2016

EPDs and Risk

L. Dale Van Vleck
University of Nebraska, dvan-vleck1@unl.edu

From: 2016 Beef Improvement Federation Annual Meeting & Symposium,
June 14 - 17, 2016, Hilton Garden Inn, Manhattan, Kansas. www.BeefImprovement.org

EPDs and Risk

Dale Van Vleck
US Meat Animal Research Center and Department of Animal Science,
University of Nebraska-Lincoln

Introduction

About 30 years ago there was concern in both the beef and dairy industries that too much emphasis was being given to accuracy of genetic evaluation. This article will discuss attempts to reduce emphasis on accuracy and, thus increase emphasis on the predictor of genetic value itself which is commonly known as estimated breeding value (EBV). Accuracy is a key component of more useful measures of risk such as standard error of prediction which can be used to create confidence ranges in units of measurement for true breeding value based on the EBV and the standard error of prediction. The concept of standard error of prediction can be extended to comparison of pairs of EBV. The influence of genomic relationships and Bayesian analyses on accuracies and standard errors of prediction will also be briefly introduced.

Accuracy

Reports of genetic evaluations, in addition to EBV (or EPD = EBV/2 or PTA = EBV/2), provide an item named 'accuracy'. Accuracy is an indicator of risk of possible change in the EBV. Accuracy is defined as the correlation between the EBV and true BV of an animal. High accuracy suggests little possible change and low accuracy suggests considerable possible change when later genetic evaluations are based on many, many more records. Accuracy of an EBV is the same as the accuracy of a corresponding EPD. A more useful measure of risk is the standard error of prediction (SEP) which depends on squared accuracy and the genetic standard deviation of the trait. [The square of the genetic standard deviation is the genetic variance of the trait. Genetic variance (symbol is Vg for this discussion) is the part of the total (phenotypic) variance for a trait that is due to effects of genes of animals.] SEP is in units of how a trait is measured (e.g., pounds) and thus is a quantitative measure of possible change. Accuracy is 'unit-less' and can range from 0.0 to 1.0. SEP can range downward from the genetic standard deviation for accuracy of 0.0 to 0.0 for accuracy of 1.0. As would be expected, the standard error of prediction for an EPD is one-half the standard error of prediction for an EBV.

Reliability

High accuracy may receive too much emphasis relative to the EBV of an animal. To reduce emphasis on traditional accuracy other measures of 'accuracy' have been proposed and reported with EBV. The dairy industry uses a method named 'reliability' which was implemented in about 1989 by Paul VanRaden and others at the Animal Improvement Programs Laboratory of the USDA which for many years did the genetic evaluations for all U. S. dairy breeds. Reliability is simply the square of traditional accuracy and represents the fraction of genetic variance accounted for by the EBV. Squared accuracy (reliability) approaches perfection (1.00) more slowly than traditional accuracy. It too is unit-less. (See Table 1.) For example with accuracy of 0.90, reliability is 0.81. Smaller reliability relative to traditional accuracy reduces the temptation to over emphasize 'accuracy' and thus will increase emphasis on the EBV.

BIF-Accuracy

The beef industry chose a different approach. Richard Willham proposed an expression which was implemented by some beef breeds about 1985. It also is unit-less and approaches 1.00 even more slowly than reliability. It has been named 'BIF-accuracy'. It has been somewhat confusing because of the name and the equation used which is based on traditional accuracy squared and the standardized standard error of prediction as well as a square root:

$$\text{BIF-accuracy} = 1.0 - \text{SQRT}[1.0 - \text{acc}^2].$$

As accuracy increases toward 1.0, BIF-accuracy increases at a rate dependent on SQRT[1.0 – acc^2]. The rate of increase toward 1.0 is much less than the rate for accuracy or the rate for reliability. [See Table 1.] The standard error of prediction (SEP) is SQRT[(1.0 – acc^2)(Vg)]. SQRT[1.0 – acc^2] is standardized SEP (that is, it corresponds to genetic variance, Vg = 1.0). Thus BIF-accuracy basically tracks the approach of SEP to zero.

Progeny accuracy

An approach, never proposed or implemented, to reduce emphasis on accuracy would have been to report traditional accuracy for a future progeny of a sire. Accuracy for a future progeny with no records is one-half the accuracy of the EBV of its sire. The reason for the one-half is that the sire is related by one-half to his progeny. Even if proposed it probably would not have been adopted because 'accuracies' less than 0.50 would, no doubt, have created doubt about validity of corresponding EBV or EPD.

Standard error of prediction

The standard error of prediction (SEP) provides a more direct measure of risk (possible change) or chance that true breeding value is so much greater or so much smaller than the EBV than does any of the 'accuracies'. SEP was introduced in discussion of BIF-accuracy as SQRT[(1.0 – acc^2)(Vg)] where Vg is the genetic variance of the trait. Prediction error is the difference between EBV and true BV:

$$\text{PE} = (\text{EBV} - \text{true BV}).$$

Variance of prediction error, V(PE), which is also often referred to as prediction error variance, PEV is:

$$\text{PEV} = \text{V(PE)} = \text{V(EBV} - \text{true BV)} = (1.0 - \text{acc}^2)(\text{Vg}).$$

A numerical value for PEV rather than SEP (the square root of PEV) comes directly from the statistical method used to obtain EBV and fortunately without having to know any true BV. Prediction error variance (PEV) is in units of measurement squared (for example, lb x lb). The square root of PEV is SEP which is in units of measurement (that is, lb) so that SEP as an indicator of risk is on the actual scale of measurement. The SEP decreases as accuracy increases. [See Table 1.] A property of Henderson's mixed model equations used to obtain EBV is that the PEV's are the diagonal terms of the inverse of the coefficient matrix of those equations. That inverse can be used to solve for EBV but, unfortunately, inverses needed for most breed evaluations are impossible to obtain with current and foreseeable computing power. Iterative

methods are used to solve for EBV. In theory the same iterative methods can be used to obtain individual PEV, but in practice are not feasible. [But see **Bayes.**] Time required would be time for solving for EBV multiplied by the number of animals in the pedigree. Usually practical values of accuracy can be approximated. From approximate accuracy, PEV can be computed. With inverse solutions to the genetic evaluation equations, PEV can be determined directly and then used to obtain traditional accuracy.

The standard error of prediction is a direct measure of possible change. Possible change is risk in units of the trait and thus has dollar value. Risk can be 'positive' or 'negative'. That is, the chance true BV may exceed EBV by a certain amount is the same as the chance true BV is less than EBV by the same amount. The up-side risk (possible gain) is the same as down-side risk (possible loss). Monetary values of up-side and down-side risk are not necessarily equivalent.

Confidence ranges

Confidence ranges are often used to determine probabilities of possible change assuming a bell-shaped distribution of true BV around the EBV. One-half of true BV would be expected to be greater than the EBV and one-half would be expected to be less than the EBV. The interval from EBV – (1)SEP to EBV + (1)SEP corresponds to 68% of possible BV for an animal centered on the EBV for the animal. The range can be shrunk or expanded corresponding to the probability of true BV being in the interval. For example, the interval from EBV – (2)SEP to EBV + (2)SEP would be expected to contain 95% of true BV. Units of SEP other than (1) or (2) would correspond to other confidence ranges.

With a 68% confidence range, 32% would be outside the range: 16% above the positive end of the range and 16% below the negative end of the range. With the 95% range, 2.5% would be expected to be greater than the upper end of the CR and 2.5% would be expected to be less than the lower end of the CR. Ranges for many combinations of EBV and SEP will overlap considerably. The more important of EBV or SEP is the EBV which centers the range. Comparison of ranges provides a more direct measure of risk than does accuracy.

Comparison of pairs of EBV

Selection decisions are essentially based on comparison of the EBV of a pair of animals. The ideas of confidence ranges and possible change can be applied to differences in pairs of EBV. The explanation becomes more complicated but the statistical principles are the same. Now there are two prediction errors:

$$PE_1 = EBV_1 - BV_1 \text{ and } PE_2 = EBV_2 - BV_2.$$

To form confidence ranges, variance of $PE_1 - PE_2$ is needed instead of V(PE). In expanded form $V(PE_1 - PE_2) = V(PE_1) + V(PE_2) - (2)COV(PE_1,PE_2)$. What is new is the covariance between the pair of prediction errors. A covariance is a measure of how two things vary together. EBV of a pair of relatives would be expected to be correlated (positive covariance) because some of the same information would be used in both EBV. Except for close relatives in the same management unit, the covariance is likely to be small relative to $V(PE_1)$ and $V(PE_2)$. [That is my expectation until proved to be different for other than close relatives.] Obtaining a prediction error covariance requires the inverse of the coefficient matrix of the genetic evaluation

equations as do the variances of prediction errors. The potential number of prediction error covariance's is much larger than number of prediction error variances: n(n – 1)/2 where n is the number of animals in the pedigree. Approximations of prediction error covariance's are probably more difficult to obtain than approximations of V(PE). If iteration were used to obtain V(PE) for an animal, a by-product would be the PE covariance's between its EBV and the EBV of all other animals in the pedigree.

If the PE covariance can be safely ignored, the standard error of the difference between a pair of EBV can be calculated easily although accuracies for two EBV must be known or approximated. Then $V(PE_1 - PE_2) = V(PE_1) + V(PE_2) = (1.0 - acc_1^2)Vg + (1.0 - acc_2^2)Vg = (2.0 - acc_1^2 - acc_2^2)Vg$. The square root of $V(PE_1 - PE_2)$ is the standard error of the predicted difference (SEPD) between EBV. [SEPD is not the standard error of an EPD.] Computation of SEPD would not be needed for most pairs of animals. The animals of most interest are potential herd sires. For pairs of interest, SEPD can be obtained using a simple table which would apply to all traits. The table values would be multiplied by the genetic standard deviation for a specific trait. The number of rows and columns of the table would correspond to ascending or descending levels of accuracy; for example, from 0.05 to 0.95 by increments of 0.05. [See Table 2.] Entries in the table would be $SQRT[(2.0 - acc_i^2 - acc_j^2)]$ for the intersection of the i^{th} row and j^{th} column. [Table values corresponding to accuracies between, for example, 0.75 and 0.85 could be obtained by interpolation although interpolation may be of little practical importance.] As an example of the use of the table if accuracy for bull 1 was 0.55 and accuracy for bull 2 was 0.95, the table entry is 1.05 [$SQRT(2.0 - 0.55^2 - 0.95^2) = 1.05$]. The second step in obtaining SEPD is to multiply the 1.05 by the genetic standard deviation of the trait, SQRT[Vg].

If SQRT[625] = 25 is the genetic standard deviation for the trait, $SEPD(EBV_1 - EBV_2) = 1.05 \times 25 = 26.25$. Confidence ranges and possible changes will now correspond to $BV_1 - BV_2$ given $EBV_1 - EBV_2$. The confidence ranges will be centered at $EBV_1 - EBV_2$. Interpretation will be as for SEP, but for $BV_1 - BV_2$ rather than for $BV_1 - 0.0$ or $BV_2 - 0.0$. Still to be determined is whether $COV(PE_i, PE_j)$ can be safely ignored. A relatively small covariance would not change SEPD of much importance. Such covariance's, will be positive and thus would make SEPD calculated not including the covariance smaller than it should be, but how much smaller would be important needs to be investigated. A large covariance (correlation) between PE of EBV of a sire and PE of EBV of his son would be expected especially if one had no progeny with records. A sire and his son would seem unlikely to be compared. Paternal half sibs would be more likely to be candidates for selection. If they had no records or progeny with records their EBV would be equal as would SEP so that SEPD would not be important.

Some of the following speculation, if confirmed, would make some of the preceding discussion irrelevant.

G-BLUP

With G-BLUP (using the genomic relationship matrix, G, rather than the identity by descent relationship matrix, A), all genotyped sires will have the same or nearly the same accuracy and the same SEP and SEPD because the same information is available for all sires (same SNPs). Exceptions are for genotyped sires having many progeny with records. Confidence ranges would differ only by the center value, EBV or $EBV_1 - EBV_2$. Different 'chips' might yield different accuracy. It would seem that covariance's between PE are more likely to be non-

zero with G-BLUP than A-BLUP. Would covariance's be the same for all pairs of genotyped animals? If so, iteration for only one column of the inverse of the coefficient matrix would be needed to obtain prediction error variances and covariance's between all pairs of prediction errors.

Bayes

Bayesian methods may make 'direct' calculation of SEP more feasible. Bayesian 'solutions' are obtained by iteration similar to iteration to obtain traditional solutions. The usual method, Gibb's sampling, produces MCMC chains of 'solutions' for an individual which correspond to a distribution around the true BV. The final 'solution' is usually taken to be the average or median of the chain of solutions after 'burn-in' and thinning. Chains could be obtained holding VC constant, that is, Gauss-Seidel iteration but with sampling for solutions but not variance components. The chains for an individual can be used to calculate something comparable to variance of prediction error from which something comparable to accuracy could be calculated as before. The covariance between pairs of predictors could also be obtained from pairs of chains which would incorporate the covariance between pairs of prediction errors.

Summary

Traditionally accuracy has been defined as the correlation between EBV and true BV and has been used as an indicator of risk of possible change.

Reliability (squared accuracy, acc^2) and BIF accuracy ($1 - SQRT[1 - acc^2]$) both go towards a maximum of 1.0 more slowly with more information than accuracy and were developed to reduce emphasis on 'high' accuracy vs. EBV.

The standard error of prediction is a more quantitative measure of risk than accuracy. It goes toward 0.00 as accuracy increases: $SEP = SQRT[(1.0 - acc^2)(Vg)]$ where Vg is genetic variance of the trait. The SEP can be used to obtain ranges such as EBV – 2(SEP) to EBV + 2(SEP) which would include true BV with confidence of 95%. Of the other 5%, 2.5% would be above the upper end of the range and 2.5% below the lower end of the range.

The concept of SEP can be extended to differences in pairs of EBV. SEPD would be the standard error of the difference between a pair of EBV. Confidence ranges would be centered on $EBV_1 - EBV_2$. If the covariance between pairs of prediction errors, $Covariance(PE_1, PE_2)$, is small relative to $V(PE_1)$ and V(PE2), SEPD can be approximated well by $SQRT[(2.0 - acc_1^2 - acc_2^2)Vg]$. A table of SEPD corresponding to pairs of accuracies can then be used to obtain SEPD for any trait and pair of EBV.

Using the genomic relationship matrix (G-BLUP) rather than the identity by descent relationship matrix is likely to result in nearly equal accuracy for many genotyped animals. Then SEP and SEPD would also be equal. Confidence ranges would also be equal but with different centers depending on EBV.

Variances and covariance's of MC-MC chains from Gibb's sampling could be used to obtain equivalents of variances and covariance's of prediction errors and from those equivalents

accuracies can be obtained without the inverse of the coefficient matrix and without approximations of accuracy.

Table 1. Comparison of accuracy, reliability, BIF-accuracy and standard error of prediction (genetic standard deviation of 25).

Accuracy	Reliability	BIF-accuracy	SEP
0.10	0.01	0.005	24.75
0.20	0.04	0.020	24.00
0.30	0.09	0.046	22.75
0.40	0.16	0.083	21.00
0.50	0.25	0.134	18.75
0.60	0.36	0.200	16.00
0.70	0.49	0.286	12.75
0.80	0.64	0.400	9.00
0.90	0.81	0.564	4.75
1.00	1.00	1.000	0.00

Table 2. Table values corresponding to accuracies for two EBV or two EPD which when multiplied by the genetic standard deviation of the trait result in the standard error of prediction of difference between the two EBV or two EPD.

	Accuracy for first EBV								
Acc-2	0.05	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95
0.05	1.41	1.41	1.41	1.49	1.38	1.35	1.30	1.22	1.09
0.25	1.41	1.41	1.41	1.40	1.38	1.35	1.30	1.21	1.09
0.35	1.41	1.41	1.40	1.39	1.38	1.34	1.29	1.21	1.08
0.45	1.40	1.40	1.39	1.38	1.37	1.33	1.28	1.20	1.07
0.55	1.38	1.38	1.38	1.37	1.35	1.32	1.26	1.18	1.05
0.65	1.35	1.35	1.34	1.33	1.32	1.28	1.23	1.14	1.00
0.75	1.30	1.30	1.29	1.28	1.26	1.23	1.17	1.08	0.93
0.85	1.22	1.21	1.21	1.20	1.18	1.14	1.08	0.98	0.81
0.95	1.09	1.09	1.08	1.07	1.05	1.00	0.93	0.81	0.61