Epidémiol. santé anim., 1997, 31-32

SOME ILLUSTRATIONS OF THE USE OF STATISTICAL PROCESS CONTROL TECHNIQUES IN MONITORING DAIRY HERD PERFORMANCE

De Vries A.[1], Conlin B.[1], Reneau J.[1], Kinsel M.[2], Marsh W.[3]

L'objet de cet article est de présenter l'utilisation de la technique SPC (Statistical Process Control) dans l'interprétation des données relatives à la gestion des troupeaux laitiers. Les systèmes d'information sur la gestion des élevages laitiers produisent une masse d'information qui se trouve souvent peu exploitée. Les données sont présentées sous forme de tableaux de moyennes calculées sur plusieurs séries temporelles. L'interprétation de ces tableaux ne permet pas de distinguer les variations aléatoires des variations associées à des causes spécifiques comme les maladies. Les techniques SPC peuvent être utilisées pour signaler rapidement tout changement anormal sur l'évolution de la production. Il existe plusieurs outils, les principaux sont les cartes de contrôle de Shewart. Le principe de ces cartes consiste tout d'abord à tracer sur un graphique l'évolution de la moyenne et des limites de contrôle qui dépendent de la variance du paramètre suivi lorsque le process est maîtrisé, puis sur le même graphique tracer la courbe de l'évolution des moyennes et des écarts types estimés à partir des échantillon du contrôle. Il existe plusieurs tests statistiques puissants permettant de localiser rapidement des déviations significatives de l'évolution des résultats au cours du temps.

INTRODUCTION

Dairy herd management information systems generate an abundance of output. Typically, mean values of parameters are reported over a series of consecutive time periods and are presented in a tabular format. It is often impossible to distinguish between inherent, random variation in the process and variation due to special causes, such as disease or management lapses, when interpreting tabulated values. Both sources of variation contribute to changes in mean values between periods.

In order to gauge performance, mean values from the latest time period are usually compared against a series of pre-determined target values and performance values below or above which productivity should not fall. Determination of appropriate target values and performance values is typically an arbitrary process. There is usually a lack of a sound statistical basis for setting these values. Discrepancies between actual and target values can only be interpreted as an incentive to search for ways of improving the production process. However, such targets are futile without a plan as to how to achieve this objective (Montgomery, 1991; Deming, 1986).

Statistical Process Control (SPC) techniques have been developed as methods of establishing and maintaining quality in production processes (DeVor et al., 1992). Shewhart control charts are one of seven SPC techniques that are widely used. Target values in control charts are generated from the process itself when it is in control and therefore have a sound statistical basis to compare actual parameter values with.

SHEWHART CONTROL CHARTS

Shewhart control charts are graphical displays of statistics plotted according to the order of their observation. Control charts provide a means for differentiating real changes (due to special causes) in a production process from inherent random variation (due to common causes only).

Dairy herd managers who make decisions are prone to making errors. A Type I error is made when inherent, random variation (common cause) is interpreted as an unexpected real change in the process. As a result, a non-existent problem is diagnosed, and an unnecessary management intervention may be implemented. Type II errors occur when the decision-maker fails to identify a real change (special cause) in the process quickly enough. By employing simple statistical and graphical techniques, control charts provide a series of decision support tools that greatly reduce the probability of committing either Type I or Type II errors.

Several types of control charts exist. The appropriate choice depends upon the nature of the data that is monitored. Generally, the mean and the variation, measured by the range, in the process are of interest. Lines indicating means, and statistically-derived control limits provide the framework for the control chart. In order for the control limits to be meaningful, it is essential that they be derived from a series of observations during which the process was known to be in control. A process in control is subject to common cause variation only. Observations used to construct initial control charts that are out of control should be excluded from the final calculations of the control limits. Once the framework of the chart has been established, consecutive values are plotted on the chart as they are observed and recorded over time.

Several tests exist which calculate the probabilities of each plotted observation. Small probabilities indicate that the process is likely affected by special causes. Such values are flagged in the control charts and indicate that special investigation is needed. Nelson (1984) described eight tests that can be used in Shewhart control charts

[1] University of Minnesota, Department of Animal Science, 1364 Eckles Avenue, St. Paul, MN 55108 U.S.A.
[2] University of Idaho, Caine Veterinary Research and Teaching Center, Caldwell, ID 83605, U.S.A.
[3] University of Minnesota, College of Veterinary Medicine, 1988 Fitch Avenue, St. Paul, MN 55108, U.S.A.

with normally distributed parameters. For example, an observation beyond one of the control limits is probably the result of a special cause.

APPLICATION AND EXAMPLES

Shewhart control charts can be applied in monitoring dairy herd management parameters. As the distribution of daily bulk milk tank somatic cell count (SCC) data is approximately normal (Schukken, 1991), control charts can be applied to monitor and control bulk tank SCC levels. For individual data, X (mean) and R_m (moving range) control charts are used. The moving range is calculated to get an estimate of the variability in individual data.

Daily SCC levels were collected from a dairy farm and means and control limits for both the average SCC level and the range were calculated and plotted in figures 1 and 2. The range was calculated as the difference between two consecutive observations. Control limits were set at three times the estimated standard deviation. Mean was 409,000. Average range was 44,400. Standard deviation was calculated as 44,400/1.128=39,400. Lower control limit for the mean was 409,000-3*39,400=290,800. Upper control limit for the mean was 409,000+3*39,400 =527,200. Lower and upper control limits for the range were respectively calculated as 0*44,400=0 and 3.267*44,400=145,100 (see DeVor et al., 1992). Next, 57 consecutive daily SCC levels and the moving ranges were plotted in the appropriate control charts.

Figure 1
X control chart for mean of daily bulk tank somatic cell count data

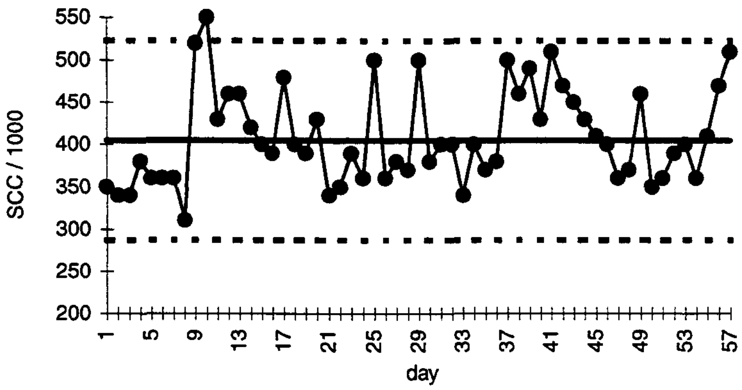
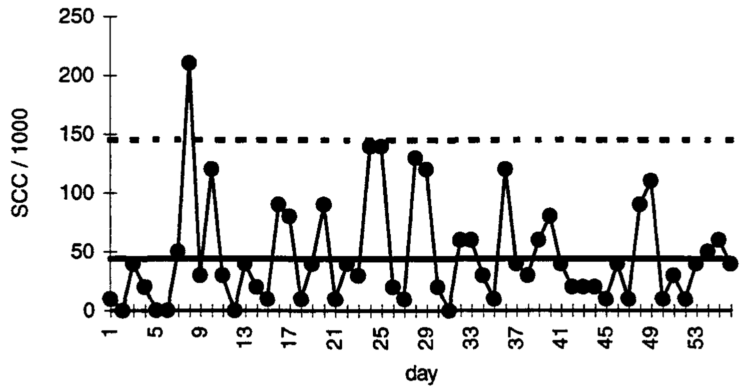


Figure 2
R_m control chart for range (n=2) of daily bulk tank somatic cell count data

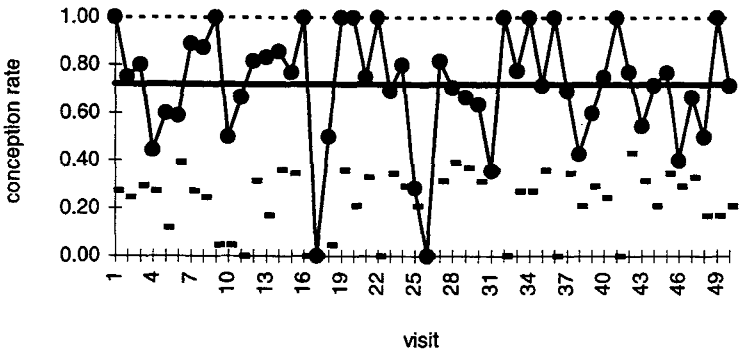


The X and R_m control charts show several observations out of control when the tests are applied. The special cause(s) that generated these observations should be investigated and removed.

Control charts can also be applied to binomially distributed parameters, like conception rate at farm visits to diagnose pregnancies in cows that were recently bred. Binomial distributions are skewed if the probability of success is not 0.5. Control limits can still be calculated, but only tests for runs and trends should be employed. Sample sizes (number of cows diagnosed per visit) need to be approximately equal. When sample sizes are not approximately equal, only observations that exceed the control limits can be regarded as out of control. Another approach is to use the normal approximation to the binomial distribution. This approach may not be feasible when too few cows are diagnosed per visit, however.

Pregnancy diagnosis data from visits to one farm was taken and a P control chart was constructed (figure 3). Average conception rate was 0.72. Standard deviations depend on number of cows diagnosed per visit (n) and were calculated as $\sqrt{(0.72*(1-0.72)/n)}$. Control limits were set at three times the estimated standard deviations from the mean with a maximum of 1.00 and a minimum of 0.00.

Figure 3
P control chart for conception rates found at farm visits for pregnancy diagnosis



The P control chart does not show clear evidence that conception rates changed significantly at a certain time or period. Sensitivity is low in this case because the number of cows diagnosed per visit is small.

It is tempting to add data from several visits together in order to get tighter control limits. The problem is that rational subgroups (visits) should be collections of individual observations whose variation is attributable only to common causes. Special causes should only arise between visits (DeVor et al.,1992). When pooling data from several visits together, special causes will more easily be introduced within each sample. Some cows will have a special cause reason for not conceiving while others do not conceive because of common causes. Control limits calculated from data from these pooled visits will therefore more likely be based on special causes as well.

DISCUSSION

X and R_m control charts may not be very sensitive for detecting shifts in the process mean level or amount of process variability, particularly when the shift in the mean or in the variability is relatively small. Another weakness of the X and R_m charts is that these two charts are not independent from each other. Exponentially weighted moving average (EWMA) control charts for both the mean and the standard deviation are more sensitive to detect trend lines in the data. A similar chart that can be effectively used for process control purposes is the cumulative sum (CUSUM) control chart (DeVor et al.,1992).

A process that is in control is considered to be in its most economical mode. However, a process in control may still not be profitable. Special causes make the economic viability of the process to be jeopardized. However, it is the attack on the common causes that is also important, since 80 to 85% of all the problems encountered are of this nature (DeVor et al., 1992).

The data should be adjusted for known causes of variation as much as possible in order to increase sensitivity of the control charts. For example, conception rates should be adjusted for known seasonal effects. In case special causes are signaled, the effects that were adjusted for can be excluded from investigation.

Another concern is the possibility of autocorrelation between consecutive observations. Tests to signal special causes require independently distributed data. Time series models could be used to remove the autocorrelation from the data. The residuals will then be independent such that the tests can be applied (Montgomery, 1991).

REFERENCES

Deming, W.E., 1986. Out of the crisis. Massachusetts Institute of Technology, Center for Advanced Engineering Study, Cambridge, MA.

DeVor, R.E., Chang, T.H., Sutherland, J.W., 1992. Statistical quality design and control. Macmillan, New York, NY.

Montgomery, D.C., 1991. Introduction to statistical quality control. Wiley, New York, NY, 2nd ed.

Nelson, L.S., 1984. The Shewhart control chart - tests for special causes. Journal of Quality Technology 16:237-239.

Schukken, Y.H., Stassen, E.N., Noordhuizen, J.P.T.M., Verheijden, J.H.M., and Brand, A., 1991. Decision making on disease occurrence in herd health and production management programs. Proceedings of the 6th ISVEE symposium, Ottawa, Canada, 12-16 August.