Normality and Skewness of Genetic Evaluations

P.M. VanRaden

Animal Improvement Programs Laboratory, Agriculture Research Service,
United States Department of Agriculture, Beltsville, MD, USA 20705-2350

Introduction

Linear models provide ideal evaluations for normally distributed traits. Somatic cell counts are very skewed, but all countries use log transformations to improve statistical properties and linear models for analysis. Longevity and fertility traits are not normally distributed. Non-linear models can provide more accurate evaluations for categorical traits such as calving ease, stillbirth, and clinical mastitis, especially when one category occurs much more frequently than others.

Differences in skewness may reduce accuracy of international EBVs when national EBVs from linear and non-linear models are mixed. When distributions with skewness of 0.5 and 0.0 are mixed, >80% of EBVs above 3 SD come from the skewed distribution, whereas nearly all EBVs below –3 SD come from the normal distribution. Possible solutions to this problem are to transform EBVs or to exchange EBVs on observed scales if underlying scales are not available from more countries. This study examined normality of EBVs for all traits evaluated by Interbull.

Methods

Data examined were from the May 2006 Interbull evaluation. On each country's scale, bull evaluations were included if that country reported progeny in at least 10 herds for the trait. All scales were aligned in the same direction before analysis so that higher evaluations were favorable. Statistics reported are for Holstein populations and not for the separate red and white Holstein scales.

Skewness and kurtosis were obtained for each country after removing birth year trend by linear regression. Skewness and kurtosis should both be close to zero if evaluations are normally distributed, except that some positive kurtosis is expected when evaluations have differing reliabilities (VanRaden *et al.,* 1984). Kurtosis was small and positive for most traits and scales, with larger values (up to 1.85) for calving traits. This study focused on skewness.

Non-linear, Bayesian models can account directly for non-normal distributions. In threshold models, the cumulative normal density function is used to convert EBVs from underlying to observed scales (Harville and Mee, 1984). Underlying BVs are assumed to be normal, but EBVs may be skewed.

Skewness is estimated as a cubic function of data and may be removed by various transformations. Subtraction of the mean and division by the SD often are the first two steps in standardizing records, and these can convert variable y with normal distribution to variable s with a t distribution.

$$s = (y - mean) / SD$$

A third step can remove the skewness if s is not normally distributed. Skewness is estimated as

$$skewness = '\ s_i^3 / (n - 2).$$

Unskewed EBVs (u) may be obtained using the general exponential transformation of Manly (1976):

$$u = (e^{as} - 1) / a$$

where the constant a can be obtained by numerical methods or approximated as skewness / –3. Statistical analyses could then be performed on the unskewed data u and the reverse transformations applied, if desired, to obtain estimates on the original units for y. Such estimates may have lower standard errors but also bias due to the nonlinear transformation. Not all skewed distributions should be normalized, because observations in the tails

may contain added information from repeated sampling (VanRaden, 2003).

International evaluations might be more accurate if all countries reported evaluations on normally distributed scales or more fair if differences in skewness were smaller. Expected responses to selection are not symmetric when EBVs are skewed.

Results

Calving ease evaluations had a range of skewness from −.81 to −.01, as shown in Table 1. Two countries (ISR and USA) reported lower evaluations as desirable, but these scales were reversed before statistics were calculated. Thus, skewness is toward difficult calving, as expected. Countries with threshold models (FRA, ITA, and USA) have more normally distributed evaluations on the underlying scale, which is reported to Interbull. The Snell score evaluations in CAN also have less skewness than several of the linear model evaluations that do not account for one category having most of the observations.

Underlying or Snell score evaluations in CAN, FRA, and USA are converted back to observed scale using non-linear formulas before being reported to breeders. Observed scale evaluations in USA have skewness of −.69 for the direct effect and −.42 for the maternal effect, which is about equal to that seen in other countries. This is in contrast to −.22 and −.03 seen in the more normally distributed underlying solutions reported to Interbull. MACE rankings from linear models already are skewed, and the skewness is doubled when foreign EBVs are transformed from underlying or Snell scores to observed scales in countries that use non-linear formulas domestically.

Stillbirth evaluations had skewness ranging from −.94 to −.41 except for one value of −.06. Stillbirth is a binomial trait with low incidence, and skewness was fairly uniform because all countries currently evaluating this trait used linear models. A threshold model tested on USA data (Cole *et al.,* 2006) for August 2006 implementation results in skewness of only −.10 for direct and −.23 for maternal effects on

underlying scale vs. −.32 and −.69 on observed scales.

Somatic cell score (SCS) evaluations have moderate, consistent skewness in almost all countries. The direction of skewness is toward higher cell count. The original research of Ali and Shook (1980) indicated that use of log transformation greatly reduced phenotypic skewness and kurtosis as compared to use of untransformed somatic cell concentration. The clinical mastitis evaluations shown in Table 2 are based on binomial data but are somewhat more normally distributed than calving ease.

Longevity and fertility evaluations are fairly normally distributed. Linear model evaluations generally had little skewness even though raw data are not normally distributed. Moderate negative skewness occurred in a few countries using survival analysis. Lack of problems may be explained by fairly high incidence rates of culling per lactation and the fact that daughters with more lactations or more inseminations provide more information about the culling rate or conception rate parameters. Thus, log transformation to make such data more normally distributed may not make evaluations more accurate (VanRaden, 2003).

Fertility evaluations had only slight skewness, ranging from −.23 to .11 for heifer conception rate, −.33 to −.09 for days to first insemination, −.20 to .10 for cow conception rate, and −.35 to −.04 for days open or calving interval. Countries are not listed individually because data were from the pilot study of Jorjani (2005, 2006).

Yield trait evaluations were fairly normally distributed in most countries and with a wider range of skewness estimates in countries with fewer bulls, as expected. Tests of non-normality generally were significant if the absolute value of skewness was >.10 in larger countries or >.20 in smaller countries. Skewness ranged from −.42 to .77 for milk, −.23 to .51 for fat, and −.47 to .93 for protein. Differences could be caused by evaluation methods but also by selection. Some countries may have a few domestic bulls far below their population average, resulting in negative skewness, or a few imported bulls of higher merit, resulting in positive skewness.

Table 1. Skewness of breeding values within years for bulls evaluated in each country.

Country	Milk	Fat	Prot	SCS	Long-evity	Direct CE[2]	Maternal CE	Direct SB[3]	Maternal SB
AUS	.10	-.06	.06	-.19	-.09	-.60			
BEL	-.16	-.12	-.47	-.31	.09				
CAN	.20	.03	.20	-.33	.00	-.47	-.34		
CHE	.08	.06	.21	-.30	-.13	-.80	-.30	-.86	-.06
CZE	-.02	.05	-.07	-.81					
DEU	.10	.02	.13	-.16	-.01				
DNK[1]	.02	-.10	.10	-.20	-.03	-.55	-.17	-.54	-.47
ESP	-.02	.00	.03	-.33	-.30				
EST	.30	.51	.54	-.41					
FIN				-.24	.00	-.65	-.67	-.52	-.69
FRA	-.01	.01	-.05	-.23	.06	-.29	-.01		
GBR	-.21	-.23	-.48	-.25	-.06				
HUN	.12	.10	.20	-.22					
IRL	-.42	-.18	-.44	-.10	.06				
ISR	-.01	-.11	.02	-.27	-.06	-.59	-.81	-.52	-.94
ITA	.17	.13	.27	-.17	-.36	-.12	-.16		
JPN	.17	.10	.18	-.25					
NLD	.02	-.06	-.13	-.24	-.15	-.37	-.21	-.48	-.62
NZL	-.16	.01	-.10	-.29	.09				
POL	.77	.45	.93						
SVN	.49	.07	.35						
SWE				-.23	.07	-.69	-.53	-.41	-.63
USA	-.14	-.03	-.16	-.26	.03	-.18	-.08		
ZAF	-.09	-.21	-.19	-.43					

[1]Milk, fat, and protein results are from the Denmark-Finland-Sweden joint evaluation.
[2]CE = calving ease ; [3]SB = stillbirths

Numbers of extreme bulls for calving ease were examined. For direct effect, the United States had 34% of all bulls, 45% of the best 100 bulls, averaged across country scales, and 20% of the worst 100 bulls, even though the mean for U.S. bulls was slightly poorer than for foreign bulls. U.S. bulls were slightly better on average for the maternal effect and supplied 22% of all bulls, 32% of the best 100, and 21% of the poorest 100 bulls. For longevity, Italy had the most negative skewness and had 6.1% of all bulls, but only 1.6% of the best bulls, and 3.9% of the poorest bulls. Thus, skewness of domestic EBVs can affect results of international evaluations.

Table 2. Skewness and kurtosis for clinical mastitis.

| Country | Within birth year | |
	Skewness	Kurtosis
DNK	-.21	.18
FIN	-.24	.49
SWE	-.23	.00

Conclusions

Even when phenotypic data are not normal, EBVs have fairly normal distributions for most traits. Evaluations tend to be more negatively skewed from linear models applied to categorical data (i.e. fewer bulls with extremely good EBVs and more bulls with extremely poor EBVs) than evaluations on the underlying scales of threshold models. Breeders tend to notice extreme bulls for particular traits, and both skewness and kurtosis affect the proportion of extremes. Breeders may suspect problems if most of the best or worst bulls for a trait are from a given country even if the means and standard deviations for each country are as expected.

Differences in skewness may not be a problem when combining different traits into a selection index, and skewed evaluations may be optimal if economic values for categorical traits are linear on the observed rather than the

underlying scale. However, international evaluations would be more useful if properties of national input data were more uniform for each trait. For calving ease and stillbirth, countries could exchange observed scale EBVs until underlying scale EBVs are commonly available. Removal of skewness by transformation or combined analysis of raw data also could provide more uniformity.

References

Ali, A.K.A. & Shook, G.E. 1980. An optimum transformation for somatic cell concentration in milk. *J. Dairy Sci. 63,* 487-490.

Cole, J.B., Wiggans, G.R. & VanRaden, P.M. 2006. Genetic evaluation of stillbirth in United States Holsteins using a sire-maternal grandsire threshold model. *Proc. 8th World Congr. Genet. Appl. Livest. Prod.* (In press).

Harville, D.A. & Mee, R.W. 1984. A mixed-model procedure for analyzing ordered categorical data. *Biometrics 40,* 393-408.

Jorjani, H. 2005. Preliminary report of Interbull pilot study for female fertility traits in Holstein populations. *Interbull Bulletin 33,* 34-44.

Jorjani, H. 2006. International genetic evaluation of fertility traits. *Interbull Bulletin 34,* 57-64.

Manly, B.F.J. 1976. Exponential data transformations. *The Statistician 25,* 37-42.

VanRaden, P.M. 2003. Longevity and fertility trait definitions compared in theory and simulation. *Interbull Bulletin 30,* 43-46.

VanRaden, P.M., Freeman, A.E. & Rothschild, M.F. 1984. Maximizing genetic gain under multiple-stage selection. *J. Dairy Sci. 67,* 1761-1766.