Theor Appl Genet (1991) 81:624–628

Selection of sires to reduce sampling variance in the estimates of heritability by half-sib correlation

L. Gomez-Raya, L. R. Schaeffer and E. B. Burnside

Centre for Genetic Improvement of Livestock, Animal and Poultry Science, University of Guelph, Guelph, Ontario N1G 2W1, Canada

Received May 29, 1990; Accepted September 19, 1990
Communicated by D. Van Vleck

Summary. Standard methods to estimate heritability by half-sib correlation are biased if selection has operated in the parental generation. In this paper a simple method to correct for selection of animals used as sires is described. By selection of both the top and the bottom ranking sires, the sampling variances of the corrected estimates of heritability are substantially reduced. Algebraic expressions to predict the sampling variance of the estimates of heritability using selected sires are derived. Theoretical predictions were checked by Monte-Carlo simulation. The results may have application in the design of experiments to estimate heritabilities.

Key words: Heritability – Selection bias – Half-sib analysis – Intraclass correlation coefficient – Experimental design

Introduction

Reliable estimates of genetic variance and heritability are needed to perform genetic evaluations, to predict genetic progress, and to help choose between alternative breeding schemes. In many practical situations, heritability is estimated either by offspring-parent regression or by correlation among sibs. Reeve (1953, 1961) showed that the sampling variance of the regression coefficient is reduced by selecting individuals with extreme high and low phenotypes as parents. Hill (1970) used this result in the optimization of the design of experiments to estimate heritability. Thompson (1976) gave optimum designs for estimation of heritability by combining regression with covariance between sibs by maximum likelihood. In both studies, the criterion was to minimize the sampling variance of the estimates for a given total number of individuals recorded in both parental and progeny generations. Hill and Thompson (1977) gave optimum designs when the number of individuals in parental and progeny generations is restricted to be the same. However, in practice, regression methods have limitations, such as an inability to correct for fixed effects or for sex limited traits. Estimates of heritability by half-sib correlation are known to be biased by selection of animals to be sires (Robertson 1977; Ponzoni and James 1978). Therefore, half-sib correlation estimates cannot be used to design experiments without accounting for selection.

The aims of this paper are: (a) to describe a simple method to correct for selection of sires in estimation of heritability by half-sib correlation; (b) to show that by selecting animals with extreme high and low phenotypes to be sires, the sampling variance of the corrected estimates of heritability by half-sib correlation is substantially decreased; (c) to discuss possible uses of selection of extreme individuals in the design of experiments to estimate heritability.

Theory

Consider a quantitative trait influenced by an infinite number of additive and unlinked loci. Selection is by truncation and population size is assumed to be infinite. Let A be the additive genetic value and $\hat{A}$ its estimate with A and $\hat{A}$ being bivariate normal variables. Now rank males according to $\hat{A}$ and select in the same proportion both the top and the bottom ranking sires. Then mate sires randomly to unselected dams. The component of variance among sires in the selected group is

$\sigma^2_{A_S}=(1/4)\,\sigma^2_A(1+k_s\,r_s^2)$,

where σ^2_A = additive genetic variance in the population, r_s = correlation between A and $\hat{A}$ (assumed to be the same for any sire), $k_s = 2\phi(z)\,z/p$ (Bulmer 1976), z = standard normal deviate, $\phi(z)$ = ordinate at cutoff points for p = total selected proportion of high and low ranking males.

The expected value of the heritability estimate ($\widehat{h^2}*$), using the component of variance between sires obtained by standard methods of analysis of variance (Falconer 1981), is

$$\mathrm{E}[\widehat{h^2}*] = \frac{\sigma_A^2(1+k_s r_s^2)}{(1/4)\,\sigma_A^2(1+k_s r_s^2)+\sigma_E^2}, \tag{1}$$

where σ_E^2 = component of error in the analysis of variance. If the selection criterion, $\hat{A}$, is the individual phenotype then $r_s^2 = h^2$ and expression (1) reduces to

$$\mathrm{E}[\widehat{h^2}*] = \frac{h^2(1+k_s h^2)}{1+(1/4)\,k_s h^4}, \tag{2}$$

where h^2 is the true heritability in the population. This expression was first given by Robertson (1977).

Corrected estimate of heritability ($\widehat{h^2}$) can be obtained by solving (with standard quadratic techniques) expression (2) for h^2 in terms of k_s and $\mathrm{E}[\widehat{h^2}*]$ and by substituting $\widehat{h^2}$ for h^2 and $\widehat{h^2}*$ for $\mathrm{E}[\widehat{h^2}*]$ in the resulting equation. This leads to

$$\widehat{h^2} = \frac{1-\beta}{2\alpha\, k_s}, \tag{3}$$

where $\alpha = (0.25\,\widehat{h^2}*) - 1$ and $\beta = [1-(4k_s\,\widehat{h^2}*\,\alpha)]^{1/2}$.

The approximate sampling variance of $\widehat{h^2}$ can be obtained by differentiating expression (3) with respect to $\widehat{h^2}*$:

$$\begin{aligned}\mathrm{Var}(\widehat{h^2}) &\approx \left[\frac{d(\widehat{h^2})}{d\widehat{h^2}*}\right]^2 \mathrm{Var}(\widehat{h^2}*) \qquad (4)\\ &\approx \left[\frac{-[2\alpha\, k_s(\widehat{h^2}*-2)]-(0.5)\,\beta(1-\beta)}{4k_s\,\alpha^2\,\beta}\right]^2 \mathrm{Var}(\widehat{h^2}*)\\ &\approx f\ \mathrm{Var}(\widehat{h^2}*).\end{aligned}$$

The first term (f) is less than one under selection of extreme high and low ranking individuals. The second term [$\mathrm{Var}(\widehat{h^2}*)$] depends on the intraclass correlation coefficient ($\hat{t}$) in the selected group of sires. Making use of Fisher's formula (1958, p. 220) of the variance of the intraclass correlation coefficient, the approximate sampling variance of the biased estimate of heritability becomes:

$$\begin{aligned}\mathrm{Var}(\widehat{h^2}*) &= 16\ \mathrm{Var}(\hat{t})\\ &\approx 16\,\frac{2[1+(n-1)(1/4)\,\widehat{h^2}*]^2(1-(1/4)\,\widehat{h^2}*)^2}{n(n-1)(N-1)},\end{aligned} \tag{5}$$

where N is the number of sires (families) and n is the number of half-sibs per sire. The effect of selection of extreme sires is that the intraclass correlation coefficient is increased. To illustrate the effect of selection of extreme sires on the sampling variance of estimates of heritability, values of $\mathrm{Var}(\widehat{h^2})$ were plotted against the total number of individuals in the progeny generation ($T = n\,N$) for varying heritabilities in the base population 0.5 (Fig. 1), 0.3 (Fig. 2), and 0.1 (Fig. 3). A design to minimize the sampling variance of $\hat{t}$ was used. Following Robertson (1959), the sampling variance of $\hat{t}$ is minimum when $n\hat{t} = 1$, approximately. If we consider a fixed total number of individuals in the progeny generation, then the optimum number of sires to be used is $N = T/n = 0.25\ T\,\widehat{h^2}*$. Four different proportions of selected sires 1%, 20%, 50%, and no selection were considered. Efficiency of selection of extreme high and low sires to reduce the sampling variance in estimates of heritability is illustrated in Figs. 1–3. The effect appears to be more important as sires are more intensely selected and total number of individuals in the progeny generation is small. Reduction in $\mathrm{Var}(\widehat{h^2})$ also increases with heritability. However, possible uses in the design of experiments must also consider the allocation of resources in the parental generation. For example, a very large number of individuals to be sires have to be measured for intense selection (e.g., 1%).

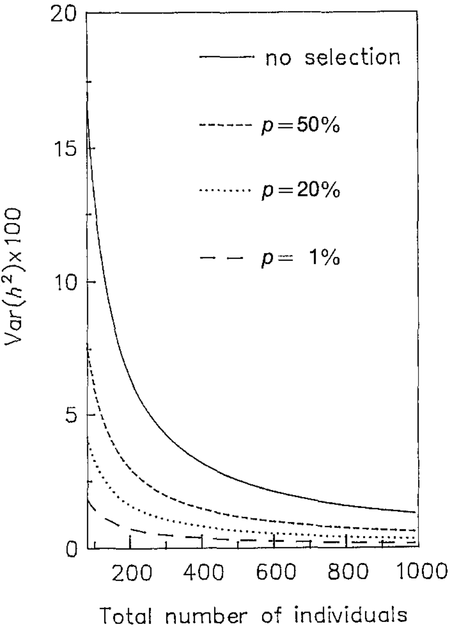

Fig. 1. Approximate sampling variance of estimates of heritability $\mathrm{Var}(\widehat{h^2})$ by half-sib correlation corresponding to total number of individuals in the progeny generation for varying proportions selected of extreme individuals to be sires [p = 50%, 20%, 1%, and no selection (*ns*)]. The family structure minimizes the sampling variance of intraclass correlation coefficient. Heritability in the base population is 0.50

Monte-Carlo simulation

The derivation of the theoretical predictions of the sampling variance of estimates of heritability using selected

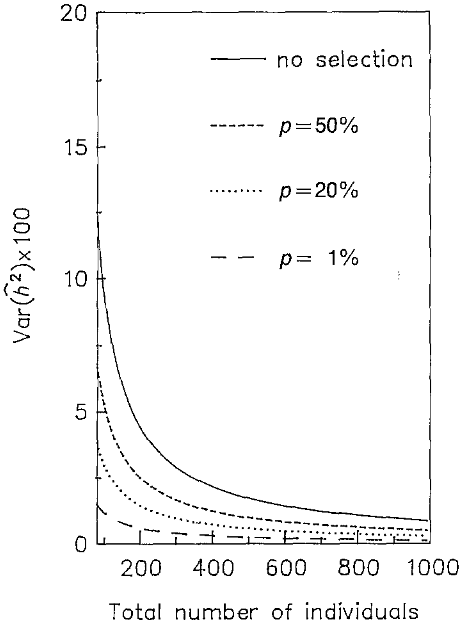

Fig. 2. Approximate sampling variance of estimates of heritability Var($\hat{h}^2$) by half-sib correlation corresponding to total number of individuals in the progeny generation for varying proportions selected of extreme individuals to be sires [$p=50\%$, 20%, 1%, and no selection (*ns*)]. The family structure minimizes the sampling variance of intraclass correlation coefficient. Heritability in the base population is 0.30

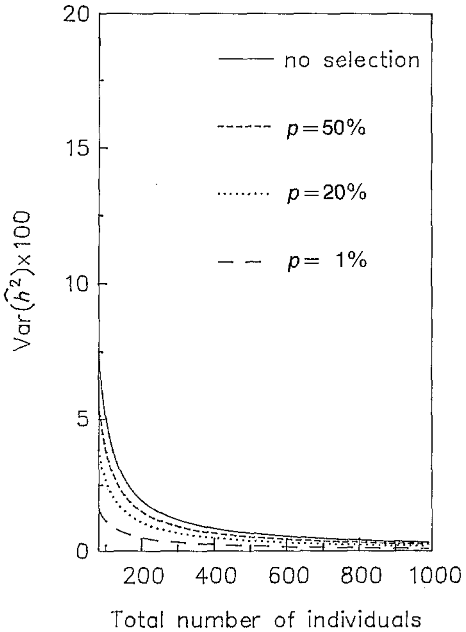

Fig. 3. Approximate sampling variance of estimates of heritability Var($\hat{h}^2$) by half-sib correlation corresponding to total number of individuals in the progeny generation for varying proportions selected of extreme individuals to be sires [$p=50\%$, 20%, 1%, and no selection (*ns*)]. The family structure minimizes the sampling variance of intraclass correlation coefficient. Heritability in the base population is 0.10

sires involves several approximations (derivation of expressions (4) and (5)). A Monte-Carlo simulation experiment was conducted to check the validity of the theoretical predictions. A population of 10,000 dams mated to 100 sires was considered. Sires were selected on their phenotypes at four intensities; 100 out of 200 ($p=50\%$), 100 out of 500 ($p=20\%$), 100 out of 10,000 ($p=1\%$), and 100 out of 100 (no selection). Each sire was mated randomly to 100 dams to produce one offspring from each mating. Heritability in the base population was assumed to be 0.10, 0.30, or 0.50. Genetic variance was fixed at 50 and environmental variance was varied to give each heritability value.

Genetic (G) and phenotypic (P) values for individuals in the parental generation were generated by:

$$G = v_1 (50)^{1/2}$$

$$P = G + v_2 (\mathrm{VE})^{1/2},$$

where v_1 and v_2 are normal deviates (0, 1) obtained from a pseudorandom number generator and VE is the environmental variance with $\mathrm{VE} = 50\,[(1-h^2)/h^2]$.

After selecting extreme ranking sires and random mating to the dams, phenotypic values of offspring were generated by:

$$0 = 0.5\ G_s + 0.5\ G_d + 0.5\ v_3 (50)^{1/2} + 0.5\ v_4 (50)^{1/2} + v_5 (\mathrm{VE})^{1/2},$$

where G_s and G_d are the genotypic values for the sire and dam, respectively, and v_3, v_4, and v_5 are normal deviates (0, 1).

To obtain the necessary components of variance to estimate heritability a minimum variance quadratic unbiased estimator (MIVQUE) was used. Given that the designs were completely balanced and without fixed effects other than the mean, standard methods of analysis of variance by half-sibs (Falconer 1981) would yield the exact same results. Each simulation set was replicated 100 times. Sampling variances of biased ($\hat{h}^{2*}$) and corrected ($\hat{h}^2$) estimates of heritability were obtained from the empirical variances among replicates for each simulation set. Mean squared error of $\hat{h}^2$ over replicates (MSE) was also computed. The reduction in variance (f) was computed as $f = \mathrm{Var}(\hat{h}^2)/\mathrm{Var}(\hat{h}^{2*})$.

Results of the simulation are presented in Table 1. In all cases, correction for selection bias was very efficient giving close agreement between the observed mean squared error (MSE) and the variance of the corrected estimates of heritability [Var($\hat{h}^2$)]. In the same way, expected and observed values of estimates of heritability were similar in both the biased ($\hat{h}^{2*}$) and corrected ($\hat{h}^2$) cases. Observed values of the reduction factor of the sampling variances (f) were also similar to expected values according to theoretical predictions. However, the expected sampling variances of the biased estimates of her-

Table 1. Expected (Exp) and observed (Obs) values in the simulation experiment of the biased ($\hat{h}^{2*}$), corrected ($\hat{h}^2$) estimates of heritability, factor of reduction of sampling variance (f), sampling variances of heritability of the biased [Var($\hat{h}^{2*}$)] and corrected [Var($\hat{h}^2$)] estimates of heritability, and mean squared error (MSE) for varying proportions of selected (p) individuals with high and low phenotypes to be sires 1% (1), 20% (20), 50% (50), and no selection (ns). Each of 100 sires was mated randomly to 100 dams to produce one offspring from each mating. Number of replicates was 100 for each simulation set

p	$\hat{h}^2$		$\hat{h}^{2*}$		f		Var($\hat{h}^{2*}$)[a]		Var($\hat{h}^2$)[a]		MSE[a]
	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs	
1	0.500	0.496	1.611	1.595	0.0724	0.0718	194.54	52.82	14.09	3.79	3.95
20	0.500	0.496	0.931	0.923	0.1751	0.1763	111.15	69.84	19.47	12.31	12.32
50	0.500	0.496	0.678	0.672	0.3787	0.3820	71.20	42.43	26.96	16.21	16.19
ns	0.500	0.491	–	–	–	–	–	–	44.72	49.16	–
1	0.300	0.298	0.830	0.822	0.0662	0.0666	95.23	39.18	6.31	2.61	2.63
20	0.300	0.295	0.478	0.468	0.2305	0.2331	41.70	25.61	9.61	5.97	6.18
50	0.300	0.302	0.370	0.373	0.4829	0.4791	27.75	21.81	13.40	10.45	10.36
ns	0.300	0.302	–	–	–	–	–	–	19.83	23.45	–
1	0.100	0.101	0.171	0.173	0.1766	0.1718	8.20	6.87	1.45	1.18	1.17
20	0.100	0.099	0.122	0.120	0.4903	0.4918	4.95	5.47	2.43	2.69	2.69
50	0.100	0.098	0.108	0.106	0.7376	0.7374	4.19	3.96	3.09	2.92	2.95
ns	0.100	0.099	–	–	–	–	–	–	3.75	5.70	–

[a] Multiplied by 10^4

itability [Var($\hat{h}^{2*}$)] were overestimated in all cases but one. The approximation for the sampling variance of the intraclass correlation coefficient may not be suitable. This effect appears to be more important at higher values of heritability.

Discussion

The estimation of heritability by half-sib correlation methods is biased if progeny result from selected sires. Correction for bias is possible if the truncation points are known. Selection of animals from the high and low extremes to be sires can be used to reduce the magnitude of the sampling variance of estimates of heritability by half-sib correlation. The procedure described in this paper to obtain unbiased estimates of heritability has been shown to be very efficient. The theoretical approximations overestimated the sampling variance of the estimates of heritability, indicating that the use of selected sires could reduce even more than expected the magnitude of the sampling variance of the estimates of heritability.

Reeve (1953) described how the sampling variance of the regression estimates of heritability with or without assortative mating would be reduced by using individuals at both extremes of the distribution of phenotypic values. Other studies showed that selection of extreme individuals in the design of experiments would reduce the sampling variance of estimates of heritability by offspring-parent regression (Hill 1970; Hill and Thompson 1977). Experiments to estimate heritability by half-sib correlation could be designed following the same principles used in those studies, but correction for the bias induced by selection should be incorporated.

The results apply when selection is carried out in one generation and the offspring are recorded in the following generation. In multigenerational experiments, mixed model methods appear unbiased by selection of animals to be parents in simulation studies (Sorensen and Kennedy 1984) provided the model is correct. This methodology may be applied if the objective is to estimate heritability in experiments with farm animals allowing for unbalanced designs, fixed effects, and multiple generations. Then, design of experiments to estimate heritability with the smallest sampling variance for given limited resources becomes very complex. A practical approach would be to select, in each generation, the extreme high and low ranking sires as intensely as possible depending on the fertility of the species and/or the size of the facilities available (assuming that the trait on which selection is based is not negatively correlated with fertility traits). An analysis to estimate components of variance using mixed model methodology (Sorensen and Kennedy 1984, 1986) should follow. Simulation work should confirm that unbiased estimates of heritability were obtained with small sampling variances in populations under continued cycles of selection of extreme ranking sires. Additionally, the use of selection could also be applied in reducing the sampling variance of estimates of heritability by full-sib correlations.

Finally, Robertson (1977) and Villanueva and Kennedy (1990) have given algebraic expressions to predict changes in genetic and phenotypic correlations induced by one cycle of selection and at equilibrium. Estimates of genetic correlations are usually subject to larger sampling variance than heritability estimates. Correcting for changes of genetic variances and covariances induced by selection of extreme individuals may yield estimates of genetic correlations with smaller sampling variance for use in multiple trait selection programmes.

Acknowledgements. We gratefully thank B. Villanueva and Dr. C. Smith for their many constructive suggestions. This research was supported by a grant from Instituto Nacional de Investigaciohes Agrarias, Spain.

References

Bulmer MG (1976) The effect of selection on genetic variability: a simulation study. Genet Res 28:101–117

Falconer DS (1981) Introduction to quantitative genetics, 2nd edn. Longman, London

Fisher RA (1958) Statistical methods for research workers, 13th edn. Oliver and Boyd, Edinburgh

Hill WG (1970) Design of experiments to estimate heritability by regression of offspring on selected parents. Biometrics 26: 566–571

Hill WG, Thompson R (1977) Design of experiments to estimate offspring-parent regression using selected parents. Anim Prod 24:163–168

Ponzoni RW, James JW (1978) Possible biases in heritability estimates from intraclass correlation. Theor Appl Genet 53:25–27

Reeve ECR (1953) Studies in quantitative inheritance. III. Heritability and genetic correlation in progeny tests using different mating systems. J Genet 51:520–542

Reeve ECR (1961) A note on non-random mating in progeny tests. Genet Res 2:195–203

Robertson A (1959) Experimental design in the evaluation of genetic parameters. Biometrics 15:219–226

Robertson A (1977) The effect of selection on the estimation of genetics parameters. Z Tierz Zuechtungsbiol 94:131–135

Sorensen DA, Kennedy BW (1984) Estimation of genetic variances from unselected and selected populations. J Anim Sci 59:1213–1223

Sorensen DA, Kennedy BW (1986) Analysis of selection experiments using mixed model methodology. J Anim Sci 63:245–258

Thompson R (1976) Design of experiments to estimate heritability when observations are available on parents and offspring. Biometrics 32:283–304

Villanueva B, Kennedy BW (1990) Effect of selection on genetic parameters of correlated traits. Theor Appl Genet 80:746–752