  $K=(f(v)/p) [(f(v)/p) - v]$ , being  $f(v)$  the value of distribution in Y-axis. With known  $v$  and  $h^2$ , the proportion of the reduction of genetic variance can be predicted as  $Kh^2$  (Robertson, 1977), so  $\delta=(1-Kh^2)$ . To simplify the estimation procedure, both males and females from parental generation were assumed affected by this  $\delta$  value (Table 2). A value of  $\delta=10^8$  was also considered for models M4 and M5 as an approach to treat base animals as fixed. REML estimations were obtained by using DFREML software (Meyer, 1989), with the corresponding  $A^{*-1}$  matrix.

**TABLE 2**  
**MODELS USED IN THE ESTIMATION OF HERITABILITY IN SIMULATED POPULATIONS**

*Modelos utilizados para estimar la heredabilidad en las poblaciones simuladas*

Model	Fixed factors		Random factors	
	Generation	Genetic groups	Animal ( $\delta$ value)	Residual
M1	—	—	✓ ( $\delta=1$ )	✓
M2	✓	—	✓ ( $\delta=1$ )	✓
M3	✓	✓	✓ ( $\delta=1$ )	✓
M4	✓	—	✓ ( $\delta=10^8$ )	✓
M5	✓	✓	✓ ( $\delta=10^8$ )	✓
M6	✓	—	✓ ( $\delta \neq 1$ ) <sup>1</sup>	✓
M7	✓	✓	✓ ( $\delta \neq 1$ ) <sup>1</sup>	✓

<sup>1</sup> For a truncation point in 9 (standard value = 0),  $\delta$  values (proportion of genetic variance after selection) were 0.68 and 0.97 for  $h^2$  values .5 and 0.05 respectively, and 0.59 and 0.96 for a truncation point in 12 (standard value = 1.12). A value  $\delta = 10^8$  was also considered.

<sup>1</sup> El valor de  $\delta$  (proporción de varianza genética tras la selección) fue de 0,68 y 0,97, para  $h^2$  y 0,05 respectivamente, cuando el punto de truncamiento era 9 (valor tipificado = 0), y de 0,59 y 0,96 cuando el punto de truncamiento era 12 (valor tipificado = 1,12). También se consideró un valor  $\delta = 10^8$ .

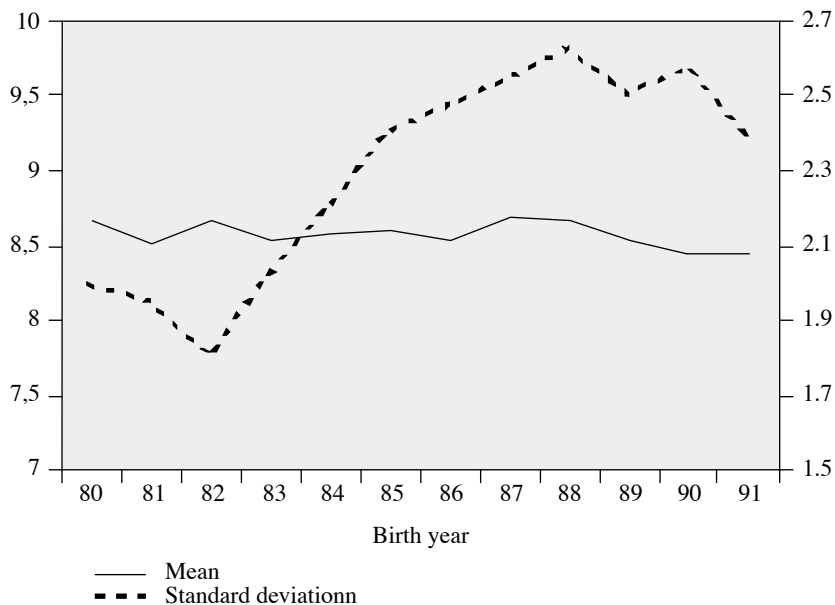
## RESULTS AND DISCUSSION

### Pig data analyses

The evolution of mean and standard phenotypic deviation of all litter size records available according to the birth year of sows (Fig. 1), showed that the variance increased significantly yearly (0.094 standard yearly deviations), whereas the mean of phenotypic values was quite constant. Older information seems to be selected in a nonrandom manner. This selection could explain the yearly increase of the phenotypic variance. The expected decrease of the phenotypic mean was not observed maybe due to a significant opposite environmental trend (0.04 piglets/litter yearly). This different evolution of mean and standard phenotypic deviation was not observed when D1986 and D1988 were considered, indicating that these two datasets were not, presumably, a selected sample of all records produced.

Parameter estimates obtained considering all the available genealogical information are reported in Table 3. When all available information (D1982) was used the estimated value of ( $h^2$ ) was 0.033 (Table 3). It was observed that estimates were greater (close to 0.06) when not all available information was considered. However, the repeatability value was rather constant indicating that the whole animal variability was similar in each data set. No differences were observed between values estimated using D1986 and D1988. D1986 comprised more data than D1988 and it was used to be compared with D1982 in terms of predictive ability.

The predictive ability of D1986 was better than that of D1982. The mean square errors were 0.93 and 0.94 and the correlation between observed and predicted records 0.27 and 0.26 for D1986 and D1982 respectively. Differences were small but similar to those obser-



**Fig. 1.**—Evolution of mean and standard deviation of litter size by year of birth of sows when all available records of pig population were considered (D1982)

*Evolución de la media y la desviación típica del tamaño de camada en función del año de nacimiento de las cerdas, al considerar todos los datos conocidos de la población de cerdos (D1982)*

ved in other works (Estany and Sorensen, 1995). Furthermore, in terms of loss of response D1986 was also better than D1982 especially for high selection pressures.

To summarise, the  $h^2$  value estimated in this population seems to be greater than the value initially estimated assuming that data was randomly collected. Selective recording of data from older animals is a likely reason to explain this underestimation.

**TABLE 3**

**PARAMETER ESTIMATES FOR DIFFERENT DATA SETS IN THE PIG POPULATION**

*Estimas de los parámetros obtenidas con distintos conjuntos de datos en la población de cerdos*

Data set <sup>1</sup>	$h^2$	se( $h^2$ )	$c^2$	se( $c^2$ )	$\sigma^2_a$	$\sigma^2_c$	$\sigma^2_p$
<b>D1982</b>	0.033	0.006	0.094	0.007	0.194	0.546	5.807
<b>D1986</b>	0.057	0.013	0.080	0.013	0.368	0.519	6.506
<b>D1988</b>	0.055	0.015	0.077	0.013	0.361	0.501	6.518

Abbreviations:  $h^2$ , heritability;  $c^2$ , fraction of phenotypic variance due to non-genetic permanent effects;  $\sigma^2_a$ , additive genetic variance;  $\sigma^2_c$ , variance due to non-genetic permanent effects;  $\sigma^2_p$ , phenotypic variance; se, approximated standard error.

<sup>1</sup> See text explanation.

Abreviaciones:  $h^2$ , heredabilidad;  $c^2$ , fracción de varianza fenotípica debida a efectos no genéticos permanentes;  $\sigma^2_a$ , varianza debida a efectos no genéticos permanentes;  $\sigma^2_p$ , varianza fenotípica; se, error típico aproximado.

<sup>1</sup> Ver explicación en el texto.

## Simulated data analyses

Results from simulation, when missing records from older dams were selected and different models used, can help to understand how this selection affects the  $h^2$  estimation. These results are collected in Tables 4 and 5.

A random model (M1) that does not account for any change on the distributions of  $a$  and  $e$ , overestimates the  $h^2$ . Residual differences between animals are assigned to genetic differences, underestimating the value of the residual variance ( $\sigma^2_e$ ) and overestimating the value of the genetic variance ( $\sigma^2_a$ ). This is especially true for low values of  $h^2$ .

The inclusion of a generation fixed factor (M2) was considered to take into account the effect of selection over the mean of  $e$ . Results show that underestimated values of  $h^2$  are obtained. These results are related to those obtained in the pig population analysed when all available data were used. The model used in the pig population included as fixed factor the herd-year-season where litters were produced. Both factors group the sample of selected records in few levels underestimating the genetic differences between animals.

The  $h^2$  underestimation produced by the inclusion of a generation fixed factor was greater when a genetic group effect was also considered (M3). This new fixed factor should account for the effect of selection on the mean of  $a$ . In general, the genetic group model (Quaas, 1988) only offers unbiased estimates when the effect of selection on the variance of  $a$  can be ignored (Pieramati and Van Vleck, 1993). In our situation this effect can not be ignored and the underestimation produced by including a generation fixed factor is weekly increased by also including the genetic group effect.

To take into account only the effect of missing information over the means of  $e$  and  $a$  does not seem to be sufficient. It is also necessary to account for the effect on the variance of  $e$  and  $a$ . To account for the effect over the variance of  $e$ , a general treatment of heterogeneous variances, as proposed by Foulley and Quaas (1995), should be considered. But, in general, it is difficult to expect that we can have sufficient information to estimate so many parameters as needed in a general treatment (R.Thompson comment in Foulley and Quaas, 1995).

One approach to account for the effect on the variance of  $a$ , could be to ignore this variance in the older animals, those with selected records, treating these animals as fixed factors (Graser *et al.*, 1987). Big overestimations were always obtained and the standard errors increased (M4 and M5 with  $\delta=10^8$ ). As when the M1 model was used, residual differences between animals are confused with the genetic differences.

Another approach is to consider the reduction of the genetic variance in base animals produced by selection, as proposed by Van der Werf and Thompson (1992). Including the effect on the variance of  $a$  (M6 and M7),  $h^2$  estimates were nearest to simulated values but estimated variances still biased. The underestimation of variance of  $e$  decreased, however, when more information in non-selected generation was known. By considering 10 females in the last generation the variance of  $e$  was less underestimated (e.g.  $6.67\pm 0.024$  and  $6.65\pm 0.024$ , for truncation points 9 and 12 respectively and simulated value 6.76). The model M7 was the model that offers the best  $h^2$  estimates of the different models compared.

In general, the effect of the simulated selective recording depended on the number of missing records and the  $h^2$  value. As greater was the number of missing records ( $\omega=12$ ) the

models that not account for reduction of  $\sigma_a^2$  often offered less biased  $h^2$  estimates. As lower was the  $h^2$  value the simple random model and the fixed base animals model offered more biased  $h^2$  estimates.

In conclusion, all the models evaluated offered biased estimates of variance components. The least biased estimates of  $h^2$  were obtained by using the model that considered the effect of selection on the mean of  $a$  and  $e$  and the variance of  $a$ . This model could be useful in some situations, especially for small  $h^2$  values, but to be applied it is necessary to know the effect of selection on the genetic variance and this value is usually unknown.

**TABLE 4**  
**PARAMETER ESTIMATES UNDER DIFFERENT MODELS IN POPULATIONS**  
**SIMULATED WITH HERITABILITY VALUES OF 0.051 ( $\sigma_a^2=0.36$ ,  $\sigma_e^2=6.76$ ).**  
**FROM BASE FEMALES ONLY THE RECORDS OVER AN ESTABLISHED**  
**VALUE ( $\omega$ ) WERE KNOWN**

*Estimas de los parámetros obtenidas mediante distintos modelos en las poblaciones simuladas con un valor de heredabilidad de 0,051 ( $\sigma_a^2=0,36$ ,  $\sigma_e^2=6,76$ ). De las hembras base sólo se conocían los datos que superaban un determinado valor ( $\omega$ )*

Truncation point*	Model**	$\delta$	$h^2$	se( $h^2$ )	$\sigma_a^2$	se( $\sigma_a^2$ )	$\sigma_e^2$	se( $\sigma_e^2$ )
$\omega=9$ ( $\nu=0$ )	M1	—	0.189	0.006	1.370	0.047	5.834	0.045
	M2	—	0.041	0.004	0.271	0.026	6.339	0.041
	M3	—	0.037	0.004	0.241	0.024	6.274	0.047
	M4	$10^8$	0.442	0.012	3.807	0.141	4.520	0.068
	M5	$10^8$	0.454	0.013	3.963	0.151	4.385	0.072
	M6	0.97	0.047	0.004	0.308	0.028	6.184	0.045
	M7	0.97	0.048	0.004	0.316	0.028	6.301	0.040
$\omega=12$ ( $\nu=1.12$ )	M1	—	0.161	0.009	1.278	0.078	6.424	0.072
	M2	—	0.046	0.005	0.324	0.033	6.719	0.047
	M3	—	0.045	0.004	0.312	0.031	6.617	0.049
	M4	$10^8$	0.211	0.014	1.692	0.126	5.431	0.076
	M5	$10^8$	0.194	0.014	1.548	0.127	5.198	0.084
	M6	0.96	0.052	0.005	0.366	0.033	6.605	0.046
	M7	0.96	0.052	0.005	0.364	0.036	6.644	0.049

Abbreviations:  $\delta$ , proportion of genetic variance after selection;  $h^2$ , heritability;  $\sigma_a^2$ , additive genetic variance;  $\sigma_e^2$ , residual variance; se, empirical standard error.

\*  $\omega$ : observed value;  $\nu$ : standard value.

\*\* See Table 2.

Abreviaciones:  $\delta$ , proporción de varianza genética tras la selección;  $h^2$ , heredabilidad;  $\sigma_a^2$ , varianza genética aditiva;  $\sigma_e^2$ , varianza residual; se, error típico empírico.

\*  $\omega$ : valor observado;  $\nu$ : valor tipificado.

\*\* Ver Tabla 2.



**TABLE 5**  
**PARAMETER ESTIMATES UNDER DIFFERENT MODELS IN POPULATIONS**  
**SIMULATED WITH HERITABILITY VALUES OF 0.5 ( $\sigma_a^2 = \sigma_e^2 = 3.56$ ).**  
**FROM BASE FEMALES ONLY THE RECORDS OVER AN ESTABLISHED**  
**VALUE ( $\omega$ ) WERE KNOWN**

*Estimas de los parámetros obtenidas mediante distintos modelos en las poblaciones simuladas con un valor de heredabilidad de 0.5 ( $\sigma_a^2 = \sigma_e^2 = 3,56$ ). De las hembras base sólo se conocían los datos que superaban un determinado valor ( $\omega$ )*

Truncation point*	Model**	$\delta$	$h^2$	se( $h^2$ )	$\sigma_a^2$	se( $\sigma_a^2$ )	$\sigma_e^2$	se( $\sigma_e^2$ )
$\omega=9$ ( $\nu=0$ )	M1	—	0.521	0.007	3.613	0.067	3.285	0.040
	M2	—	0.418	0.008	2.682	0.062	3.689	0.043
	M3	—	0.397	0.008	2.486	0.058	3.735	0.044
	M4	$10^8$	0.678	0.005	6.151	0.093	2.703	0.036
	M5	$10^8$	0.671	0.004	5.973	0.102	2.871	0.014
	M6	0.68	0.537	0.009	3.756	0.087	3.173	0.048
	M7	0.68	0.488	0.007	3.235	0.058	3.365	0.039
$\omega=12$ ( $\nu=1.12$ )	M1	—	0.592	0.007	4.668	0.073	3.187	0.049
	M2	—	0.464	0.009	3.216	0.075	3.664	0.052
	M3	—	0.448	0.009	3.053	0.072	3.719	0.052
	M4	$10^8$	0.741	0.010	7.145	0.162	2.382	0.068
	M5	$10^8$	0.725	0.010	6.773	0.158	2.453	0.069
	M6	0.59	0.542	0.010	4.069	0.096	3.370	0.056
	M7	0.59	0.548	0.008	4.048	0.079	3.296	0.047

Abbreviations:  $\delta$ , proportion of genetic variance after selection;  $h^2$ , heritability;  $\sigma_a^2$ , additive genetic variance;  $\sigma_e^2$ , residual variance; se, empirical standard error.

\*  $\omega$ : observed value;  $\nu$ : standard value.

\*\* See Table 2.

Abreviaciones:  $\delta$ , proporción de varianza genética tras la selección;  $h^2$ , heredabilidad;  $\sigma_a^2$ , varianza genética aditiva;  $\sigma_e^2$ , varianza residual; se, error típico empírico.

\*  $\omega$ : valor observado;  $\nu$ : valor tipificado.

\*\* Ver Tabla 2.

## FINAL REMARKS

REML estimates of genetic parameters in the animal populations can be significantly affected when data are selectively recorded. Although mixed model equations of Animal Model can easily accommodate different kinds of selective breeding, when data are selectively recorded all different approaches that do not consider the reduction produced in the residual variance offer biased estimates of  $h^2$ . From a practical point of view, when no other selection has been produced an alternative could be to ignore all non random records. Unfortunately, it is not obvious to detect which records are not randomly registered, and in small populations this alternative may not be possible because there are not enough data.

## ACKNOWLEDGEMENTS

We gratefully thank M.Baselga for very useful comments. These studies were partially funded by CICYT (project AGF94-1016) and INIA (project 9084), Spain.

## RESUMEN

### Efecto de la selección de datos sobre las estimaciones de heredabilidad

En este trabajo se analiza el efecto de la selección de datos sobre las estimaciones de heredabilidad. Se estimó el valor de heredabilidad del tamaño de camada en una población porcina en la que los datos correspondientes a las cerdas más viejas eran una muestra seleccionada. Las estimaciones se obtuvieron usando distintos conjuntos de datos derivados de toda la información disponible. Esos conjunto de datos se compararon evaluando su capacidad predictiva. Se vio que las estimaciones de heredabilidad obtenidas utilizando todos los datos disponibles correspondían a valores infraestimados. También se simuló un carácter materno y se generó un conjunto de datos seleccionados eliminando aquellos correspondientes a las hembras sin padres conocidos. Distintos modelos, habitualmente empleados cuando no existe selección de registros, se consideraron para estimar el valor de heredabilidad. Los resultados mostraron que ninguno de esos modelos ofrecía estimaciones insesgadas. Sólo los modelos que tenían en cuenta el efecto de la selección sobre la media residual y la media y varianza genéticas ofrecían estimaciones poco sesgadas. Sin embargo, para poder aplicarlos se debe conocer la selección realizada. El problema de la selección de datos es difícil de abordar cuando se desconoce cual es el proceso de selección que se ha realizado en una población.

**PALABRAS CLAVE:** Heredabilidad  
Selección de datos

## REFERENCES

- ESTANY J., POMAR J., ALFONSO L., BABOT D., NOGUERA J.L., 1992. A recording data system for efficient genetic evaluation in pig breeding. Proc. Satellite symposium on pig management information systems, Madrid. Monografías INIA, 68: 221-228.
- ESTANY J., SORENSEN D., 1995. Estimation of genetic parameters for litter size in Danish Landrace and Yorkshire pigs. *Anim. Sci.* 60(2):315-324.
- ESTANY J., ALFONSO L., BABOT D., NOGUERA J.L., 1995. Genetic evaluation of production and reproduction traits from field pig data in Spain. In: Application of mixed linear models in the prediction of genetic merits in pigs (E. Groeneveld, Ed.). Commission of the European Communities, 26-31
- FOULLEY J.L., QUAAS R., 1995. Heterogeneous variances in Gaussian linear mixed models. *Gen. Sel. Evol.* 27: 211-228.
- GRASER H.U., SMITH S.P., TIER B., 1987. A derivative-free approach for estimating variance components in animal models by restricted maximum likelihood. *J. Anim. Sci.* 64:1362-1370.
- GRONEVELD E., KÓVAC M., WANG T., 1990. PEST, a general purpose BLUP package for multivariate prediction and estimation. Proc. of 4th. World Congress on Genetics Applied to Livestock Production, Edinburgh, XII:488-491.
- HENDERSON C., 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics.* 31: 423-447.
- HENDERSON C., 1984. Applications of linear models in animal breeding. Guelph University Press, Guelph.
- IM S., FERNANDO R.L., GIANOLA D., 1989. Likelihood inferences in animal breeding under selection: a missing-data theory view point. *Gen. Sel. Evol.* 21:399-414.
- MEYER K., 1989. Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm. *Gen. Sel. Evol.* 21:317-340.
- NOGUERA J.L., ESTANY J., BABOT D., ALFONSO L., GALOFR R., POMAR J., 1992. GTEP-IRTA: a pig information system for technical and selection management and research. Proc. Satellite symposium on pig management information systems, Madrid. Monografías INIA, 68: 49-57.
- PIERAMATI C., VAN VLECK L.D., 1993. Effect of genetic groups on estimates of additive genetic variance. *J. Anim. Sci.* 71: 66-70.
- QUAAS R.L., 1988. Additive genetic model with groups and relationships. *J. Dairy Sci.* 71: 1338-1345.

- ROBERTSON A., 1977. The effect of selection on the estimation of genetic parameters. *Z. Tierzüchtg. Züchtgsbiol*, 94:131-135.
- RUIZ J., 1993. Heredabilidad del tamaño de la camada en poblaciones porcinas españolas. Proyecto Fin Carrera. Universitat de Lleida.
- SAS INC., 1987. SAS User's guide: Statistics. SAS Institute Inc., Cary, NC.
- VAN DER WERF J., THOMPSON R., 1992. Variance decomposition in the estimation of genetic variance with selected data. *J. Anim. Sci.*, 70:2975-2985.