Is Off-Target Performance a Real Problem? How You Can Help Your Clients Avoid Being Tricked

William E. Marsh, Alvaro J. Soler
Mark L. Kinsel and **Jeffrey K. Reneau**
University of Minnesota
St Paul, MN

Abstract

The veterinary profession continues to make greater use of production databases in serving the dairy industry. Increasingly sophisticated database management systems allow producers and practitioners to consider the biological, statistical, and economic implications of management interventions. Frequency distributions and simple descriptive statistics can be very helpful in understanding patterns that underlie commonly used measures of dairy herd productivity.

Once underlying patterns are understood, statistics can be used to set feasible production targets and confidence intervals to accommodate the variability inherent in biology. This is particularly appropriate for smaller herds or when one wishes to analyze performance over short time periods. Confidence intervals and interference levels can be adjusted to lessen the chances of erroneous conclusions being drawn when sample sizes are small. Consequences of drawing the wrong conclusions are implementing changes when no intervention is warranted, or failing to act quickly enough in response to falling productivity.

Practicing veterinarians can apply these simple-to-use practical techniques which help avoid the trap of confusing normal variation with real changes in productivity.

The computer software programs used in the dairy industry to monitor productivity today are so sophisticated that users can be overwhelmed by the numerous options and reports. Increasingly, producers are turning to their veterinarians to help them interpret the production information generated by their database programs. Properly applied, this information can be enormously valuable in identifying and diagnosing management problems. As in any biological system, however, there will be some normal variability in the production parameters for a dairy herd. The challenge facing the practitioner is to differentiate between production values that represent normal biological variation and those that truly warrant interference.

Statistics help us interpret production values in meaningful ways. By understanding and using simple statistics, one can determine:

- How confident one can be that the data for a subpopulation is not different from the long-range goals for the productivity of the entire herd (i.e., the **confidence interval**); and
- at what point productivity falls outside of this confidence interval, meaning that productivity is truly off-target and interference is warranted (i.e., the **interference level**).

Of course, deciding where to set an interference level can never be based purely on statistics. Since management changes usually involve an expenditure of time or money, one must always assess the financial significance and riskiness of continuing to operate at or below the chosen interference level. Because the productivity information on which we base management decisions is never complete and certain, we always run the risk of "diagnosing" a management problem that doesn't really exist (Type I error), or failing to detect an emerging management problem that truly does exist (a Type II error). These errors are particularly likely when management decisions must be based on data from a small sample size or short time period, because the fewer the number of observations, the more difficult it is to distinguish between real differences and normal biological variation. However, using statistics to compensate can greatly reduce the likelihood of such errors.

Consider this case

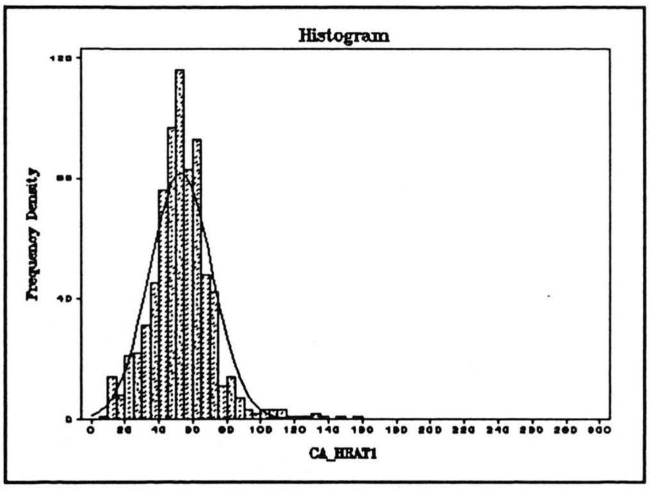
One month's reproductive performance data for a herd show a calving to first service interval of 82 d, and a first service conception rate of 52.5%. Given the long-run reproductive performance goals for calving to first service interval (target: 75 D; interference level: 79 d) and first service conception rate (target 60%; interference level 55%), do last month's figures truly indicate a problem of poor reproductive performance or not?

Normally-distributed variables

Our immediate task is to determine whether there is a problem with reproductive management. The next step, then is to plot a frequency distribution of the data for the last (say) 100 calving to first heat intervals to give a pictorial representation of the underlying distribution. We recommend that you plot data for at least

100 observations because only when you have an adequate number of observations will you realistically be able to assess whether the distribution it follows is normal (Fig. 1). Frequency distributions are obtained by plotting the range of possible values (in this case, calving to first heat intervals) along the horizontal (X) axis against the number or proportion of the population that falls within each interval on the vertical (Y) axis (in this case, percentage of the last 750 calving to first heat intervals).

Figure 1. Frequency distribution of calving to first observed heat intervals.



Remember that your challenge is to differentiate real dips in productivity from inherent biological variability. Your second step in interpreting this client's production data, then, is to verify whether the data for the previous month is normally distributed (i.e., approximates a normal bell-shaped curve). From the plot, we can see that the **calving to first heat** intervals from the most recent 750 observations are reasonably close to being normally distributed. Observations which comprise a similar plot of **calving to first service** intervals for the same population of cows are shown in Fig. 2. While these data do not conform quite so well to the mathematical properties of the normal distribution, the shape of this distribution is close enough to normal for our purposes. On the other hand, a plot of **calving to conception** intervals (Fig. 3) shows that the distribution is definitely not normally distributed. Because of this, more sophisticated techniques than those described in this paper are required to calculate confidence intervals and interference levels for calving to conception intervals.

If our distribution is approximately normal, our next step is to determine the degree of variance in our data (how the individual values are dispersed about the mean). The classic way to measure variance is to calculate the standard deviation. The larger the variance or standard deviation, the more scattered the individual data points. Mathematically, in a normal distribution:

Figure 2. Frequency distribution of calving to first observed service intervals.

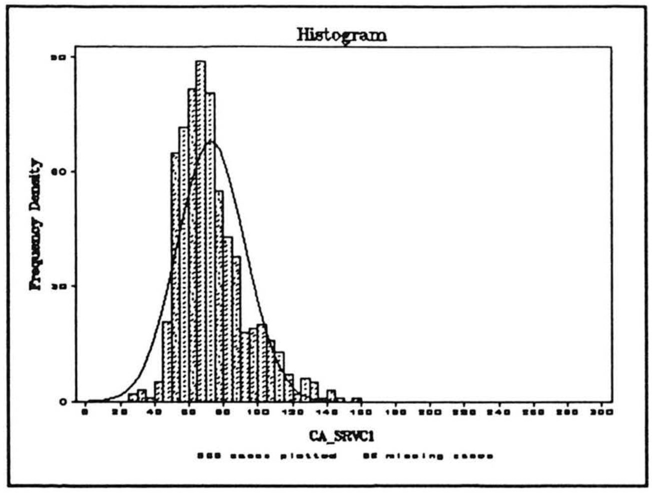
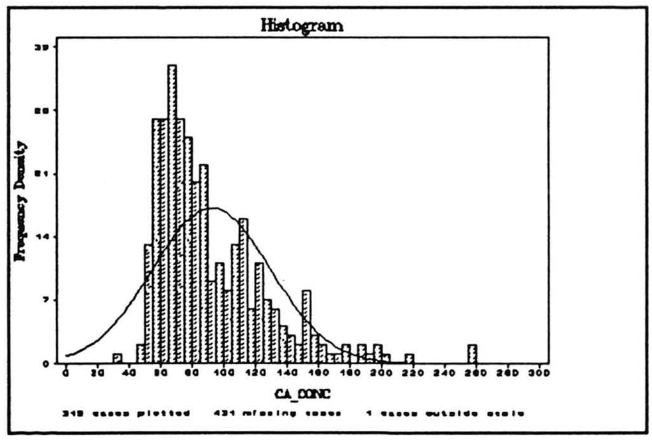


Figure 3. Frequency distribution of calving to conception intervals.



- 68% of the observations will fall within ± 1 standard deviation of the mean;
- 95% of the observations will fall within ± 2 standard deviations of the mean; and
- 99% of the observations will fall within ± 3 standard deviations of the mean.

Determining the standard deviation allows us, in turn, to calculate the confidence interval--i.e., how confident we can be that our small subpopulation (calving to first service intervals of cows that were inseminated in the previous month) accurately represents the entire herd. Essentially, the confidence interval is a mathematical expression of the relationship between the mean, the standard deviation and the sample size.

Computing confidence intervals

To compute confidence intervals, we need four pieces of information:

- The mean (our target value)
- Standard deviation
- The sample size
- A "constant"

Statisticians have computed "constant" values from the shape of Normal distribution. These "constants" are the number of standard deviations from the mean within which specified proportions of the distribution are expected to occur. For example, to compute a 95% confidence interval, we use the constant 1.96 because 95% of the observations fall within the range of -1.96 to +1.96 standard deviations.

The 95% confidence interval is calculated as:

Mean±(1.96**Standard Deviation) / √(Sample Size)*

For the calving to first service data, we use:

$$75 \pm (1.96 \times 20) / \sqrt{30}$$
$$=75 \pm 7.157$$

The result is expressed as: (67.8 < 75.0 < 82.2)

i.e. our estimate of the mean is 75 d, and there is a 95% probability that the true population mean lies between 67.8 and 82.2. For a given sample size, other confidence intervals can be calculated simply by substituting the appropriate constant into the equation:

99% C.I: 75.0 ± ((2.58 * 20) / √ 30) = (65.6 < 75.0 < 84.4)
95% C.I: 75.0 ± ((1.96 * 20) / √ 30) = (67.8 < 75.0 < 82.2)
90% C.I: 75.0 ± ((1.64 * 20) / √ 30) = (69.0 < 75.0 < 81.0)
80% C.I: 75.0 ± ((1.28 * 20) / √ 30) = (70.3 < 75.0 < 79.7)

Thus, as we move from the 99% to the 80% confidence interval, the width of the interval narrows, reflecting the decreasing (but still relatively high) probability that the true population mean lies within the calculated range.

Does this herd have a problem with calving to first service interval? The answer must be based on the confidence interval for the herd, which takes into account the sample size of the sub-population e.g., calving to first service intervals of cows that were inseminated in the previous month (Table I). The values in Table I were computed for the target calving to first service interval of 75 d that was set for the herd using the formula to generate confidence intervals. The standard deviation (20 d) used in the Table was derived by running the

Table I. Interference levels at various sample sizes, assuming a target calving to first service interval of 75 d, and a standard deviation of 20 d.

STD DEV	NUMBER OF OBSERVATIONS									
	10	20	30	40	50	60	70	80	90	100
0.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0
0.5	78.2	77.2	76.8	76.6	76.4	76.3	76.2	76.1	76.1	76.0
1.0	81.3	79.5	78.7	78.2	77.8	77.6	77.4	77.2	77.1	77.0
1.5	84.5	81.7	80.5	79.7	79.2	78.9	78.6	78.4	78.2	78.0
2.0	87.6	83.9	82.3	81.3	80.7	80.2	79.8	79.5	79.2	79.0

DairyCHAMP® STATISTICS report type under Database Applications (which will report the mean and the standard deviation). If you do not have access to Dairy CHAMP®, you may use the following approximation to calculate the standard deviation:
Standard deviation ≈ (Highest value - Lowest value) ÷ 6

To use the table:

- read down the column headed "100" to the number closest to the long-run interference level the herd manager has set. In this case, a calving to first service interval of 79 d corresponds to 2.0 standard deviations above the mean (extreme left-hand column)
- read across the 2.0 standard deviations row to the column headed with the number closest to the sample size for the period of interest (30 in this case).

The number in this table at this point is 82.3. A calving to first service interval of 82.3 d for a sample size of 30 is equally likely to be found as a calving to first service interval of 79 d for a sample size of 100 observations, as they both represent a deviation of 2.0 standard deviations from the mean, adjusted for sample size.

The table also shows that in a period with only 10 observations, a mean calving to first service interval as high as 87.6 d should be considered to be in the acceptable range, while in a period containing 50 observations, the interference level adjusts to 80.7 d. Our conclusion in this example is that an average calving to first service interval of 82 d among a sample of 30 observations is not inconsistent with a target of 75 d among 100 observations. As we cannot be sure that anything is really wrong, no intervention is warranted.

Binomial variables

We would apply the same statistical principles if our task were to investigate a problem with heat detection or conception rates (i.e. a binomial variable). Unlike the normal distribution, the binomial distribution is not usually described by a mean and standard deviation. Rather, it is based on a mathematical expression that considers the probability of a number of successes (conceptions) occurring out of a number of trials (inseminations).

Consider this case

Let's look at the same herd and determine whether it has a problem with conception rate (based on 30 inseminations per month).

How do we calculate confidence intervals for productivity parameters that follow a binomial distribution? First, we must determine whether we have an adequate number of observations to draw safe conclusions. We do so before performing the following "sample size" calculation:

Target Rate x (1 - Target Rate) x sample size must be ≥ 5.

In the case of a 60% Target Conception Rate (TCR):

(0.6 x (1 - 0.6) x minimum sample size) - 5
(0.6 x 0.4) x minimum sample size - 5
Minimum sample size = 5 / 0.24
Minimum sample size = 21

An adequate sample size would require a minimum of 21 cows inseminated per period. A 15% or 85% TCR would require a minimum of 50 cows inseminated - the sample size increases as the TCR tends towards 0 or 100%. Because of the characteristics of the binomial distribution, the smallest sample size occurs when the target value is 50%. Thus, targets that are very low, e.g., 5% calf mortality rate, will require larger populations if the Normal Approximation is to be used (105 calves!)

Once we have an adequate number of observations, we can perform a simple calculation called the Normal Approximation to compute our confidence interval. Using the Normal Approximation, the approximate 80% confidence interval is determined by:

TCR + 1.28 x (TCR x(1 - TCR) x sample size)

(TCR is target conception rate.)

Example: to calculate the 80% confidence interval for an 60% conception rate with a sample size of 50 inseminations:

$0.60 \pm 1.28 \times ((0.60) \times (1 - 0.60) \div 30)^{0.5}$
0.65 ± 0.089
80% confidence interval: (0.511 < 0.60 < 0.689)

. . . indicating that the interference level is 51.5%.

Thus, if the long-run (100 females mated) targets and interference levels for conception rate were 60.0% and 55.1% respectively, then intervention would be warranted only if the conception rate fell below 51.5% in a sample of 30 inseminations (Table II). Table III may be used in a similar fashion where targets for either conception rate or heat detection efficiency are more modest.

Table II. Interference levels at various sample sizes, assuming a target conception rate of 60%.

STD DEV	NUMBER OF OBSERVATIONS									
	10	20	30	40	50	60	70	80	90	100
0.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0
0.5	••••	54.5	55.5	56.1	56.5	56.8	57.1	57.3	57.4	57.6
1.0	••••	49.0	51.1	52.3	53.1	53.7	54.1	54.5	54.8	55.1
1.5	••••	43.6	46.6	48.4	49.6	50.5	51.2	51.8	52.3	52.7
2.0	••••	38.1	42.1	44.5	46.1	47.4	48.3	49.0	49.7	50.2

Table III.

STD DEV	NUMBER OF OBSERVATIONS									
	10	20	30	40	50	60	70	80	90	100
0.0	45.0	45.0	45.0	45.0	45.0	45.0	45.0	45.0	45.0	45.0
0.5	••••	39.4	40.5	41.1	41.5	41.8	42.0	42.2	42.4	42.5
1.0	••••	33.9	35.9	37.1	38.0	38.6	39.1	39.4	39.8	40.0
1.5	••••	28.3	31.4	33.2	34.4	35.4	36.1	36.7	37.1	37.5
2.0	••••	22.8	26.8	29.3	30.9	32.2	33.1	33.9	34.5	35.1

If you do not have an adequate sample size, we recommend that you enlarge the time period of the report showing the reproductive statistics to ensure that the number reported is always based on a minimum of 30 inseminations in each period. If you do not have at least this many observations, you would need to calculate the confidence interval using a calculation called the "Exact Binomial," which is so complex it requires a computer to calculate it and thus is of limited practical value in a real-life situation.

Setting appropriate productivity targets and interference levels.

Setting productivity targets is a fairly straightforward process: consider historic performance and expectations of future productivity based on anticipated improvement, together with performance of other, comparable herds. Setting the appropriate interference level, however, is a more personal and subjective process. Some dairy producers tend to over-manage by immediately reacting to very subtle changes in productivity, while others with a more relaxed approach tend to be more willing to ride out fairly major productivity drops. Effectively, members of the latter group tend to tolerate wider confidence intervals (e.g. 95%) than over-managers, who may be inclined to interfere when productivity approaches the lower 80% confidence limit.

Ultimately, of course, determining when to interfere in the herd should take into account the economic importance of the problem. Even though experienced managers tend to ponder economic considerations when adjusting targets and interference levels, it is important to consider the financial repercussions of alternative interventions (or doing nothing), rather than interfere as a matter of course. In cases where the herd is already managed very efficiently, the cost and additional risk of interference may outweigh the potential expected benefits. As a practitioner, you will need to understand not only how to exploit statistics to help your clients interpret their productivity data, but you will also need to tailor your recommendations to the season, the particular management practices in the herd, and the personality of the producer or herd manager.

The simple procedures we have outlined (from plotting data to showing distributions, setting feasible production targets, calculating confidence intervals, to adjusting interference levels to compensate for small sample sizes) can improve effective communication between the herd manager and consultant. Properly applied, these techniques can help us to avoid becoming distracted by spurious production changes while improving our power to detect emerging problems before they cause severe economic loss.