



Effect of non-normality and inadmissible estimates on estimation of heritability

AMRIT KUMAR PAUL¹, RANJIT KUMAR PAUL² and WASI ALAM³

Indian Agricultural Statistics Research Institute, New Delhi 110 012 India

Received: 4 May 2013; Accepted: 15 August 2013

Key words: Heritability, Inadmissible estimates

Heritability is one of the most important genetic parameters widely used in plant and animal breeding genetic improvement studies. In literature, several methodologies are available for estimation of heritability for different experimental situations. Unfortunately none of these provide always a valid estimate of heritability and the estimate is so inadmissible that no conclusions can be drawn for the inheritance of the trait under consideration. Further in particular, there is no unique methodology, which is suitable for estimation of heritability in unbalanced situations. Keeping this in view, need has been felt that with the help of computer, the sensitivity and robustness of the very widely used genetic parameter, heritability might be studied at length. The sensitivity here refers to that how the estimate of heritability is dependent upon the aberrant or outliers. The paper contains some of the results as obtained by Bhatia *et al.* (2003).

Methodology: As the heritability is a function of variance components so the following 3 methods are used for variance components estimation and subsequently the estimate of heritability was obtained from the estimates of variance components: ANOVA (Henderson 1963); REML (Patterson and Thompson 1971), and ML (Harville 1977).

Following Bhatia *et al.* (2003), in the present study, the parametric value of heritability was considered as 0.1, 0.25, 0.5 and 0.75. Computer programme for analysis were done in SAS-IML(1995).

Effect of inadmissible estimates: In order to gain information about the effect of inclusion of inadmissible estimates of heritability, different data sets were simulated with various parametric values. Different samples were generated for different parametric values of heritability and for these samples estimates of additive variance components were obtained by using ANOVA, ML and REML methods. and subsequently the estimates of the heritability were worked out. The results for both balanced and unbalanced situation with inclusion and rejection of inadmissible

estimates of h^2 are given in the Table 1, which showed that for both balanced and unbalanced situation, the results obtained are more or less similar for the cases of inclusion and rejection of inadmissible estimates of heritability. Since the occurrence of inadmissible estimates is due to inherent biology of the experimental material and the layout of the design, it is, therefore, not desirable to discard these type of estimates but in fact they must be included in arriving at the final estimate. Non-inclusion will certainly violate the statistical assumption and will also induce a bias in the

Table 1. Estimates of heritability obtained by methods in balanced and unbalanced situations

Situation	Parametric value	Methods	With inadmissible estimates
Balanced	0.1	ANOVA	0.0750 (0.1665)
		REML	0.1099 (0.1665)
		ML	0.0878 (0.1488)
	0.25	ANOVA	0.2201 (0.2509)
		REML	0.2347 (0.2729)
		ML	0.2011 (0.2483)
	0.50	ANOVA	0.4638 (0.4876)
		REML	0.4681 (0.4876)
		ML	0.4217 (0.4486)
	0.75	ANOVA	0.7077 (0.7156)
		REML	0.7085 (0.7156)
		ML	0.6525 (0.6658)
Unbalanced	0.1	ANOVA	0.0795 (0.1661)
		REML	0.1192 (0.1661)
		ML	0.0879 (0.1418)
	0.25	ANOVA	0.2356 (0.2873)
		REML	0.2468 (0.2773)
		ML	0.2102 (0.2532)
	0.50	ANOVA	0.4809 (0.5070)
		REML	0.4900 (0.5051)
		ML	0.4392 (0.4623)
	0.75	ANOVA	0.7340 (0.7471)
		REML	0.7349 (0.7348)
		ML	0.6749 (0.6817)

*Figures in the brackets indicate estimates of heritability obtained only for the case of admissible estimates.

Present address: ¹Senior Scientist (pal@iasri.res.in), ^{2,3}Scientist (ranjitstat@iasri.res.in, wasi@iasri.res.in).

Table 2. Estimates of heritability obtained by different methods from contaminated populations

Proportion of contamination	Parametric value	Weighted average	Methods	With inadmissible estimates
0.8, 0.2	0.1, 0.75	0.23	ANOVA	0.1519 (0.2243)
			REML	0.1797 (0.2174)
			ML	0.1507 (0.2093)
	0.1, 0.5	0.18	ANOVA	0.1304 (0.1136)
			REML	0.1549 (0.2093)
			ML	0.1278 (0.1853)
	0.25, 0.1	0.22	ANOVA	0.1866 (0.2523)
			REML	0.2046 (0.2589)
			ML	0.1735 (0.2256)
	0.25, 0.5	0.30	ANOVA	0.2626 (0.3049)
			REML	0.2745 (0.3049)
			ML	0.2377 (0.2764)
	0.25, 0.75	0.35	ANOVA	0.2996 (0.3401)
			REML	0.3095 (0.3401)
			ML	0.2707 (0.3042)
	0.50, 0.75	0.55	ANOVA	0.5091 (0.5282)
			REML	0.5124 (0.5282)
			ML	0.463 9(0.4883)
0.8, 0.2	0.75, 0.1	0.62	ANOVA	0.5468 (0.5665)
			REML	0.5495 (0.5665)
			ML	0.4995 (0.5203)
	0.5, 0.1	0.42	ANOVA	0.3672 (0.3984)
			REML	0.3745 (0.3984)
			ML	0.3323 (0.3652)
	0.1, 0.25	0.13	ANOVA	0.0989 (0.1777)
			REML	0.1287 (0.1777)
			ML	0.1043 (0.1630)
	0.5, 0.25	0.45	ANOVA	0.4106 (0.4383)
			REML	0.4173 (0.4383)
			ML	0.3722 (0.4002)
	0.75, 0.25	0.65	ANOVA	0.5969 (0.6171)
			REML	0.5989 (0.6171)
			ML	0.5467 (0.5636)
	0.75, 0.50	0.70	ANOVA	0.6573 (0.6719)
			REML	0.6584 (0.6719)
			ML	0.6042 (0.6165)

*Figures in the brackets indicate estimates of heritability obtained only for the case of admissible estimates.

estimate of heritability. Because of this reason, it is concluded that in the further analysis of the data, the discarding of the inadmissible estimates should not be carried out.

Effect of non normality and mixture of two populations: For the methods of ML and REML, we generally assume that the records along with other effects are distributed normally but this assumption may not hold true always and the resultant distribution of the observations may be non-normal. Keeping this in mind, various non-normal populations were simulated by mixing 2 populations with the different proportions and having different parameters. The resulting mixture population may thus be non-normal

Table 3. Effect of contaminated population on the estimates of heritability with mixture of two populations with heritability 0.10 and 0.25 having different proportion of contamination and with different location parametric value

Proportion of contamination	Parametric values	Weighted average	Location parameters	Methods	With inadmissible estimates
0.8, 0.2	0.10, 0.25	0.22	0, 10	ANOVA	0.0040 (0.1437)
				REML	0.0661 (0.1467)
				ML	0.0516 (0.1341)
0.8, 0.2	0.10, 0.25	0.22	10, 0	ANOVA	-0.0097 (0.1320)
				REML	0.0594 (0.1320)
				ML	0.0453 (0.1191)
0.2, 0.8	0.10, 0.25	0.22	0, 10	ANOVA	0.0093 (0.1451)
				REML	0.0683 (0.1452)
				ML	0.0534 (0.1302)
0.2, 0.8	0.10, 0.25	0.22	10, 0	ANOVA	-0.0031 (0.1360)
				REML	0.0639 (0.1360)
				ML	0.0490 (0.1256)
0.8, 0.2	0.25, 0.75	0.35	10, 0	ANOVA	0.0055 (0.1413)
				REML	0.0692 (0.1413)
				ML	0.0530 (0.1205)
0.8, 0.2	0.25, 0.75	0.35	0, 10	ANOVA	0.0162 (0.1407)
				REML	0.0718 (0.1407)
				ML	0.0564 (0.1374)
0.2, 0.8	0.75, 0.25	0.35	10, 0	ANOVA	0.0309 (0.1648)
				REML	0.0857 (0.1648)
				ML	0.0677 (0.1410)
0.2, 0.8	0.75, 0.25	0.35	0, 10	ANOVA	0.0388 (0.1594)
				REML	0.0845 (0.1594)
				ML	0.0670 (0.1456)

*Figures in the brackets indicate estimates of heritability obtained only for the case of admissible estimates.

and these were simulated for both the situation of zero location and non-zero location parameters. Here also, initially the effect of inclusion and discarding of the inadmissible estimates were examined and the results are presented in Table 2. From the Table 2 it is once again reiterated that

estimates are more close to true parametric values (weighted average) and one should consider the final estimate by including all the estimates whether admissible or inadmissible. Though the resultant populations are non-normal so strictly theoretically speaking REML and ML are not applicable but just to get inside, the empirical estimates have been obtained. For studying the role of location parameter also, 2 populations with different location as well as other parameters were mixed in different proportions and resultant non-normal populations were simulated and the results thus obtained are presented in Table 3. From the Table 3 it is seen that in none of the case, the estimate is close to the parametric value, which thus advocates that the methodology for the mixture of the 2 populations need to be re-examined.

SUMMARY

For studying the effect of inclusion of inadmissible estimates and as well as effect of normality on estimate of heritability, different data sets were simulated with various parameteric values. Different samples were generated for

different parametric values of heritability and for these samples estimates of variance components were obtained and subsequently the estimates of the heritability were worked out by using ANOVA, ML, REML and MIVQUE methods. It was found that REML estimates were more close to true parametric values and one should consider the final estimate to including all the estimates whether admissible or inadmissible.

REFERENCES

- Bhatia V K and Paul A K. 2003. *On Robust Estimation of Heritability*. Unpublished Project Report, IASRI, New Delhi.
- Harville D A 1977. Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association* **72**: 320–40.
- Henderson C R. 1963. Selection index and expected genetic advance. *Statistical Genetics and Plant Breeding*. National Research Council Publication 98 L. Pp. 141–63.
- National Academy of Sciences. Patterson H D and Thompson R 1971. Recovery of inter block information when block sizes are unequal. *Biometrika* **58**: 545–54
- SAS-IML. 1995. *SAS documentation*. SAS institute Inc. Cary, NC, USA