

Fundamentals of Expected Progeny Differences

Darrh Bullock, Extension Professor, University of Kentucky

Beef cattle genetic evaluation is the process of taking all of the relevant information on an animal and converting it into a useful tool for selection. This process was started long ago when livestock producers began to realize that progeny often performed similar to their parents for certain traits. Even though they did not know this phenomenon was due to genetics, as we know today, the practice of selecting superior animals to produce the next generation is the foundation for animal breeding and is the purpose for beef cattle genetic evaluation. The first genetic evaluations were simply based on visual appraisal and progress was usually slow and limited. Through scientific discovery and applied practices beef cattle genetic evaluation has evolved into a sophisticated methodology that incorporates pedigrees, phenotypic data and the potential use of genomic information to provide producers with accurate selection tools for a wide variety of economically important traits. The purpose of this paper is to give insight into the current method of computing Expected Progeny Differences (EPD).

The computation of EPD requires three elements: pedigree, phenotypic data and an estimate of the heritability for the trait. The pedigree is used to determine the relationship of each animal in the data set to the other animals in the data set. This is useful because it is known that closer relatives typically have more genes in common and are therefore more likely to perform similarly for that trait.

The phenotype is simply the measurement of the trait of interest; for example the weaning weight of a calf. The more phenotypic data that is available in a data set improves the estimates of the genetic evaluation. Heritability is the proportion of variation in a trait that can be attributed to additive genetics, or the genetics whose effects get passed from one generation to the next. If a trait is highly heritable then the offspring tend to perform similar to their parents for that trait. If it is lowly heritable then the environment tends to play a larger role and there is little similarity between parents and offspring for that trait. With this knowledge it is easy to see that when computing EPDs using this basic model the best success will occur when you have a data set with complete relationship information, a highly heritable trait and a trait that is easily measured.

The first EPDs to be computed by breed associations were birth weight, weaning weight and yearling weight. These were economically important traits (birth weight by association to calving ease), moderate heritability, relatively easy to measure and most breeds had good pedigree information. Different methodologies and models have been used over the years and many additional traits have been added to the evaluation, but the product that has been seen by producers (EPDs) has stayed basically the same with the improvements being in the reliability of the estimates and which animals could have EPD computed.

In other words, the appearance of EPDs and how to use them has not changed over the past 30 years, they have simply gotten better.

The information that is currently used to compute an EPD is: actual measurement of the animal, measurements of any ancestors (parents, grandparents, etc.), measurements of any collateral relatives (siblings, cousins, etc.), measurements on descendents (offspring, grand offspring, etc.) and measurements on correlated traits. For a particular animal, varying amounts of information may be used to compute its EPDs and a measurement on the animal is not necessarily a requirement. The benefit of adding more information to the computation on an animal's EPD is an increase in accuracy, which will be discussed later in the paper.

The basic model for genetic evaluations is:

$$\text{Phenotype} = \text{Known Effects} + \text{Genetic Merit} + \text{Unexpected Variation}$$

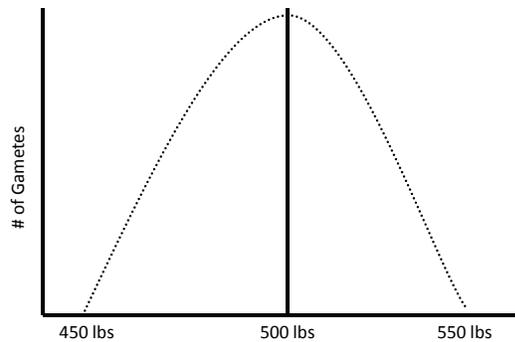
In this equation we know what the phenotype is, because we measured it, and we estimate the Known Effects through our knowledge of contemporary groups and how sires used across multiple contemporary groups perform differently. This leaves us with two unknowns: Genetic Merit, which is ultimately what we are interested in, and Unexpected Variation. Using simple algebra we can attempt to solve this equation to estimate the genetic merit:

$$\text{Phenotype} - \text{Known Effects} = \text{Genetic Merit} + \text{Unexpected Variation}$$

Unfortunately, we cannot completely isolate Genetic Merit and therefore our estimate or EPD will always be compromised. When confronted with this situation the logical approach to achieve the best estimate of Genetic Merit is to try to reduce the Unexpected Variation to as close to zero as possible. This is accomplished by the inclusion of more data; the more data that exists on an animal the greater the reduction in Unexpected Variation. For practical purposes EPDs relate to Genetic Merit and Accuracy relates to Unexplained Variation.

Differences in two bulls EPDs for a particular trait is the best estimate of the average difference we would expect to see in the progeny produced by those two bulls. In effect, EPDs are estimating the average value of the individual gametes that are being produced by an animal, sperm in the case of a bull. Because of random segregation each sperm produced by a bull has the potential to have a unique set of genes; that is why flush mates have different genetic make-up even though they are produced by the same sire and dam. If the actual genetic merit of each individual sperm could be determined from the ejaculate of a bull the results might resemble Chart 1. The bull's genetic potential is for 500 lbs of weaning weight and the largest proportion of his sperm would result in 500 lbs calves. However, the bull has some sperm that would result in a much lighter calf and other sperm that would result in a much heavier calf. In this example, on average his calves would weigh 500 lbs with some weighing as little as 450 lbs and some weighing as much as 550 lbs.

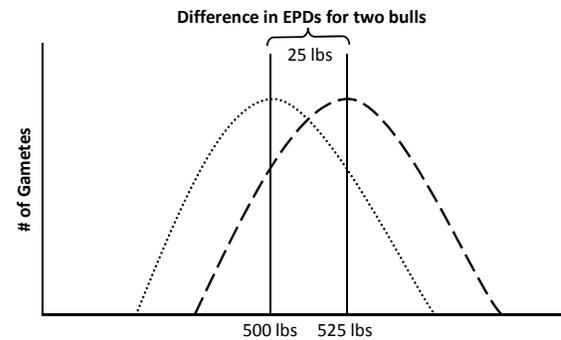
Chart 1: Example distribution of genetic potential of individual gametes produced by an animal.



Weaning Weight Genetic Potential

Compare the first bull with genetic potential of 500 lbs with a second bull with the genetic potential of 525 lbs (Chart 2). On average the second bull's calves exceed the first bull's calves by 25 lbs, but note there is considerable overlap in the genetic potential of the two bulls. That is why an inferior bull can produce calves that exceed the performance of some of the calves of the superior bull, but on average the superior bull puts an additional 25 lbs on his calves and he should produce progeny that exceed the heaviest progeny of the inferior bull. In this scenario, assuming no unexpected variation, the superior bull would have a weaning weight EPD that was 25 lbs greater than the inferior bull.

Chart 2: Example distribution of genetic potential of individual gametes produced by an animal.



Weaning Weight Genetic Merit

One of the reasons that EPDs are not a perfect science, and sometimes yield incorrect results, can be explained in a similar manner. Assume that the bull's true genetic potential is depicted in Chart 1, with the average of his sperm resulting in 500 lbs of genetic potential. If we had all of this information and conducted the genetic evaluation then we would get a correct EPD. However, let's assume this is a young sire and he has only produced 5 calves. Let's further assume that the 5 calves he produced happened to get his best mix of genes and averaged 550 lbs, instead of the expected 500 lbs. Because of other information that goes into the computation of the bull's EPD he would not end up with an EPD that was 50 lbs larger than correct, but it could be significantly larger than his true genetic potential. Under this scenario let's assume that many other producers use semen from this bull and the next analysis he has 100 calves represented and the average of these calves would likely be closer to the expected 500 lbs, resulting in a much smaller weaning weight EPD, but more correct and with a higher accuracy.

As mentioned earlier, accuracy is a reflection of the potential unexplained variation associated with EPDs and is dependent on the amount of data available for the computation of the EPD. The methodology used to compute accuracy is irrelevant to most beef producers, but knowledge on how to use this information may be beneficial in the risk management of selection decisions. In beef cattle genetic evaluations the accuracy value is a reflection of the range of potential change in the EPD as it approaches true genetic merit. Each breed prints a potential change table in their sire summary to indicate the range of potential change for each accuracy value. Table 1 is a fictitious example of such a table. In this example let's assume that a bull has a birth weight EPD of +3.0 lbs and an accuracy of .10. The potential change for birth weight associated with an accuracy of .10 is ± 2.7 lbs. In reality, this means that there is a 2/3 probability that the bull's true EPD for birth weight is between +.3 lbs and +5.7 lbs. If management dictated that a 4 lb birth weight EPD was the greatest that should be used in this herd, then there is potential that this bull would exceed that. As additional information is gained on an animal the accuracy values of their EPDs also increases, which results in a decrease in potential change. Let's assume that a second bull with a 3.0 lbs birth weight EPD had an accuracy value of .70. Now the 2/3 probability of potential change is $\pm .9$ lbs for a range of +2.1 to +3.9 lbs for the true birth weight EPD. Even though the two bulls had the same EPD of 3.0 lbs the second bull could be used with greater confidence of not exceeding the 4.0 lbs birth weight EPD maximum for this scenario.

Table 1: Example of potential change in EPD associated with different accuracy values.

For most commercial producers the purchase of young, unproven, sires is reality and low accuracy values should not be a discouragement from using EPDs in selection decisions. Low accuracy EPDs are still the best source of information available for making selection decisions on that trait. However, they can be used to help manage risk on traits that are of extreme importance to the beef cattle operation. In cases where there is little room for error it is sometimes necessary to go to greater extremes on the EPDs for certain traits when using low accuracy bulls.

Expected Progeny Differences provide beef producers with a reliable tool to make selection decisions. They are not perfect and there is error associated with all EPDs; however, every EPD has an accuracy value associated with it to help producers manage the level of risk they are willing to take for each trait they are selecting for. The methodology for computing EPDs will likely change over time and could incorporate genomic information in the future; with all of the successes associated with genetic evaluations there is potential that they can become even better and more accurate in the future.