Understanding and Managing Variation in Nutrient Composition

Bill Weiss
Department of Animal Sciences
Ohio State University
Wooster 44691
Phone: 330-263-3622
email: weiss.6@osu.edu

Normand St-Pierre
Department of Animal Sciences
Ohio State University
Columbus 43210
Phone: 614-292-6507
email: st-pierre.8@osu.edu

Summary

The nutrient composition of all feeds vary, and nutritionists must learn to accommodate variation when formulating rations. Variation in feed composition increases risk and that has a cost. If the composition of a feed is highly variable, diets either have to be over supplemented to avoid a deficiency (i.e., increased feed costs) or production may decrease because at times the diet does not provide adequate nutrients. Feeds with large variation in nutrient composition are worth less than feeds with less variation. If adequate nutrient composition data are obtained and summarized for specific ingredients, a reliable mean and standard deviation can be calculated. Those values, rather than values from a single sample should be used to formulate diets. The standard deviation can be used to 'adjust' diet formulation so that the risk of a deficiency of a specific nutrient can be greatly reduced. The impact of variation in nutrient composition of a feedstuff is reduced when diets contain a wide variety of different ingredients.

Introduction

The approach followed by most nutritionists is to sample a feed, have it analyzed, and then formulate a diet based on that information. When a new analysis is obtained, the previous data is eliminated and a new diet is formulated based on the new composition data. The inherent assumption of this practice is that the new data better represents the feed than did the old data. This may or may not be true. When new analytical data are obtained, the user should ask one simple question: Is there a good reason why the composition changed? Possible answers to that question include: I changed suppliers, I started feeding a new cutting of alfalfa, or probably most commonly, I don't know. If you cannot think of a good reason the composition change, the change may simply be a random event. The difference could be caused by load to load random variation, by within load (i.e., sampling) variation, or both. In this case, the new number may be no better than the old number but the mean of the two numbers has the lowest probability of being substantially wrong.

The purpose of this paper is to discuss how feed analysis data should be used and method of reducing variation in the composition of diets. With proper sampling techniques, adequate number of samples, and appropriate data handling, you can reduce the uncertainty associated with feed analysis data which can result in increased profitability.

Elementary Statistics

We need to start thinking about feed composition data in terms of probabilities rather than actual, absolute concentrations. In other words, how confident are you (or should you be) that the number you have actually represents the true concentration of a nutrient in a feed? Because we are working with probabilities, a basic understanding of some statistical principles and terminology is needed.

Populations and Samples

The ultimate goal of feed analysis is to obtain an analytical value from a sample that reflects the actual value of a 'population'. Examples of a population include a truckload of distillers grain, all the distillers grain produced by a specific distillery, or perhaps all the distillers grain produced in the U.S. In statistical terms, a population is defined as a larger set from which samples are obtained. In other words, a population is defined by sampling. If distillers grain from a single distillery were sampled extensively, we would have a good estimate of the average nutrient composition of distillers grain produced at that plant. However, since other distilleries were not sampled we should be very hesitant to extrapolate the data obtained from a single distillery (i.e., a population) to the larger population of all distilleries.

Central Tendency and Dispersion

Because of inherent variation and variation caused by sampling and analytical procedures, we know that all the sample values taken from a population will not be the same. Rather than one single value, you will obtain a distribution of values. The two most important pieces of information we need from that set of samples are a measure of central tendency and a measure of dispersion. For observations that follow a normal statistical distribution, the mean (in this discussion average and mean will be used interchangeably) is the best measure of central tendency. The mean of a normal distribution is not the absolute 'right' answer, but rather it is the value that has the lowest probability of being substantially wrong. The concentrations of most nutrients in plant-based feedstuffs fit an approximately normal distribution; therefore the mean is the best measure of central tendency for those nutrients. With a normal distribution, approximately one-half of the samples will have values lower than the mean and one-half will have concentrations higher than the mean.

Although many people are familiar with and often use measures of central tendency (i.e., the mean) in ration formulation, fewer people consider or use measures of dispersion in ration formulation. In simple terms, a measure of dispersion should be used to determine how much confidence you have when using a mean value. When a distribution of values has a large dispersion, the probability of being substantially wrong increases when using the mean. For a normal distribution the most common measure of dispersion is the standard deviation (SD). In a normal distribution approximately 38% of all observations will be within $\pm$ 0.5 SD units of the mean, 68% of all observations will be within $\pm$ 1 SD of the mean and approximately 95% of the observations will be within $\pm$ 2 SD of the mean. For example, if the mean concentration of crude protein in a population of brewers grains was 25% and the SD was 2 we would expect that about 68% of the samples from that population would contain between 23 and 27% CP and 95% of the samples would contain between 21 and 29% CP. The smaller the SD, relative to the mean, the less likely it will be that using the mean value will cause a substantial error in diet formulation.

Sources of Variation

Understanding potential sources of variation in feed composition data will help determine which data to use and how to use it. The nutrient composition of feeds can be influenced by plant genetics (hybrid, variety, etc), growing and harvesting conditions (climate, soil fertility, maturity, storage conditions, etc.), and manufacturing techniques. The above sources of variation are considered fixed, i.e., they can be described and replicated). For example, Hybrid X may have been bred to produce corn silage with higher than average NDF digestibility, or Distillery Y may dry their distillers grains at very high temperatures causing high concentrations of acid detergent insoluble protein. Another possible fixed source of variation is the analytical lab. Although great progress has been made in standardizing methods, labs often use slightly different analytical techniques to measure nutrients. Other sources of variation are considered random. We do not know why the values differ, they just do. If you sample a load of brewers grains 10 times and send those 10 samples to a lab, you will probably get back 10 different concentrations of protein. The variation could be caused by variation within the load of brewers grain or it could be caused by random errors at the lab.

Ideally, random variation would be considered within population variation and fixed variation would be considered as variation between populations. For example, because of manufacturing differences, distillers grains from distillery X has consistently higher NDF concentrations than distillers grains from distillery Y. If distillers grains from X and Y were considered separate populations, the SD within each population would be expected to be lower than the SD when the distillers grains from both distilleries were combined. Because of blending grains and multiple sources of feedstock for manufacturing facilities, many fixed sources of variation become blurred (you probably will not know the variety of the soybeans used to make the soybean meal you purchased or whether the gluten feed you purchased was made from drought-stressed corn grain). In these situations, the fixed sources of variation become random sources resulting in an increase in the within population variation. Nonetheless, accounting for as many fixed sources of variation as possible by defining separate populations will reduce the dispersion of the data and reduce the potential of being substantially wrong when using the mean.

Variation in Feed Composition

The largest publicly-available data base of feed composition can be found in the NRC dairy publication (NRC, 2001). That database contains both means and SD for measured nutrients in most common feeds (some example feeds are shown in Table 1). The data used to calculate those means and SD came from a wide array of sources. Samples came from across the U.S. and over several years. For some feeds and nutrients, the number of samples used to calculate the mean and SD is quite limited and those values should be used carefully. For other feeds, the sample size is quite large and the mean and SD are probably good estimates for the broad population from which the samples were drawn. However, it is important to remember that the broad population represented in the NRC tables may not be an accurate estimate for a specific source of a feed. For example, the mean concentration of CP in dried distillers grain in NRC is 29.7% with a SD of 3.3. DePeters et al (2000) and Belyea et al. (1989) collected multiple samples of distillers grains from a single distillery in California and Missouri, respectively (Table 2). The mean CP concentration from the CA and MO distilleries were 30.6 and 31.2%; both were very similar to the NRC mean. However, the SD from the CA and MO datasets was 1.6 and 0.6; both substantially lower than the SD in NRC (Figure 1). The lower SD for the CA and MO data is expected since many of the sources of variation that contribute to the SD for the NRC data are not included in the CA and MO data.

Based on the mean CP concentrations (and assuming the concentrations of other nutrients are similar), all three distillers grains (NRC, CA, and MO) are worth approximately the same per ton, The true economic value of the three distillers grains, however, is not the same. If distillers grain averaged 30% CP and it had an SD of 3.3 (i.e., the SD in NRC), you would expect that about one-third of the time, the CP concentration would be less than 28%. For distillers grain with a mean CP of 30% and an SD of 0.6, you would expect that the CP concentration would almost never be less than 28%. With the high SD distillers, you are at much greater risk of feeding a diet with insufficient CP (resulting in lower milk yield) than if you fed the low SD distillers. To avoid a protein deficiency you might want to feed a higher CP diet (resulting in increased feed costs) when you used the high SD distillers. This is an example of the cost of variation.

To accurately calculate the cost associated with variability you need a good estimate of the variation in the feeds. For some feeds, NRC values (mean and SD) are adequate. Feeds with low expected variability (e.g., corn grain, soybean meal) do not have to be analyzed routinely and in some cases not at all. Sampling and analytical errors become small when large numbers of samples are analyzed. For these feeds, a mean derived from a large number of samples may actually be better than a single observation or a mean from a small set of samples. Book values can be used unless you have good reason to believe your particular feed is different (for example, if you grow or buy high oil corn, the mean values for regular corn would not be appropriate). However for other feeds, the SD obtained from a broad population (e.g., NRC) may not accurately reflect variation in specific situations. If no other measure of dispersion is available to you, the SD in the NRC table can be used; however, you must remember that for many feeds, the actual variation you will observe could be substantially less than the SD in the NRC table (Table 3). Taking adequate samples and calculating a SD for each specific situation is usually preferred. To do this, data from each sample is collated into a spreadsheet and the mean and SD are calculated as data accumulate (all spreadsheets have a function to calculate SD). For example, if you sample alfalfa silage once a month, data from those samples are entered into a spreadsheet and mean and SD are calculated after each entry. If you have reason to believe the population changed (for example, you think you reached a new cutting in the silo), a new column of data is started and you start calculating the mean and SD for the new population. The more samples you have, the more accurate your estimates of the mean and SD will be. There is no specific number of samples needed (highly variable populations require more samples than less variable populations) but generally 5 to 10 samples will yield reasonable estimates of the SD.

How Should You Handle Variation in Feed Composition

As mentioned several times above, the SD is an important statistic. It is an indicator of how wrong you could be. In Table 1, corn gluten feed had a mean CP concentration of 23.8 and an SD of 5.7. Assuming a normal distribution and assuming you received totally random loads of corn gluten (i.e., not from a single source), approximately 16% of the loads would have a CP concentration less than 18.1% and 16% of the loads would have a CP concentration greater than 29.5%. If a particular load of corn gluten had 18% CP and you used the mean concentration and corn gluten made up 10% of the diet DM, the actual CP concentration of the diet would be about 0.6% units lower than the formulated value. An error of this magnitude or larger would be expected 16 out of every 100 loads. If you are willing to accept this risk, then using the mean is the best option. However, if based on your experience, you conclude that milk production will drop 2 lbs (or some other number) if the diet contains 0.6 percentage units less crude protein than formulated and you are unwilling to accept that risk (even though this will happen only 16% of the time), you need to adjust for variation. You can reduce your risk of substantially under feeding CP by 'adjusting' the mean value based on its SD. Based on a normal distribution, if you use the mean minus 0.5 X SD, rather than the mean, you reduce the risk of making the error discussed above from 16% of the time to 7% of the time. If you use the mean minus 1 SD unit, you reduce the risk of making the above error to just 2% of the time. In the example above, mean CP for corn gluten was 23.8 (SD = 5.7). If I was willing to risk being substantially wrong 7 out of every 100 loads of corn gluten feed, I would use 23.8 - (0.5 x 5.7) or 21.0% CP for corn gluten feed when I balanced the diet. If I only wanted to be substantially wrong 2% of the time, I would use 23.8 - 5.7 = 18.1% CP. By using a lower CP concentration for corn gluten feed, I have substantially decreased the probability of being substantially deficient in CP; however, I will be over supplementing CP most of the time. You will need to determine how much risk you are willing to accept and balance that against increased feed and environmental costs.

The problem with this approach is that it only considers variation in a single ingredient but the composition of all ingredients in a diet will vary. What really matters is not the variation in a single ingredient but rather the variation and mean for a diet. A software program is being developed at Ohio State that will calculate the expected variation in nutrient composition of a mixed diet based on variation in composition of the feedstuffs.

Ways to Reduce the Impact of Variation

The variation in nutrient composition of a diet is a function of the variation in nutrient composition of the ingredients, the inclusion rate of each ingredient, and the number of ingredients in the diet. Obviously using highly variable ingredients increase the variation in nutrient composition of the diet but the effect of individual ingredient variation on variation in the diet is strongly influenced by the inclusion rate. On a theoretical basis, the contribution of a feedstuff to the variance of the total diet grows with the square of its inclusion rate. In other words, if you double the inclusion rate of a feed its impact of diet variation increases by a factor of 4. If you double the inclusion rate of a highly variable ingredient you will substantially increase diet variation but if you double the inclusion rate of a very consistent ingredient you will substantially decrease the diet variation. Highly variable ingredients can be used successfully and can be profitable (assuming the price is right) as long as inclusion rates are low. If high inclusion rates are used, the costs associated with the increase in uncertainty regarding diet composition may overwhelm any apparent savings in ingredient costs. The last method of reducing variability of diets is to increase the number of ingredients in the diet. Variation among feeds is independent. In other words, if the corn silage you fed today had lower than average CP, the alfalfa, corn grain, and soybean meal you feed today will not necessarily have lower than average CP. Independence of variation across feeds is a powerful tool for reducing variation in nutrient composition of diets. To illustrate this principle, assume that a diet has only one ingredient that provides nutrient X. Each time you make up the diet you have a 50% chance that the concentration of nutrient X will be greater than expected (i.e., the mean) and a 50% chance it will be less than expected. If you put together a diet with two ingredients that provided nutrient X, you have a 25% chance that both ingredients will have lower than expected concentrations of nutrient X, a 25% chance that both will have higher than expected concentrations of nutrients X and a 50% chance that one ingredient will have higher and one ingredient will have lower than expected concentrations of nutrient X. If you had 10 ingredients that provided nutrient X, the chances that all 10 ingredients had higher than expected concentrations of nutrient X at the same time is 1 out of 1000 (if that happens on your farm, go buy a lottery ticket). Increasing the number of ingredients will reduce diet variation only if good feeding management practices are in place. Inaccurate weighing, mixing, etc, of ingredients will increase variation in the diet.

The effect of increasing the number of ingredients on reducing variation in the diet decreases as the number of ingredients increase. The greatest reduction in variation is observed going from one ingredient to two ingredients whereas going from 9 ingredients to 10 ingredients would have a minimal effect on diet variation. A simple example will be used to illustrate the effect of inclusion rate and number of ingredients on diet variation (Figure 2). The example diet has 60% forage and has either one forage (60% of diet is forage A), two forages (30% of the diet is forage A and 30% is forage B), or three forages (20% of the diet is forage A, B, and C). For this example, all forages contain on average 40% NDF with an SD of 2.5. Since all forages have the same average NDF, all diets also have the same average concentration of forage NDF (24%). The SD for the diets with one, two, and three forages is 1.50, 1.08, and 0.89. For the diet with a single forage, the SD for the diet is the SD of the forage times its inclusion rate (i.e., 2.5 X 0.6). The SD for the diets with multiple sources of forage reflects inclusion rate of each forage and the effect of multiple ingredients. Increasing the number of forages from one to two (remember the total amount of forage in the diet did not change and all forages had the same SD for NDF) reduced the SD for forage NDF in the diet by 28%. Going from 2 to 3 forages reduced the diet SD by only 17%. In this simple example, all forages had the same variation which is unlikely in the real world. Does increasing the number of ingredients still reduce variation if certain ingredients have more variation? Often times, the answer to that question is yes. In the example above the SD for forage NDF in the diet was 1.5 when a single source of forage that had an SD for NDF of 2.5 was used. If you fed a diet with 30% of that forage and 30% of a more much more variable forage (SD = 4.0, a 60% increase in variation), the expected variation in the concentration of dietary forage NDF is 1.40 or about 7% less than a diet with only a single forage even though that forage was fairly consistent. This illustrates the importance of relying on multiple sources of nutrients when the goal is to provide a more consistent diet.

References

Belyea, R. L., B. J. Steevens, R. J. Restrepo, and A. P. Clubb. 1989. Variation in composition of by-product feeds. J. Dairy Sci. 72:2339-2345.

DePeters, E. J., J. G. Fadel, M. Arana, et al. 2000. Variability in the chemical composition of seventeen selected by-product feedstuffs used by the California dairy industry. Prof. Anim. Sci. 16:69-99.

National Research Council. 2001. Nutrient Requirements of Dairy Cattle. Natl. Acad. Press, Washington DC.

Table 1. Average (Avg) concentrations and standard deviations (SD) for CP, NDF, and ether extract (EE) in selected feeds. Data are from NRC (2001) and represent very diverse populations.

	CP		NDF		EE	
	Avg	SD	Avg	SD	Avg	SD
Grains						
Barley	12.4	2.1	20.8	8.6	2.2	0.6
Corn	9.4	1.3	9.5	2.3	4.2	1.0
Sorghum	11.6	1.8	10.9	5.0	3.1	0.8
Byproducts						
Wet brewers	28.4	4.0	47.1	6.8	5.2	1.6
Corn gluten feed	23.8	5.7	35.5	6.8	3.5	1.1
Dry distillers grain	29.7	3.3	38.8	7.8	10.0	3.4
Potato waste	10.5	8.4	22.1	14.3	10.8	7.8
Rice bran	15.5	2.2	26.1	6.8	15.2	4.2
Soyhulls	13.9	4.6	60.3	7.4	2.7	1.4
Soybean meal-48	53.8	2.1	9.8	5.6	1.1	0.4
Wheat midds	18.5	2.1	36.7	7.5	4.5	1.3
Forages						
Corn silage	8.8	1.2	45.0	5.3	3.2	0.5
Alfalfa silage (AS)	20.6	3.0	45.7	6.5	3.1	0.7
Mid-maturity AS	21.9	1.8	43.2	1.5	2.2	0.3

Table 2. Mean ˜() concentrations and standard deviations for selected nutrients and selected feeds. The California data are from Depeters et al. (2000) and the Missouri data are from Belyea et al. (1989). Within experiment and feed, samples originated from the same production facility (i.e., limited populations). These values should be compared to those in Table 1 (a broad population).

	CP		NDF		EE	
	˜	SD	˜	SD	˜	SD
CA						
Brewers grain, wet	27	2.2	37.3	3.4	6.3	0.4
Corn gluten feed, wet	22.9	4.3	38.8	3.8	3.4	0.4
Distillers grain, dried	31.2	0.6	35.6	8.2	13	1.3
MO						
Corn gluten feed, dry	23.3	1.4	51.9	2.3	6.6	1.9
Distillers grain, dried	30.6	1.4	33	1.5	7.4	0.9
Rice bran	19.1	0.4	21.8	1.3	17.3	1.9
Soybean hulls	11.8	0.2	72.5	0.8	0.8	0.3

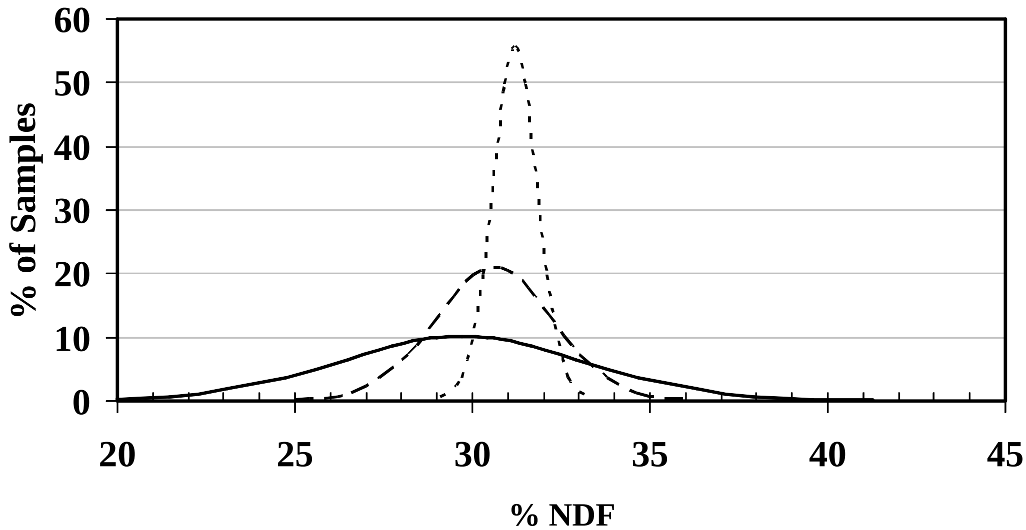

Figure 1. Distributions of crude protein concentrations in dried distillers grains. The solid line represents data from a nationwide population (NRC, 2001). The small-dashed line represents samples from a single source in Missouri (Belyea et al., 1989) and the large-dashed line represents samples from a single source in California (Depeters et al., 2000). Although mean concentrations were similar among populations, note that dispersion is substantially less for the limited populations (CA and MO) compared to the broad population (NRC). Distributions were calculated based on the mean and SD.

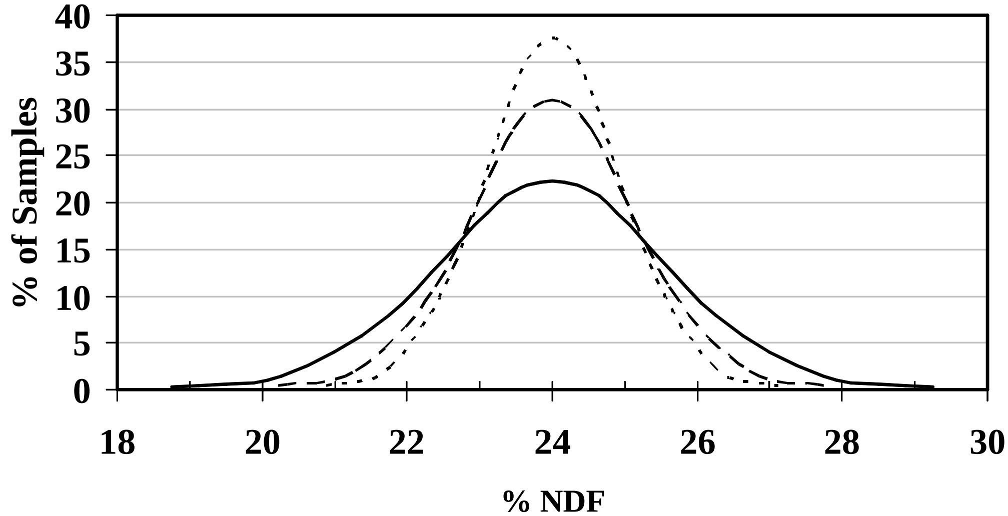

Figure 2. The effect of increasing the number of forages on variation in the concentration of forage NDF in a diet. The solid line indicates the variation in the concentration of forage NDF in the diet when a single forage is fed (60% of diet dry matter). The large dashed lines indicates the variation when two sources of forage is fed (30% of diet DM for each forage). The small dashed line indicates the variation when three sources of forage is fed (20% of diet DM for each forage). In this example all three forages had an average NDF concentration of 40% and an SD of 2.5.