Paper No. 963086
An ASAE Meeting Presentation

SAMPLE SIZE CALCULATIONS FOR STRAY VOLTAGE MILK REDUCTION STUDIES

by

Steven D. LeMire
Graduate Research Assistant
and
Douglas J. Reinemann, Ph.D.
Associate Professor

Department of Biological Systems Engineering
University of Wisconsin-Madison, Milking Research and Instruction Lab

Written for presentation at the
1996 ASAE Annual International Meeting
Sponsored by
ASAE: The Society For Engineering in Agriculture, Food and Biological Systems
Phoenix, Arizona
July 14-18, 1996

Summary:

Both positive and negative hypothesis tests should have associated error rates stated in order for the researcher to draw inferences from an experiment. To insure minimally acceptable error rates for conclusions, adequate sample sizes are needed. This paper presents methods of calculating sample size and several examples of sample size calculations. Scenarios considered are for experiments involving reduction in the mean milk production of dairy cows for both field and controlled experiments.

Keywords:

Sample Size, Statistical Power, Response Variation, Mean Difference, Type I error, Type II error, Mean Milk Production.

SAMPLE SIZE CALCULATIONS FOR STRAY VOLTAGE MILK REDUCTION STUDIES

S.D. LeMire and D.J. Reinemann
UW Milking Research and Instruction Lab

INTRODUCTION

Milk reduction studies normally conclude that some treatment effect has either reduced or not reduced mean milk production. If a researcher concludes that a treatment has reduced milk production, a Type I error (false positive) is given as evidence for this conclusion. If an experiment is run, and a researcher fails to find a statistically significant reduction in milk output, no conclusions or inferences should be drawn from the experiment without a statement of Type II error (false negative). Type II error can be presented as one minus Type II error to obtain a statement about the power of the test. Power can be defined as the probability of concluding that there is a difference if the difference exists. Common values of a Type I error rates are 5 percent or less. Common values for Type II error rates are 20 percent or less. When experiments are designed, adequate sample sizes should be used to obtain the required power of the test.

This paper discusses some of the design considerations for making sample size calculations on mean milk reduction experiments. Three example calculations are given.

Hypothesis: Most mean milk reductions experiments have one of the two hypotheses shown in Table 1.

Table 1 Common Hypothesis for Mean Milk Reduction Experiments.

	Null Hypothesis	Alternative Hypothesis
Two tailed	Ho: $\mu_{trt} = \mu_{cont}$	H1: $\mu_{trt} \neq \mu_{cont}$
One tailed	Ho: $\mu_{trt} \geq \mu_{cont}$	H1: $\mu_{trt} < \mu_{cont}$

The notation definitions for the above table are as follows:

- Ho: Null hypothesis.
- H_1: Alternative hypothesis.
- μ_{cont}: Population mean of milk production of Control group.
- μ_{trt}: Population mean of milk production of Treatment group.

When making conclusions on these tests shown in Table 1, the possible decisions that we can make are shown in Table 2 [1. page 282].

Table 2 The possible decisions.

		True Condition	
	Hypothesis	Ho	H1
Researcher Decision	Ho	Correct 1-α	Incorrect β
	H_1	Incorrect α	Correct 1-β

Type I Error: When we reject the null hypothesis, Ho, and conclude a statistically significant difference between the means of the control and treatment groups, we state the Type I error α . The Type I error α, is the probability of a false positive. Often, when reporting test results, a measure called the P-value (or level of significance) is reported. The P-value is the minimum size of Type I error at which one can reject Ho. When we state that the type I error is less than 0.05 we say that we expect to be wrong less than 1 in 20 times when we "decide" that the two groups are different. A researcher establishes the burden of proof by stating the probability of being wrong [2.].

Type II Error: When one accepts Ho , we must also state the probability of being wrong. This is the Type II error which is denoted by beta (β) . Beta is the probability of a false negative or equivalently the probability of accepting Ho when it is not true.

Statistical Power: When we reject Ho and conclude that there is a statistically significant difference between control group and treatment group mean milk production, we look at the difference in the means and determine if measured differences are practically significant. If we found that the mean difference was 1 percent, we may conclude that even though we found a statistically significant difference, the 1 percent difference may have little practical meaning. On the other hand if we had found a 10 percent difference to be statistically significant, we might conclude that we have both a statistically significant difference and a practically significant difference.

When we fail to reject Ho, we also need to ask a similar question. For example, we might ask what chance was there to find a mean difference of 10 percent of milk production. If we had little or no chance of finding a 10 percent difference in mean milk production, one should not state that there was no difference in mean milk production based on the failure to reject Ho. The probability of detecting a given difference when the difference exists is called the power. Power is equal to 1-β (1 - Type II error). If a researcher fails to reject Ho, and calculates that he/she has little or no power (ie. less than 50 %) to find a practically significant difference, then failure to reject Ho should not be taken as support of either hypothesis.

To avoid designing inconclusive studies, the researcher needs to design experiments with adequate power. The power of the test depends on the values of alpha, effect size, response variability, and sample size.

Steps for Calculating Sample Size: The following material will lay out the basic steps a researcher should use in calculating sample sizes for mean difference experiments.

The Main Question: Any research Project should start with one well-defined question [3. page 11]. This question should be the primary motivation for the physical research. The question should be specific and stated in advance of the analysis. Other ancillary questions may be asked and considered, but the researcher should determine the sample size (and consequentially the cost) of the research based on the most important question. If the researchers have a number of questions that they want to answer for a given research project, sample size calculations should be done for each questions and the largest sample size should be used. Then, the most important question is now the question with the largest sample size. This is because the cost of the experiment now has to be justified based on the importance of this question.

The Main Response Variable: Integral to the main question is the main response variable. The main question should mention the main response variable.

The Effect Size: Most stray voltage research depends on a change in a response variable. As strange as it sounds, it is impossible to prove something is the same--this is because as the difference between two measurements goes to zero, the sample size goes to infinity. Typically, the size of the effect is the mean difference in the response variable. The main question should state the main response variable and the difference that the researcher is looking for in it. In this paper, we are defining effect size as the mean difference between the two groups of data. In some text books, the effect size is defined as the mean difference between the two groups divided by the standard deviation or the standardized effect size.

Selecting a Design: The intent of this paper is not to explain all possible designs, but to give the researcher a feel for the necessary steps that they should take to do sample size calculations for stray voltage experiments. However, the type of design used can have a significant effect on the sample size needed to conduct a successful experiment. Paired and Independent Two sample designs will be used for examples.

We have not included an example with a one-sided test. One-sided tests should be used sparingly. The problem with one-sided tests is that if the mean difference goes the opposite direction than anticipated, it is considered unethical to calculate P- values for the other side of the test or for a two-sided test. In some cases, the one tailed test could be applied, however, a strong justification that the response variable cannot change in one direction (e.g., cannot decrease) or has no practical meaning is required, and this determination must be made before the experiment is performed.

Estimate Response Variability: Estimating the variability of the main response can be a difficult task. One of the best methods is to use actual data. Data from past experiments can be used or trial experiments can be run. One important thing to remember is that your sample size is just an estimate. If the variability of the response that you used for the main questions increases, you will end up with less power for a given sample size than you had predicted. This is why some researchers design experiments with more power than was originally thought necessary in case the actual experimental variation is more than estimated.

Set Type I and Type II Error Rates: The accepted design values for most experiments are a 5 percent Type I error and less than 20 percent Type II error. It is often a good idea to base sample sizes on a Type II error rate of less than 20 percent. This is because unexpected errors will lower the power of the final test.

Sample Size Estimate: The formula for the sample size equation is given below. When this equation is used for independent sample size calculations, it assumes that equal sample sizes will be used for both the control and treatment groups. Equal sample size experiments are statistically the most efficient. However, equal sample size experiments may not be the most cost effective--for example, the cost to treat cows might be significantly more expensive than to obtain control cows. For information on unequal sample size calculations see [4. page 98].

When two independent group designs are being used, the N is the sample sized needed for each group. When doing a paired test, N/2 is the number of pairs needed. If the paired design uses the same unit (in our case cow) for both the control and treatment observations, N/2 is then the number of cows needed. For both types of pairing, N is the number of observation.

$$N = \frac{2(Z_{1-\alpha/2} + Z_{1-\beta})^2 S_x^2}{(\overline{X}_1 - \overline{X}_2)^2}$$

Where: N: Number
$Z_{1\text{-alpha}/2}$: Z for Type I error
$Z_{1\text{-beta}}$: Z for Type II error
$X_{bar\ 1} - X_{bar\ 2}$: Difference in Means

Common Z values for Type I and II errors are given in Table 3.

Table 3 Common Z alpha and Z beta values used in sample size calculations.

Type I Error	**$Z_{1\text{-alpha}/2}$**	**Type II Error**	**$Z_{1\text{-beta}}$**
0.20	1.282	0.20	0.842
0.10	1.645	0.10	1.282
0.05	1.96	0.05	1.645
0.01	2.576	0.01	2.326

Most of the statistical comparisons done for mean milk studies will use the t.test. Since the sample size estimate is based on the normal distribution, the actual power will be underestimated.   To avoid this underestimation at the Type I error rate of 0.05 add 2 pairs (Four cows if you are pairing cows, or 2 cows if you are using individual cows as their own comparison).   For independent sample size calculations add 1 cow to each group [5. page 104]. To make this simple, we will just add 2 samples to the estimate sample sizes.

What If Your Sample Size Is Too Large: If your sample size is too large, you can consider doing one of the following to reduce sample size:

> Reduce Response Variability: This may be achieved by changing methods to reduce variability from sources other than the treatment. For example mean milk studies, this might be accomplished by assuring more consistent milking routine and milkout of the cows.
>
> Look for a Larger Difference: Large differences between control groups and treatment groups require small sample sizes to find them if these differences really exist.
>
> Choose a Different Design: Paired designs can reduce the sample size of an experiment drastically if the pairing of the cows or the observations from a single cow are highly correlated.   If there is zero or negative correlation between the pairs, a paired design will not reduce the required sample size over independent tests. In mean milk studies, pairing observation on a single cow can produce high correlations and is often the most effective design.

You can also ask a different question or pick a different response variable.

EXAMPLE OF INDEPENDENT TWO SAMPLE COMPARISON SAMPLE SIZE CALCULATION FOR A FIELD STUDY OF ROLLING HERD AVERAGE

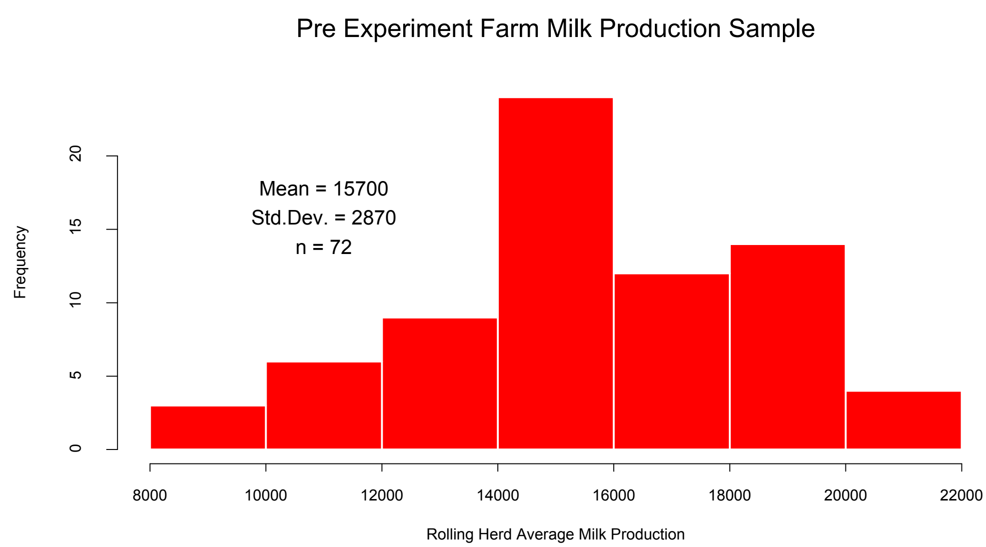
The following is an example of the steps to complete a sample size calculation for an independent experimental design.

The Main Question: Say that we wanted to study the effect of isolation of farm neutral systems on milk production on farms that have been identified as having stray voltage in the cow contact area of over 0.5 volts.   Therefore, we might ask a question like, ”How much does electrical isolation affect milk production?”. We will need to refine this question after we decide on a main response variable and determine its variability.

The Main Response Variable: The next step is to determine what the response variable will be. Examples of possible measures include, daily milk production per cow, ME305 on sample day (to reduce variability from other sources), or rolling herd average on the study farms could be used. The choice of the response variable may be influenced by when, and over what period the data is collected. If the response is expected to be affected very quickly by the stimuli, a response variable which responds quickly is desired. If the response is expected to occur slowly and continue for an extended period of time, a response variable which integrates over time may be desired. Since rolling herd average is an easily obtained and standardized measure of milk production on a given farm, we might decide to use this as the main response variable.

The Effect Size: After stating the question in general terms and determining what our main response variable will be, we now need to define a minimum important change or difference in the variable. Say that we consider that a 10 percent change in milk production to be a significant change.

Estimate Response Variability: We now have to estimate the variability in the main response variable. We can do this in a number of ways. It can be done from past records or by taking a current sample of measurements in the projected study area. For example, say that we take a random sample of farms and obtained the data shown in Figure 1 below. We could use the standard deviation of 2870 from this sample to estimate the standard deviation that might occur in the actual study.



Refine Main Question: At this point we can restate the main question. For example, we could ask, "Does electrically isolating dairy farms with stray voltage over 0.5 volts in the cow contact area change mean rolling herd average by 10 percent?".

Set Type I and II Error Rates: The commonly accepted Type I error is usually 5 percent. Acceptable Type II error is usually less than 20 percent. If this study was going to be run in the summer where a large number of heat stress days might be anticipated, one might expect an increased variability. One way to protect for unexpected events that increase variability would be to lower the Type II error in the sample size equation. In this example, a Type II error of 20 percent will be used. The power for this sample size will then be 1-.020= 0.80 or 80 percent.
Sample Size Estimate: We are now ready to make the sample size calculation. In our case we wanted to detect a 10 percent change in mean rolling herd average. To calculate the difference,

or what we are calling the effect size, we just multiply the mean of our sample by 0.10 (15700*.01=1570).

$$N = \frac{2(Z_{1-\alpha/2} + Z_{1-\beta})^2 S_x^2}{(\overline{X}_1 - \overline{X}_2)^2}$$

Where:	N:	Number of farms per group
	$Z_{1\text{-alpha}/2}$:	Z for 1.96 Type I error
	$Z_{1\text{-beta}}$:	Z for 0.84 Type II error
	$X_{\text{bar } 1}$ - $X_{\text{bar } 2}$:	Difference in Means
	S_x :	Standard deviation of independent sample or of differences

Z 1-alpha/2	: 1.96
Z 1-beta	: 0.84
$X_{\text{bar } 1}$ - $X_{\text{bar } 2}$	: 1570
S_x	: 2870

$$54.2 = \frac{2(1.96 + 0.84)^2 * 2870^2}{(1570)^2}$$

Therefore, we would need 53 + 2 farms in the control group and 53 + 2 farms in the treatment group. By adding two to each group, we have adjusted for the use of a t.test in the comparison.

When calculating sample size estimates for proposed studies, it is often productive to produce a table with estimated sample sizes for given effect sizes. Table 4 below shows sample sizes for varying effect sizes. As one can see, sample size increases drastically when one wants to find small differences between control and treatment groups.

Table 4 Sample size estimates for study that uses rolling herd averages as main outcome.

For Two Independent Groups with 0.05 Type I error and 0.20 Type II error			
Percent difference	Effect Size lbs. milk	Number of Farms per Group	Total # of Farms
1%	157	5244	10488
2%	314	1314	2628
3%	471	586	1172
4%	628	332	664
5%	785	212	424
10%	1570	55	110

EXAMPLE OF INDEPENDENT TWO SAMPLE COMPARISON SAMPLE SIZE CALCULATION FOR CONTROLLED MILK REDUCTION STUDIES

Say that we wanted to run a controlled experiment to investigate the effects of stray voltage on dairy cow milk production. If we wanted to do an independent test, we could do the following.

The Main Question: Does placing voltage on the water bowl of a cow change its milk production?

The Main Response Variable: For the main response variable, we will use the daily milk production in pounds of milk.

The Treatment: If 0.5 volts put on a water bowl is considered a problem, we might want to apply twice that amount to determine if milk loss will occur.

The Effect Size: Say that we believed that the average dairy farm had a profit margin of about 5 percent. If the presence of stray voltage could reduce milk production by 5 percent, it would be considered a serious problem. We will estimate the sample size to detect a 5 percent difference in mean milk production.

Estimate Response Variability: We now have to estimate the variability in the main response variable. For this example, we will use five days of daily milk weights from 6 cows. To obtain an estimate for the variability for the daily milk weights, we will take the average for the five days that we have for each cow. For a mean milk study, you will only obtain one or two means from a given cow. One if you are running an independent study and two if you are running a paired study. The reason for this is that daily milk weights are serial correlated within a cow. If a variance was calculated on daily milk weight of one cow and used to estimate a sample size for a study on mean milk change for other cows, the sample size would likely be too small. This is because the serial correlation would have caused the researcher to underestimate the true variability.

Now that we have the means of the six cows, we will calculate the average and the variance of the individual cow means. Since we are doing an independent test, we can use this in the sample size equation. Table 5 shows the data used to estimate the variability of mean milk weights for the sample size calculation. These data are from the control phase of an actual stray voltage experiment.

Table 5 Example of six cows daily milk weights over five days.

	Cow 1 lbs. milk/day	Cow 2 lbs. milk/day	Cow 3 lbs. milk/day	Cow 4 lbs. milk/day	Cow 5 lbs. milk/day	Cow 6 lbs. milk/day
Day 1	66	73	58	91	86	101
Day 2	68	67	52	85	80	92
Day 3	72	72	57	91	81	97
Day 4	73	76	60	98	82	101
Day 5	70	74	51	90	78	99
Mean	69.8	72.4	55.6	91	81.4	98

The average of the daily mean milk weights is then 78.0 lbs. and the sample standard deviation is 15.4 lbs.

Refine Main Question: We can now state the main question. The question could be, "Does exposing dairy cows to 1 volt on the water bowl change milk output by 5 percent?".

Set Type I and Type II error Rates: The standard Type I error is usually 5 percent. Twenty percent is generally the minimum Type II error used to estimate a sample size. Since the estimate for the variability came from cows that were not receiving any treatment, and the treatment might increase the variability in milk output for a given set of cows, it is recommended for this test that a Type II error of something less than 20 percent be used. Because of this potential increase in mean milk variability, a Type II error of 10 percent will be used for 90 percent power.

Determining Sample Size Estimate: We are now ready to calculate the sample size.

$$N = \frac{2(Z_{1-\alpha/2} + Z_{1-\beta})^2 S_x^2}{(\overline{X}_1 - \overline{X}_2)^2}$$

Where:

N:	Number of cows per group
$Z_{1\text{-alpha}/2}$:	1.96 Z for type I error
$Z_{1\text{-beta}}$:	1.282 Z for Type II error
$X_{\text{bar }1}$ - $X_{\text{bar }2}$:	Difference in Means for 5 percent

Z 1-alpha/2	: 1.96
Z 1-beta	: 1.282
$X_{\text{bar }1}$ - $X_{\text{bar }2}$	: 3.9
Sx	: 15.4

$$327.8 = \frac{2(1.96 + 1.282)^2 * 15.4^2}{(3.9)^2}$$

Therefore, we would need 328 + 2 farms in the control group and 328 + 2 farms in the treatment group. By adding two to each group, we have adjusted for the use of a t.test in the comparison.

Table 6 gives sample size estimates for varying effect sizes for the independent sample size calculation. Two samples have been added to each sample N to account for a t.test comparison.

Table 6 Sample size estimates for study that uses daily milk productions as main outcome.

For Two Independent Groups with 0.05 Type I error and 0.10 Type II error			
Percent difference	**Effect Size lbs. milk/day**	**Number of cows per group**	**Total # of Cows**
1%	**0.8**	**7794**	**15588**
2%	**1.6**	**1952**	**3904**
3%	**2.3**	**946**	**1892**
4%	**3.2**	**490**	**980**
5%	**3.9**	**330**	**660**
10%	**7.8**	**86**	**172**

The Estimated Sample Size is Too Large: This experiment has been designed to find a 5 percent difference in mean milk production between two groups of cows. The sample size calculation tells us that we need 330 cows in both the control group and the treatment group. At the University of Wisconsin Madison barn, we have about 90 milking cows. Therefore, it is doubtful that this experiment could realistically be run. The first option is to try to reduce the variability of the main response. In this case, we would have to look at the milking system. However, even if we could cut the variability in half from a standard deviation of 15.4 to 7, we would still need a sample size of 70 cows in each group for a total of 72 study cows.

A second option would be to look for a larger effect size. If we would settle for a 10 percent difference, we would need 86 cows in each group for a total of 172 cows. This is still too many cows. A third option would be to look at a different design. If we used a paired test, we could take advantage of the correlation between daily milk weights and eliminate the between cow variation in this study.

PAIRED TEST EXAMPLE

Say that we still wanted to run a controlled experiment to investigate the effects of stray voltage on dairy cow milk production. The sample size for the independent groups design was too large to enable us to run the experiment. Since we can use each cow as its own control, we will now explore a paired design.

The Main Question: "Does exposing dairy cows to 1 volt on the water bowl change milk output by 5 percent?" The main question has not changed.

The Treatment: If 0.5 volts on a water bowl is considered a problem, we might want to apply twice that amount to determine if milk loss will occur. In the independent groups design, one group would have received the voltage on the water bowl and the other group of control cows would have been kept under similar conditions expect they would have not received shock on the water bowl.

In the paired experiment, each cow will act as its own control. To avoid confounding with time and other unknown influences, it is often best to run a switch back design where half the cows receive treatment first then the control phase second and the second half receive the control first and the treatment second. A period of washout between treatment and control should be allotted if residual from the treatment could affect the cow after it has come off treatment. This period of time must be determined by the researcher. The cross over or switch back design should not effect your sample size estimate.

The Main Response Variable: For the main response variable for the cross over design will be the mean difference between the control phase and the treatment phase for individual cows.

The Effect Size: We want to find 5 percent change in milk production (78 lbs * 0.05 = 3.9 lbs).

Estimate of Response Variability: We now have to estimate the variability in the main response variable. For a cross over design, We need to know the variability of the mean difference between treatment phases of the same cow. The equation for the variance of the differences is shown below [7].

$$\sigma^2_{x1-x2} = \sigma^2_{x1} + \sigma^2_{x2} - 2Cov(X_1, X_2)$$

If we assume that the variance of the control phase and the treatment phase are the same and that the correlation between the mean milk output during treatment and during control is zero, the variance of the differences is then:

$$\sigma^2_{x1-x2} = 2\sigma^2_{x1}$$

This would then double the variance of the test sample and increase the sample size considerably. Therefore, it is important to have an estimate of the correlation between the control and treatment phases of the experiment.

One way to do this is to run small trials and determine the correlation experimentally. We will use the same trial data as used in the independent calculation. See Table 7 for the data used for the control phase of the sample size calculation. Table 8 shows the data for the treatment phase of the experiment. Table 9 shows the means for individual cows for the control and treatment phases. The variances of the differences will be calculated from the means in Table 9.

Table 7 Example of six cows daily milk weights over five days of control.

	Cow 1 lbs milk/day	Cow 2 lbs milk/day	Cow 3 lbs milk/day	Cow 4 lbs milk/day	Cow 5 lbs milk/day	Cow 6 lbs milk/day
Day 1	66	73	58	91	86	101
Day 2	68	67	52	85	80	92
Day 3	72	72	57	91	81	97
Day 4	73	76	60	98	82	101
Day 5	70	74	51	90	78	99
Mean	69.8	72.4	55.6	91	81.4	98

Table 8 Example of six cows daily milk weights over five days of treatment.

	Cow 1 lbs milk/day	Cow 2 lbs milk/day	Cow 3 lbs milk/day	Cow 4 lbs milk/day	Cow 5 lbs milk/day	Cow 6 lbs milk/day
Day 1	60	65	58	76	70	95
Day 2	52	60	52	79	71	99
Day 3	70	74	55	94	83	101
Day 4	64	69	51	90	71	92
Day 5	65	67	50	89	76	95
Mean	62.2	67	53.2	85.6	74.2	96.4

Table 9 Mean milk weights for the control and treatment phases of pre experiment test.

	Cow 1 lbs milk/day	Cow 2 lbs milk/day	Cow 3 lbs milk/day	Cow 4 lbs milk/day	Cow 5 lbs milk/day	Cow 6 lbs milk/day
Means control phase	69.8	72.4	55.6	91	81.4	98
Means treatment phase	62.2	67	53.2	85.6	74.2	96.4
Mean differences	7.6	5.4	2.4	5.4	7.2	1.6

One could now take the variance of the two samples and the covariance between the means of the control phase and treatment phase and calculate the variance of the differences. However, this is not necessary when you have the actual data. You can just calculate the standard deviation of the mean differences and use this in the sample size equation. Both methods will give you the same answer. The standard deviation of the mean differences from Table 9 is 2.5 lbs milk.

Set Type I and Type II Error Rates: The Type I error for this experiment will be set at 5 percent. The Type II error will be set at 10 percent.

Determining Sample Size Estimate: We are now ready to calculate the sample size. N represents the number of observations [5]. In our case, we will get two mean observation per cow. Therefore, we will need N/2 cows.

$$N = \frac{2(Z_{1-\alpha/2} + Z_{1-\beta})^2 S_x^2}{(X_1 - X_2)^2}$$

Where:	N:	Number of Observations
	N/2	Number of cows
	$Z_{1\text{-alpha/2}}$:	1.96 Z for Type I error
	$Z_{1\text{-beta}}$:	1.282 Z for type II error
	$X_{bar\ 1} - X_{bar\ 2}$:	Difference in Means for 5 percent

Z 1-alpha/2	: 1.96
Z 1-beta	: 1.282
$X_{bar\ 1} - X_{bar\ 2}$	: 3.9
Sx	: 2.5

$$8.6 = \frac{2(1.96 + 1.282)^2 * 2.5^2}{(3.9)^2}$$

Table 10 shows the sample size estimates for studies that use daily milk productions as main outcome. Each sample size calculation had 2 added to it to correct for the t.test comparison.

Table 10 Sample sizes for a paired study.

Paired Test with 0.05 Type I error and 0.10 Type II error		
Percent difference	**Effect Size lbs milk/day**	**Total # of Cows**
1 %	**0.8**	**105**
2 %	**1.6**	**28**
3 %	**2.3**	**14**
4 %	**3.2**	**8**
5 %	**3.9**	**6**
10 %	**7.8**	**3**

By pairing and taking advantage of inter cow correlations we have reduced the number of cows needed from 660 for the independent test to 6 for the paired test. It should be noted that when sample sizes are as small as the paired sample above, a researcher should be cautious in the inferences they draw to the greater population of animals.

SUMMARY AND CONCLUSION

The material in this paper has covered a few of the basic methods of sample size calculations as they apply to mean milk studies.

No experiment that will use substantial resources should be designed without consideration of sample size. An experiment that is designed and executed without the chance of finding a meaningful difference is poor science.

REFERENCES

1. Hayes HW : Statistics. 5th ed. New York: Harcourt Brace College. 1994.

2. Cohen, J: Statistical Power Analysis for the Behavioral Sciences. New Jersy: Lawrence Erlbaum Associates, 1988.

3. Friedman, L., M. , Furberg, C., D., DeMets, D., L: Fundamentals of Clinical Trials. Massachusetts: PSG Publishing Company, 1985.

4. Lachin JM: Introduction to Sample Size Determination and Power Analysis for Clinical Trials. Controlled Clinical Trials 2, 93-113 1981.

5. Snedecor GW, Cochran WG: Statistical Methods, 8th ed. Ames: Iowa State University Press. 1989.