Department of Grain Science and Industry

Sampling: Statistical and Economic Analysis

Tim Herrman
Extension State Leader
Grain Science and Industry

Every ingredient included in a feed formulation possesses important nutrients that maximize animal performance at the lowest possible cost. Sampling and sample evaluation enables the feed manufacturer to make inferences about the quality of incoming grain, protein sources, micro-nutrients, and finished feed.

Increasingly, quality assurance personnel must justify the number of samples collected and assayed. Statistical and economic analysis should be applied to a quality assurance program to determine sampling and evaluation frequency. Important factors that determine the design and implementation of a sampling program involve shipment size, ingredient variability, laboratory accuracy and cost of the assay, and value of the ingredient.

Definitions

A **population** is the set of all measurements (frequently characterized as infinite) that is sampled. An **element** is an object on which a measurement is taken, and **sampling units** are non-overlapping collections of elements from the population. A **frame** is a list of sampling units, and a **sample** is a collection of sampling units drawn from the frame.

The **population mean**, μ, is the average of the population. The **population variance**, σ^2, is the average of the square of the deviations about μ. Because the population is infinite, or at least too large to measure, utilize sample statistics to represent the mean (Equation 1) and variance (Equation 2):

Equation 1. Mean

$$\bar{y} = \frac{y_1+y_2+y_3 \ldots y_i}{n}$$

Equation 2. Variance

$$s^2 = \frac{[\sum y_i^{\,2} - (\sum y_i)^2 \div n]}{n-1}$$

Where:
$\bar{y}$ = sample mean
s^2 = sample variance
y_i = individual sample measure
n = number of samples
$\sum$ = sum

Sample Size for Estimating Population Mean

Sampling patterns, sample size, and number of samples are frequently specified in official procedures by scientific associations or agencies (Herrman 2001). These techniques facilitate trade (typically for commodities) and assist in settling disputes between buyers and sellers. Individuals looking to optimize their sampling program can utilize the following statistical procedures when deciding how many sample units to collect from a population.

Random Sample

A random sample is one in which a sample of size n is drawn from a population of size N in such a way that every sampling unit has an equal probability of being selected. Example 1 illustrates how to calculate sample size for bagged feed ingredients.

Example 1: A truckload containing 400 bags of fish meal arrives at the feed mill: 12 bags are randomly selected and probed, and fish meal in each sampling unit is analyzed for protein content (Table 1).

Table 1. *Fish meal protein content from 12 bags.*

65.1	64.1	65.5	66.0
64.8	65.1	65.0	67.0
65.7	65.4	64.9	66.9

The mean and variance for fish meal protein content is calculated as follows:

$$\bar{y} = \sum \frac{y_1 + y_2 + y_3 \ldots y_i}{n} = \frac{785.5}{12} = 65.5$$

$$s^2 = \frac{[\sum y_i^{\,2} - (\sum y_i)^2 \div n]}{n-1} = \frac{51{,}425 - (617{,}010 \div 12)}{11} = 0.7$$

The mean and variance for the fish meal protein content are used to calculate the number of bags for future sampling (Equation 3). Additionally, Equation 3 includes the assumption that we are 95 percent confident that the sample number will enable us to predict the mean protein content within 0.5 percent.

Equation 3. Random sample bagged feed

$$n = \frac{N \times s^2}{(N\text{-}1)\, D + s^2}$$

Where:
D = $B^2 \div 4$
B = 0.5 percent (bound on the error estimate of the mean) and the divisor 4 represents two standard deviations above and below the mean (95 percent confidence).

Placing these values in Equation 3 results in the following:

$$n = \frac{400 \times 0.7}{(400\text{-}1) \times 0.0625 + 0.7} = 10.9$$

$$\text{where } D = \frac{0.5^2}{4} = \frac{0.25}{4} = 0.0625$$

Therefore, sample 11 bags to estimate the average protein content within 0.5 percent.

Suppose the fish meal is delivered in bulk. The number of samples that need to be collected to accurately (95 percent confidence) estimate the protein content within 0.5 percent can be calculated using the t statistic and the sample variance (0.7 percent protein) as follows:

Equation 4. Random sample bulk feed

$$t_{\alpha/2} \sqrt{\frac{s^2}{n}} = 0.5$$

$$1.96 \sqrt{0.7/n} = 0.5$$

Rearranging the equation produces the following results:

$3.84 \text{ x } 2^2 \text{ x } 0.7 = n$
n= 10.7

Each of these techniques yield the same number of sample units (11) needed to estimate the average protein content of fish meal within 0.5 percent.

Companies can develop their own tables that provide variance measures for different feed ingredients. Table 2 provides variance estimates for eight feed ingredients. Soybean meal exhibited a low variance estimate, thus requiring two samples to accurately estimate (95 percent confidence) the average protein content within 0.5 percent protein. The highly variable ingredients, meat and bone meal and cottonseed meal, require more extensive sampling (43 and 62 sample units, respectively). For these highly variable ingredients, estimating the average protein content within 0.5 percent may be unnecessary or impractical.

Table 2. *Variation within an ingredient and recommended number of samples to estimate the average protein content within 0.5 percent.*

Protein Source	Variance Estimate	Sample Number
Soybean Meal	0.1	2
Rapeseed Meal	0.3	5
Corn	0.3	5
Fish Meal	0.7	11
Wheat Midds	0.7	11
Corn Gluten Meal	1.1	17
Bakery Meal	1.3	20
Meat and Bone Meal	2.8	43
Cottonseed Meal	4.0	62

Note that the procedure described above requires that each sampling unit be assayed individually. Future assays can be performed on composite samples (combination of all the sampling units).

Other Sampling Techniques

Stratified Random Sample: A stratified random sample is obtained by separating the population elements into non-overlapping groups, called strata. A simple random sample is then collected from each stratum. Use the procedures for selecting a random sample number and apply this to each stratum.

Systematic Sampling: Systematic samples are easier to perform than a simple random sample and often provide greater information per unit cost than does simple random sampling. An example of a systematic sample involves the use of a diverter-type (D/T) mechanical sampler while loading a bulk shipment of soybean meal. The D/T is mounted in grain spouts, at the end of belts, or at the head of elevator legs, and the diverter moves through the soybean meal (takes a cut) at timed intervals.

To calculate the sample size for estimating the population mean, use Equation 4 for a random sample.

Sampling Economics

Elements within each sampling unit (e.g., protein content) represent nutritional and economic value to the nutritionist. This value can be applied to sampling decisions.

Moisture Content

Table 3 contains bakery meal moisture content from 15 trucks (samples). The **average range** (**AR**) of the moisture content between truckloads is calculated by measuring the difference between each sample (in chronological succession), totaling the differences, and dividing the sum by the number of measurements (14) in the range column.

The AR is 1.8 percent, which means there is an average of 1.8 percent difference in moisture content between each successive load of bakery meal.

The AR statistic can be applied to the economic analysis to assess the potential value of nutrient variability between truckloads of an ingredient. Table 4

Table 3. *Moisture content for bakery meal and calculation of the average range.*

Sample ID	Moisture	Range
1	11.1	
2	10.0	1.1
3	4.1	5.9
4	9.3	5.2
5	9.7	0.5
6	8.8	0.9
7	8.6	0.2
8	8.6	0.1
9	7.1	1.5
10	7.2	0.1
11	9.2	2.0
12	8.4	0.9
13	4.8	3.6
14	8.6	3.8
15	8.8	0.2
Sum	124.3	26.0
Average	8.3	1.8

Table 4. *Sample mean and average range for wheat midds, soybean meal, corn, and bakery meal.*

Ingredient	% Moisture		% Protein		% Fat	
	Mean	Range	Mean	Range	Mean	Range
Wheat Midds	12.2	0.8	17.1	0.7	4.5	0.4
Soybean Meal	12.0	0.3	48.7	0.5	—	—
Corn	13.6	0.8	7.6	0.3	3.8	0.3
Bakery Meal	8.3	1.8	11.1	1.3	8.6	1.8

Equation 5. 1 percent Mst Value = [(Specified Mst – Actual Mst) ÷ Specified DM] *x* \$ per ton

Where: Mst = moisture
DM = dry matter
Contract = 8 percent moisture
Actual = 9 percent moisture
Specified dry matter = 92 percent
Cost of bakery meal = \$90 per ton

Solution:
1 percent Mst Value = [(8% - 9%) ÷ 92% x \$90 per ton] = -\$0.98 per ton

Example 2: Calculate the economic value of analyzing a truckload of bakery meal for moisture content. Begin this process by substituting the AR for the difference between specified moisture and actual moisture (1.8%) as follows:

AR Value = 1.8% ÷ 92% x \$90/ton = \$1.78/ton

Assuming the truck contains 20 tons of bakery meal, multiply the \$1.78/ton by 20 tons per truck to arrive at a potential moisture value of \$35.60 per truck.

Moisture Value per Truck = \$1.78/ton x 20 tons/truck = \$35.60/truck

Equation 6. 1 percent protein = (20 lbs/ton ÷ CP in SBM) x ($/lb SBM)
Where:
lbs = pounds
CP = crude protein
SBM = soybean meal (in this example 47.5 percent CP)

Example 3: Calculate the value of an additional one percent protein in corn. In this example, suppose corn contains 9 percent CP versus the table value of 8 percent used in a least cost formulation program. Substituting 47.5 percent CP for the SBM and $0.10/lb for the value of SBM the value of 1 percent protein (20 lbs CP per ton) is calculated as follows:
Assumptions:
SBM is 47.5 percent CP
Cost of SBM = $200/ton

Solution:
1 percent CP = (20 lbs/ton ÷ .475) x $0.10 = $4.21 per ton

Example 4: The AR value for protein content in wheat midds is 0.7 percent. What is the potential value of sampling and analyzing a truckload of wheat midds for protein?
Assumptions:
47.5 percent protein SBM costs $300 per ton
One truckload of wheat midds carries 18 tons of wheat midds

Solution:
0.7 percent CP = (14 lbs/ton ÷ 0.475) x $0.15/lb = $4.42/ton
$4.42/ton x 18 tons/truck = $79.58 truck

Note: 0.7 percent CP is equivalent to 14 pounds of CP in one ton of SBM
To assign the value of performing a fat analysis, use a similar procedure only substitute the value of the standard fat ingredient (e.g., choice white grease).

Example 5: Calculate the value of an additional percent fat in corn.
1 percent Fat = 20 lbs/ton x $0.11/lb = $2.20/ton
Where:
Choice white grease costs $0.11 per lb

contains sample data for feed ingredients that were analyzed for moisture, protein, and fat from one feed company.

The economic value for 1 percent moisture content in bakery meal is calculated using Equation 5.

Protein and Fat

Equation 6 is used to calculate the value of 1 percent protein in feed ingredients and finished food. Applying the value of protein to a sampling program for corn and wheat midds are presented in Examples 3 and 4 respectively. Example 5 illustrates the use of Equation 6 to estimate the value of 1 percent fat in corn.

Protein and Fat Analysis

The value of protein content in a feed ingredient can be calculated using the price of soybean meal (see page 5).

Managing Quality Data

The above examples provide a mechanism for estimating the potential value associated with nutrient variability between truckloads of a feed ingredient. An analysis of truck to truck variability over an entire year can help the quality manager assess the potential savings associated with a sampling program, presented in Example 6.

Example 6: The following case study contains sample data from a feed manufacturer who utilizes 10,000 tons of SBM per year. Table 5 presents the assumptions underlying the economic analysis. The value of 1 percent CP is $4.21, and 1 percent moisture content is $2.27.

Note: The protein values were adjusted to 12 percent moisture content. Figure 1 reports the protein and moisture assay results from 108 loads of SBM in histogram form. The value associated with receiving SBM possessing protein that exceeded contract specification was about $39,024. The cost of receiving SBM that exceeded the maximum moisture contract specification was $3,678. Realizing part or all of this value requires a quality management program in which lots of SBM with different levels of protein are held in different bins and rations are reformulated based on actual protein content. Additionally, rapid quality assessments using a near infra-red (NIR) instrument are required to provide protein and moisture content of the feed ingredient upon arrival.

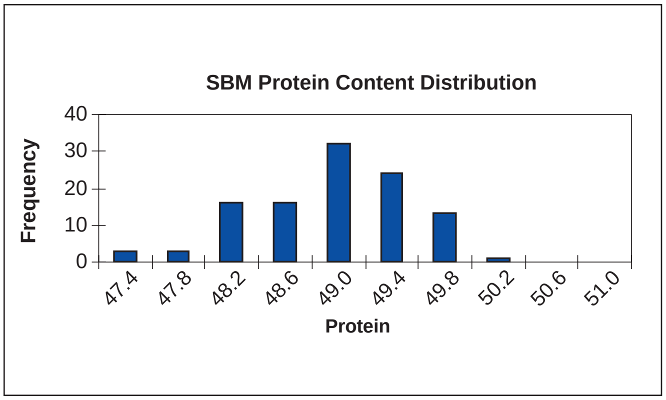
Table 5. *Protein and moisture specifications for SBM.*

Category	Limit
Protein, Lower	47.5
Protein, Upper	48
Moisture	12
Tons, Annual Purchases	10000
Value for 1 percent Protein	$4.21
Value for 1 percent Moisture	$2.27

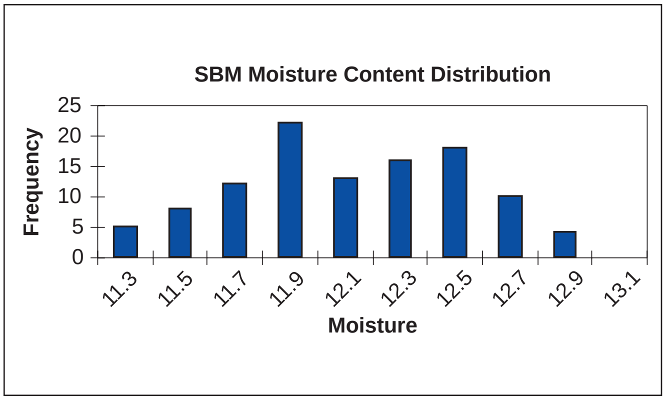
Using the results of Figure 1, one could perform a net present value analysis on the purchase of a NIR instrument and justify the cost of this investment to management.

Figure 1. *Protein and moisture analysis results for SBM*

Protein	Loads	Value $
47.4	3	-$157
47.8	3	$0
48.2	16	$1,034
48.6	16	$3,529
49.0	32	$12,048
49.4	24	$12,778
49.8	13	$8,948
50.2	1	$844
50.6	0	$0
51.0	0	$0
	108	$39,024



Moisture	Loads	Value $
11.3	5	$236
11.5	8	$41
11.7	12	$0
11.9	22	$0
12.1	13	$0
12.3	16	$0
12.5	18	-$1,799
12.7	10	-$1,420
12.9	4	-$736
13.1	0	$0
	108	–$3,678



Net Present Value of NIR Equipment

The data in Example 6 provides an estimate of the value associated with ingredient variability currently not captured by one feed manufacturing company. As previously mentioned, purchase of a near infrared (NIR) analyzer to rapidly measure ingredient protein, fat, and moisture content would greatly assist a program designed to control ingredient variability. The Net Present Value (NPV) calculation can help answer the question, "Can the money I save using a NIR instrument to measure and manage ingredient variability provide sufficient return on investment to purchase this equipment?" The following example explains how to answer this question.

Example 7. Assumptions for this example include the following:

1: An NIR instrument will enable a feed company to capture half of the lost value calculated in Example 6.

2: Soybean meal protein variation will follow the same pattern during the next three years as was observed in Example 6, and the average value for 1 percent protein is $4.21 per ton.

3: The cost of capital is 12 percent per year.

4: The cost of one NIR is $50,000.

Equations 7 and 8 are used to calculate the present value (PV) and NPV, respectively. The PV calculation adjusts the dollars saved over the next three years to current dollars. The NPV is calculated by subtracting the cost of a NIR from the PV.

Equation 7. PV = cash flow ÷ $(1 + r)^n$

Where:

Cash flow = $39,024 per year × 0.5 = $19,512

r = cost of capital at 12 percent

n = number of years

Thus:

Year 1 PV = $19,512 ÷ 1.12	=	$17,421
Year 2 PV = $19,512 ÷ $(1.12)^2$	=	$15,555
Year 3 PV = $19,512 ÷ $(1.12)^3$	=	$13,888
Total		$46,864

Equation 8. NPV = PV – investment

Where:

The NIR investment is $50,000

Thus:

NPV = $46,864 – $50,000 = –$3,135

Therefore, purchasing a NIR would return –$3,135 less than investing $50,000 capital at 12 percent interest over 3 years. Note: The manager may know of other opportunities to capture value using the NIR that could improve cash flow and make this a profitable investment.

Summary

Techniques are presented for optimizing sampling and maximizing the value of data collected from a sampling program. Quality assurance managers can show a positive return on investment for their sampling program and utilize these techniques to increase the company's profitability. Additionally, the benefits associated with a sampling program can be used to justify investment in rapid quality detection equipment.

Literature Cited

Herrman, Tim. 2001. *Sampling: procedures for feed.* MF-2036. Kansas State University Research and Extension, Manhattan.

Acknowledgements

Thanks to Jody Morgan, Carl Akey Inc., for helping to develop economic calculations for ingredient variation.

This, and other information, is available
from the Department of Grain Science at
www.oznet.ksu.edu/grsiext,
or by contacting Tim Herrman,
Extension State Leader
E-mail: tjh@wheat.ksu.edu
Telephone: (785) 532-4080

Kansas State University Agricultural Experiment Station and Cooperative Extension Service

MF-2506 May 2002