Validation of Genetic Evaluation Methodology Using the Nonparametric Bootstrap Method

Joel Ira Weller* Ulf Emanuelson,† and Ephraim Ezra‡
**Institute of Animal Sciences, ARO, The Volcani Center, Bet Dagan, Israel*
†Interbull Centre, Uppsala, Sweden
‡Israel Cattle Breeders' Association, Caesaria, Israel

Abstract

Interbull currently uses three methods to validate genetic evaluation methodologies of member countries. In Method 3, the current genetic evaluation of each bull is analyzed as a function of its genetic evaluation four years ago. The δ factor is a function of the number of new daughters per bull during the past four years. If the evaluations are unbiased, the expectation of δ is zero. The official criterion for Method 3 acceptance is: $|\delta| < 0.01$ (0.02) genetic SD if transmitting abilities (breeding values) are analyzed. Since the standard error of δ is a function of the population size and the number of bulls returned to service, the current criterion penalizes small populations. We therefore propose a new criterion based on the empirical confidence interval for δ computed by the nonparametric bootstrap. This criterion should more correctly reflect validity of the evaluation method for both small and large populations. The current and proposed criteria were evaluated on the Israeli Holstein population, and four anonymous populations. In many cases the current criterion would reject even though δ was not significantly different from zero.

Keywords: Genetic evaluation; validation; dairy cattle; nonparametric bootstrap

---

1. Introduction

Boichard *et al.* (1995) presented three methods to validate genetic evaluations, based on functions of observed vs. expected genetic trends. In all three cases, parameter estimates derived from the sire evaluations are compared to expected values, based on the null hypothesis that the genetic evaluations are unbiased. Analytical methods to compute the standard errors were not presented, although for all three methods the standard errors will be functions of the sample size, *i. e.*, the number of bulls included in the analysis. In the Interbull guidelines for analysis method validations (http://www-interbull.lu.se/service_documentation/General/framesida-general.htm), tolerance values for the three methods are given relative to the population genetic standard deviation, without consideration of the sample sizes, or the standard errors of the estimates for a given population.

The question of the appropriate tolerance value is especially acute for Method 3, in which the current genetic evaluation of bull i, Y_i, is analyzed as a function of its genetic evaluation four years ago, X_i. The Method 3 analysis model is:

$$Y_i = a + bX_i + \delta t_i + e_i \qquad [1]$$

The δ factor in the analysis equation is a function of the number of new daughters per bull during the four years previous to the current genetic evaluation (t_i). If the evaluations are unbiased, the expectation of δ should be zero. However, δ is determined only by those bulls that are returned to general service, which may be a small fraction of the AI bulls with genetic evaluations. The official criterion for Method 3 acceptance is: $|\delta| < 0.01$ (0.02) genetic SD if transmitting values (breeding values) are analyzed.

Since the standard error of δ is a function of the number of bulls returned to service, the current criterion penalizes small populations. We therefore propose that a criterion based on the confidence interval for δ, would more correctly reflect validity of the sire evaluation method for both small and large populations.

If the residuals of equation [1] are normally and independently distributed, then analytical standard errors for both b and δ can be derived from the inverse of the coefficient matrix, and these values can be used to compute confidence intervals (CI) for both statistics. However, this is not the case, because of the highly asymmetric distribution of t_i. The objective of the current study was to derive empirical 95% CI for both parameters using the nonparametric bootstrap (Efron and Tibshirani, 1993), and to compare these CI to the 95% CI derived from the analytically computed SE.

2. Material and Methods

The genetic evaluations of the Israeli Holstein population for milk, fat, and protein production were analyzed. Predicted differences were computed for 172 bulls by a multitrait animal model including parities 1 through 5. Of these 14 bulls (8.1%) were returned to service (*i. e.*, $t_i > 0$), Lactation records were included only if there were valid production records for all three traits. Thus the numbers of bulls and records per bull are the same for all three traits. Four additional (anonymous) populations of different sizes were also analyzed for milk production.

All five populations were analyzed by equation [1]. The details of the method are given in Boichard *et al.* (1995). Analytical SE were derived by PROC REG of SAS (SAS Institute, 1999). Analytical 95% CI were derived as the parameter estimate ± 2 SE. Empirical 95% CI were derived by generating 1000 bootstrap samples for each trait and population. The CI limits were set at the lower and upper 2.5% of bootstrap samples.

3. Results and Discussion

The results of the Method 3 analysis of the Israeli Holstein population for milk, fat, and protein production are presented in Table 1. None of the three traits meet the current Interbull criterion that $|\delta| < 0.01$ genetic SD. The analytical standard errors, derived under the assumption of normality, are also presented in Table 1.

Table 1. Estimates of B and δ derived from the Israel Holstein population (172 bulls).

		Parameter estimates		Analytical SE	
Trait	0.01*genetic SD	*b*	δ	*b*	δ
Milk	6.36	0.9450	-12.55	0.0167	5.022
Fat	0.248	0.9730	-0.514	0.0169	0.146
Protein	0.162	0.9676	-0.244	0.0199	0.146

The lower and upper bounds for the analytical 95% CI for δ, and the bootstrap 95% CI for both b and δ are given in Table 2. For all three traits, the analytical 95% CI are greater than the Interbull criterion of ± 0.01 genetic SD. As expected for b (both X and Y have near normal distributions), the bootstrap confidence intervals were very similar to the analytical values. Therefore, only the bootstrap CI are presented. The bootstrap CI for δ are about twice as large as the analytical CI, and all the bootstrap CI include zero. Based on the empirical CI, none of the Israeli δ values were significantly different from zero. The ratio: CI/(0.01*GSD) is in the final column of this table. These values range from 4.6 to 7.2.

Table 2. Analytical 95% CI for δ, and empirical bootstrap 95% CI for b and δ for milk, fat, and protein production of the Israel Holstein population.

	Analytical 95% CI (δ)		95% bootstrap CI[a]				Ratio: CI/(0.01*GSD)
			b		δ		
Trait	Lower	Upper	Lower	Upper	Lower	Upper	
Milk	-22.59	-2.51	0.910	0.982	-32.9	12.7	7.17
Fat	-0.806	-0.222	0.945	0.999	-1.05	0.09	4.60
Protein	-0.536	0.048	0.927	1.006	-0.70	0.36	6.54

[a] 1000 bootstrap samples were generated.

The reason that the CI are much greater than the analytical values can also be illustrated with the following example. The relatively high negative values for δ in the Israeli population are due to two specific bulls, 3308 and 3338, both born in 1991. For both of these bulls the evaluations based on the second daughter crop was much lower than the original evaluation. Sire 3338 had only 49 daughters in its X_i evaluation. We ran a second set of regressions with sire 3308 deleted. The δ values were -4.0, -0.27, and -0.06. These values are quite different from the values computed with sire 3338 included, and none of these values are significantly different from zero, using CI derived from the analytical SE.

The Method 3 results for the four anonymous populations are given in Table 3. These populations differ greatly in the number of bulls and the fraction of bulls that were returned to service. As was the case for the Israeli population, the analytical and empirical CI for *b* were quite similar for all four populations. For the criterion of $|\delta| < 0.01$ (0.02) genetic SD if transmitting values (breeding values) are analyzed to have validity, the ratio CI/th should be ≤ 2. Based on the analytical CI for δ, the ratio CI/th < 2 for the three larger populations, but not for population 4. However, using the empirical CI, the ratio CI/th < 2 only for the two larger populations. That is the two smaller populations could be rejected by the current Interbull Method 3 criterion even though the estimate of δ is not significantly different from zero. The difference between the analytical and empirical CI for δ is greatest for population 2, which had the lowest fraction of bulls returned to service.

4. Recommendations

1. The Method 3 criterion should be revised so that the 95% CI for δ should include zero.
2. The CI of b and δ should be estimated by the nonparametric bootstrap method.

Acknowledgements

This study was supported by the Israel Milk Marketing Board.

References

Boichard, D., Bonaiti, B., Barbat, A. & Mattalia, S. 1995. Three methods to validate the estimation of genetic trend for dairy cattle. *J. Dairy Sci. 78,* 431-437.

Efron, B. & Tibshirani, R.J. 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York, NY.

SAS Institute. 1999. *SAS/STAT® User's Guide, Version 8*. SAS Institute, Cary, NC.

Table 3. Method 3 results for milk production of four populations.

Population	th[a] × Genetic SD	Number of bulls	$t_i>0$ [b] (%)	Parameter	Estimate	SE	Analytical			Bootstrap		
							95% CI		Ratio CI/th	95% CI		ratioCI/th
							lower	upper		lower	upper	
1	11.2	3395	46	b	1.014	0.005	1.005	1.023		1.004	1.023	
				δ	-2.087	1.532	-5.151	0.976	0.547	-6.483	2.753	0.825
2	13.8	1971	21	b	0.997	0.002	0.992	1.001		0.992	1.001	
				δ	-15.751	1.726	-19.203	-12.298	0.500	-25.336	-6.020	1.400
3	11.0	134	65	b	0.985	0.017	0.951	1.019		0.958	1.011	
				δ	18.859	4.165	10.530	27.189	1.521	4.406	37.059	2.982
4	7.8	42	86	b	1.024	0.076	0.872	1.176		0.911	1.149	
				δ	-41.706	9.773	-61.252	-22.160	5.025	-57.478	-22.145	4.542

[a] 0.01 or 0.02 genetic SD, depending on whether evaluations are given in units of breeding value or transmitting abilities.

[b] the fraction of bulls with new daughters during the last four years.