Expected Progeny Differences

Matthew L. Spangler, University of Nebraska-Lincoln

It is impossible to visually determine the genetic potential of an animal as a parent for traits that are controlled by numerous genetic variants, as is the case for fertility, growth, carcass merit, and other trait complexes of economic importance. Consequently, predictions of genetic merit have evolved over the last several decades and now include phenotypic information, pedigree information, and more recently genomic information. These predictions are called Expected Progeny Differences (EPD) and have been proven to be the most reliable tool to generate change from selection.

What Are Expected Progeny Differences?

Expected Progeny Differences are predictions of genetic merit of an individual as a parent. As the name would imply, they are predictions of the differences in individuals' offspring performance. Historically, most beef breed associations conducted a genetic evaluation twice annually, meaning that EPD were updated twice a year. This schedule was due to the fact that new data were generally available twice a year, to correspond with two general calving seasons (spring and fall). However, with the advent of genomic information, new data are continually available, and producers wish to see the changes in EPD that result from the new data. This has necessitated weekly genetic evaluations, and thus updated EPD are available on a weekly basis for the majority of beef cattle breeds. In other words, more frequent genetic evaluations mean more current predictions of the genetic merit of animals.

How Do You Use EPD?

Expected Progeny Differences are tools designed to compare animals based on their genetic potential as parents and to make directional change for a particular trait. Simply knowing an animal's EPD for a given trait has no meaning without something to compare it to. This comparison can be between animals or between an animal and a point of reference, such as the average of a particular breed. Breed averages are rarely 0. Rather they represent either a point in time or a set of reference animals (i.e., historic set of high accuracy sires). Knowledge of breed average is helpful in determining how an animal ranks within a given breed for a particular trait. Most breeds publish a percentile rank table which allows producers to determine how an animal ranks for a particular trait within a particular breed. Expected Progeny Differences are reported in units of the trait. For example, weight traits (e.g., birth, weaning, yearling) are reported in pounds. However, some traits are reported as percentages (e.g., heifer pregnancy, docility).

	Weaning Weight EPD
Bull A	50
Bull B	60

With this in mind, the interpretation of the difference in EPD between two bulls is the average difference in performance of their offspring if the bulls were mated to the same cows and the calves were reared in the same environment. Following is an example.

Based on this example, on average, we expect the offspring of Bull B to weigh 10 pounds more than the offspring of Bull A. This does not mean that every calf from Bull B will weigh more than every calf sired by Bull A. There will be variation in the weights of calves produced by both bulls, but with large enough groups of offspring the average difference will be reflected by the difference in sire EPD.

Calculating EPD

The actual calculation of EPD requires the use of sophisticated statistical approaches and modern computational resources. To put the task into perspective, larger breed organizations calculate EPD for approximately 12 to 20 traits for more than10 million animals on a weekly basis. This is not a trivial task. However, the calculation of reliable EPD begins at the ranch level. Accurate phenotypes and correct accounting for management differences is the responsibility of the breeder. Although advancements in the use of artificial intelligence (AI) algorithms are beginning to penetrate genetic evaluations as a means of categorizing data in terms of quality, the fundamental responsibility will always belong to the breeder to ensure that records are accurate. Records collected at the ranch level are sent to breed organizations where they are adjusted for effects such as age of the animal, age of the animal's dam, and breed composition. These adjusted records are then used in the genetic evaluation.

The genetic evaluation itself uses a system of equations referred to as the mixed model equations (MME). This system of equations uses phenotypes of animals from across the country, and in many cases internationally, to estimate the genetic value of animals. This method requires that animals are linked through relationships, either pedigree or genomic based. Given these linkages, an animal's genetic merit is informed not only by its own phenotype but also by the phenotypic records of relatives from other herds and across time. The more closely related two individuals are, the more they contribute to the other's EPD.

What Are Accuracies?

Accuracy is the theoretical correlation between an animal's EPD and their true genetic merit and can range between 0 and 1. In the U.S. beef industry, Beef Improvement Federation (BIF) accuracy is used, which is much more conservative than "true" accuracy. Expected Progeny Differences are predictions and thus are not known with complete certainty. They are updated, and become more accurate, when additional data becomes available. For example, a young non-parent animal may have a record for their own weaning weight. If the animal becomes a parent and has offspring with recorded weaning weights, their offspring inform their EPD. This increases accuracy. Another source of data that increases accuracy is genomic data. Genomic information, in the form of SNP markers, is routinely included in the genetic evaluation of all major U.S. beef cattle breeds. This enables higher accuracy predictions, particularly for non-parent animals. One way that genomic information is used to increase accuracy is by improving the estimates of relationships between animals. Instead of relying solely

on pedigree information to inform kinship, genomic data can be used to determine the relationship between animals. For example, although the expectation (pedigree) of the relationship between an individual and their grandparent is 0.25, the true relationship (genomic kinship) can range between 0 and 0.5 due to sampling of alleles inherited by different animals from their parents. By estimating relationships more accurately, EPD become more accurate.

Contemporary Groups

A contemporary group represents a set of animals that were given an equal opportunity to perform and shared a common environment. The foundation for a contemporary group includes animals born in the same year, season, herd, and who were treated equally. In other words, if a subset of animals is fed differently (given preferential treatment) they should become a separate contemporary group. Admittedly there is an optimization between accounting for environmental effects through contemporary groups and allowing contemporary group size to be large enough to compare animals (and parents). At the limit, a contemporary group size of one would perfectly account for the unique environmental effects experienced by the animal. However, single animal contemporary groups are not useful for genetic evaluation as the animal's genetic merit becomes completely confounded with the environmental effects.

It is critical to report data on all animals in a contemporary group. Not doing so leads to biased estimates of genetic merit. If only the heaviest 50% of calves have weaning weights reported, then the magnitude of the differences between each animal and the average of the contemporary group is shrunk, incorrectly suggesting that the animals reported are not as superior for pre-weaning growth as they actually are.

Direct vs. Maternal EPD

Some phenotypes are influenced by both the genetics of the individual (direct) and genetics of the dam (maternal). Examples include weaning weight and calving ease. The EPD for weaning weight direct is simply called weaning weight whereas the maternal EPD for weaning weight is called milk. In beef cattle, milk EPD is expressed in pounds of weaning weight due to maternal influences, principally lactation. Milk EPD can be thought of as the comparison of a bull's grand-progeny that are products of his daughters. Calving ease also has a maternal genetic component. Calving ease direct EPD represent the probability of how easily a bull's calves will be born when he is bred to heifers. Calving ease maternal EPD are a misnomer in the sense that they reflect total maternal merit. Total maternal is the sum of maternal EPD and half of the direct EPD and represents the probability of unassisted births of a bull's daughters during their first parturition. Although calving ease maternal EPD are not labeled as such, the majority of beef breed associations publish total maternal calving ease.

Multiple-trait Analysis

Many traits are genetically correlated to each other. As such, knowledge of the performance of one trait informs the genetic prediction of another, correlated trait. Growth traits are a good example. Birth, weaning, and yearling weight are all genetically correlated with each other and as a consequence are evaluated in the same multiple-trait model. This has two primary benefits. First, it enables early growth traits to inform the EPD of later growth traits before the later growth traits are observed. Secondly, it mitigates the impact of selection that has occurred earlier in life (sequential selection) on EPD. In the case of yearling weight, it is conceivable that animals with low weaning weights were culled prior to the collection of yearling weight. Accounting for this selection decision is critical to avoid bias in traits measured later in life, in this example yearling weight. Using a multiple-trait model accounts for the fact that selection occurred and some animals were culled while others were not. An important caveat is that although yearling weight EPD are reported, the actual trait analyzed is post-weaning gain. Resulting EPD for weaning weight and post weaning gain (adjusted to 160 days) are then summed and reported as yearling weight EPD. Another example of a multiple-trait model is calving ease and birth weight. Birth weight is a useful indicator of calving ease and is thus included in the same model as the economically relevant trait of calving ease. This means that resulting calving ease EPD incorporate birth weight observations, and selecting on both calving ease and birth weight EPD results in overemphasizing birth weight.

Multi-breed Analysis

In the U.S. beef industry there is a mixture of single- and multi-breed genetic evaluations. Single breed genetic evaluations utilize data from only one breed, while multi-breed genetic evaluations utilize data from multiple breeds. Currently the largest multi-breed genetic evaluation is International Genetic Solutions (IGS). The goals of multi-breed genetic evaluations are sharing of data across breeds and the ability to report EPD across multiple breeds that are directly comparable to each other. The underpinning of a multi-breed genetic evaluation is pedigree ties across breeds and contemporary groups that include animals from more than one breed (or crossbred animals). Pedigree ties across breeds enable the sharing of data across breeds. Generally speaking, Angus and Red Angus serve as the links that tie multiple breeds together largely due to composite programs such as LimFlex, Balancer, and SimAngus. Having contemporary groups that contain more than one breed enable the estimation of breed differences, which are needed to conduct a multi-breed genetic evaluation. Without this, breed differences must be obtained from external sources (e.g., U.S. Meat Animal Research Center).

Summary

Expected Progeny Differences enable genetic selection decisions for multiple traits. Core to accurate EPD are well-formed contemporary groups. Expected Progeny Differences change over time as additional information is available. These changes are more frequent with weekly genetic evaluations. Genomic data that is integrated into EPD allows accuracy of non-parent animals to increase.