Center for Quality and Productivity Improvement
University of Wisconsin
610 Walnut Street
Madison, Wisconsin 53705

(608) 263-2520
(608) 263-1425 FAX
quality@engr.wisc.edu

Report No. 170

Quality Quandaries*
Use of Cusum Statistics in the Analysis of Data and in Process Monitoring

George Box

October 1998

*Edited by George Box and Søren Bisgaard

CQPI
University of Wisconsin

Report No. 170

Report Series
in Quality and Productivity

Quality Quandaries*
Use of Cusum Statistics in the Analysis of Data and in Process Monitoring

George Box

Center for Quality and Productivity Improvement

University of Wisconsin–Madison, USA

ABSTRACT

The uses of the Cusum Chart are discussed. It can be used for on-line monitoring of an operating process or for post-mortem analysis of data. The Cusum statistic is discussed as a particular example of a Cuscore statistic for the detection of a signal in noise.

KEYWORDS: Cusum, Process Monitoring, Data Analysis

*Edited by George Box and Søren Bisgaard

Quality Quandaries*
Use of Cusum Statistics in the Analysis of Data and in Process Monitoring†

George Box

The uses of the Cusum Chart are discussed. It can be used for on-line monitoring of an operating process or for post-mortem analysis of data. The Cusum statistic is discussed as a particular example of a Cuscore statistic for the detection of a signal in noise.

Cusums were originally devised by Page[1] and Barnard[2] and put to extensive use both in England and the United States particularly for the monitoring of synthetic fiber manufacture. (See for example Goldsmith and Whitfield[3] and Lucas and Crosier[4]). They provide an extremely sensitive on-line procedure for detecting small shifts in the process mean; that is, for the detection of "step changes." In its simplest form, the Cusum S is an accumulated sum of the deviations $y_1 - T$, $y_2 - T$,... from the process target value T. Thus $S_t = (y_1 - T) + (y_2 - T) + \ldots + (y_t - T)$, and a Cusum chart is obtained by plotting S_t against t. If the mean is on target then S_t will simply be the sum of a series of deviations having zero mean some of which are positive and some negative, but as soon as a shift in the mean away from the target occurs, the Cusum plot exhibits a change in *slope*. This is so because as soon as an increase in the mean of d units occurs then as each new observation is obtained quantities d, $2d$, $3d$, ... are added to successive values of the Cusum..

For example, Figure 1(a) shows observations plotted on a Shewhart run chart for which the target value is T = 30. Figure 1 (b) shows the Cusum chart in which, for the same observations, the successive sum S_t of the deviations from target is plotted against t. The marked change in slope that occurs in the Cusum plot strongly suggests that a shift in mean has occurred close to observation number 16. Also,

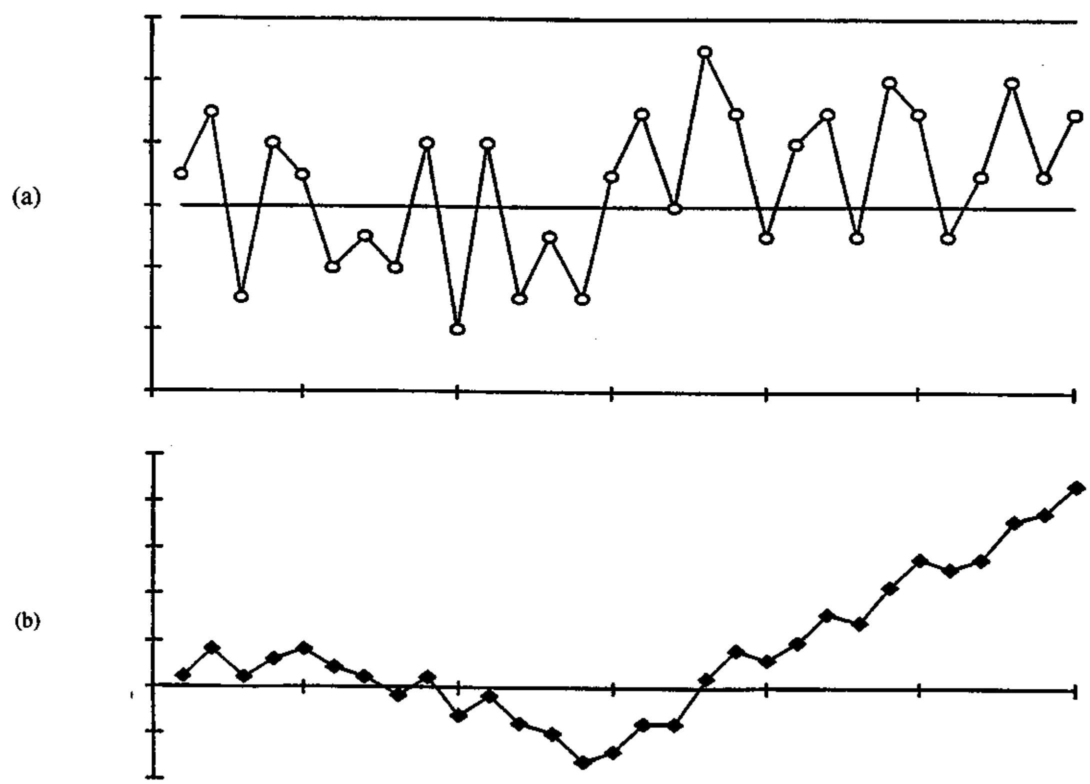

Figure 1. *A series with a one-standard deviation step change occurring at observation 16 plotted (a) as a Shewhart chart and (b) as a Cusum chart.*

*Edited by George Box and Søren Bisgaard

†The examples and illustrations used in this column are taken by permission from the book *Statistical Control by Monitoring and Feedback Adjustment* by George Box and Alberto Luceño.

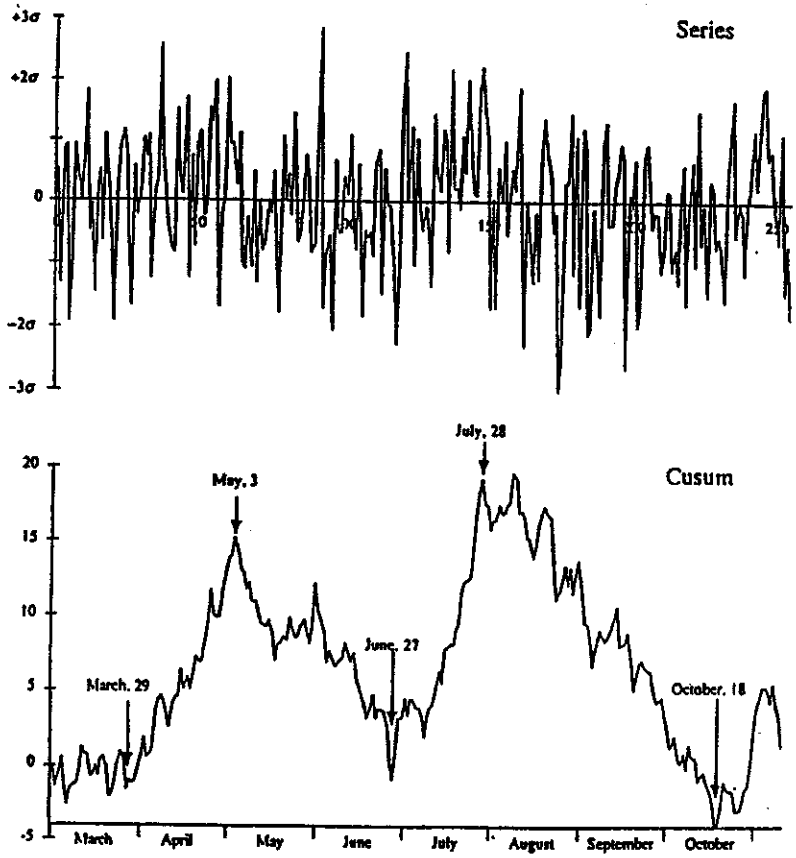

Figure 2 (a) Daily readings of gas flow over an eight-month period, (b) A Cusum chart for this data.

over the last 15 observations a total increase of about 32 occurs in the Cusum so that the amount of shift can be estimated to be about 32/15 = 2.1 units. It is true that by using only the Shewhart chart, for example, with extended Westinghouse rules a change can also be detected. However, the Cusum characterizes the point of change and its magnitude more distinctly.

Although over the last few decades there has been extensive development of the formal theory of Cusum tests, the usefulness and simplicity of Cusums in the analysis of past data has sometimes been overlooked. In particular their ability to determine when and by how much the process mean has changed is often of great use in tracking down assignable causes of trouble.

An interesting example showing the Cusum used as a diagnostic tool occurred in the analysis of data from one stage of a large chemical process in which two gases were brought together at very high temperatures to form a desired product. In order to avoid production of impurities it was necessary that the gas flows were maintained as close as possible to specified constant levels. Deviations from target for daily flow readings of one of the gases over an eight-month period are shown in Figure 2a together with 2σ and 3σ limits. The run chart in Figure 2(a) does not immediately bring to attention any features of particular interest. However, in the Cusum chart for the same data, shown in Figure 2(b) a series of slope changes are clearly seen, indicating possible changes in the mean in the original data. The chart suggests that these changes occurred close to the following dates: March 29, May 3, June 27, July 28, and October 18. After some thought, it was realized that these times were close to those when the meters measuring the gas flow were recalibrated. As a result of this discovery, the calibration system was drastically modified and the problem eliminated.

The Cusum as an Example of a Cuscore Statistic

The Cusum is a particular example of what Box and Ramirez[5] called *Cuscore* statistics. They regarded Shewhart and other quality control charts used for process monitoring as particular devices for detecting a *signal* buried in *noise*. The signal represents a particular kind of pattern in the data, characterizing a malfunction, which we need to detect. The noise is the background variation in which the signal is hidden. This background variation might, for example, be characterized by "white noise": that is, by *independently* distributed deviations from the mean drawn from an approximately normal distribution. Alternatively it might be characterized by dependent deviations

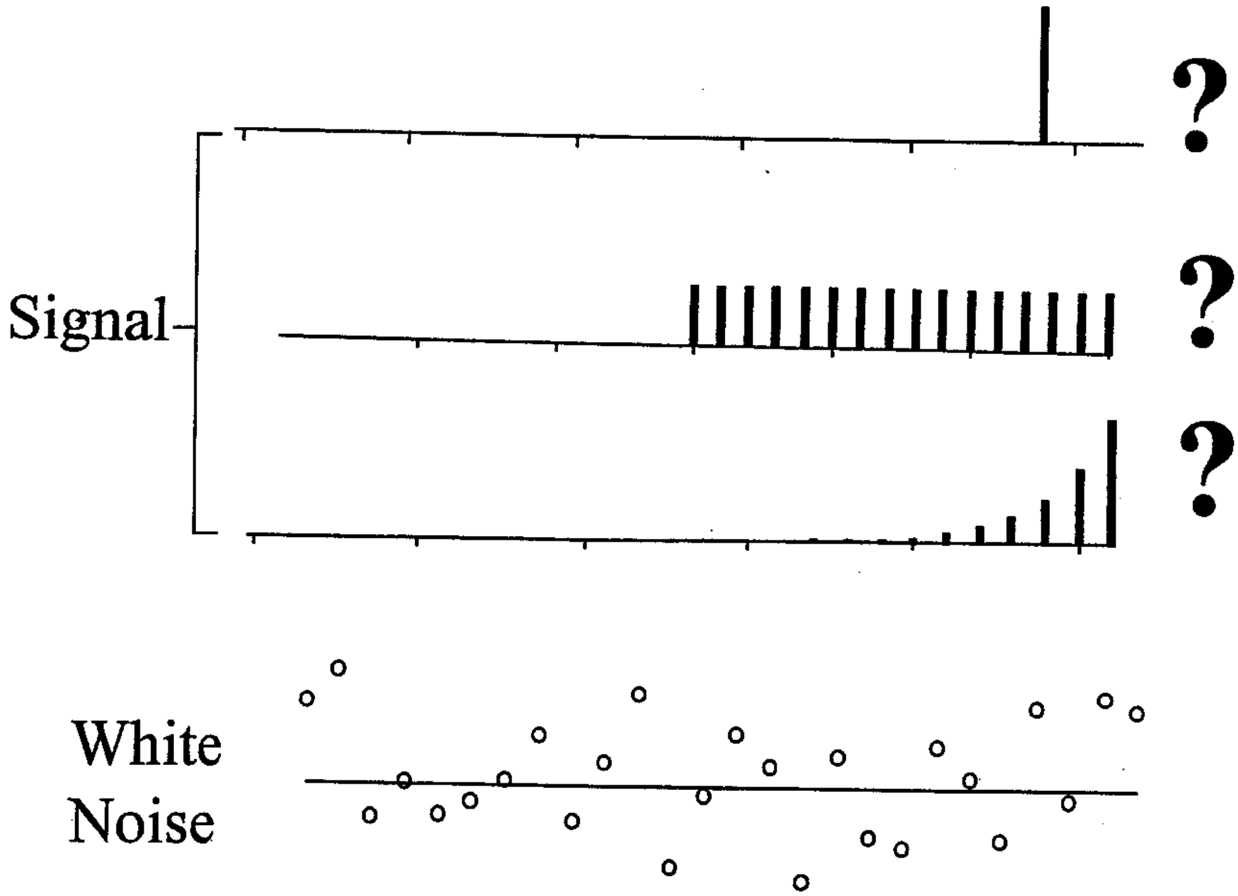

Figure 3. Three signals which when buried in white noise are best detected by: a Shewhart chart, a Cusum chart and an EWMA chart.

generated by some kind of time series model and possibly by distributions other than the normal.

In any case, given a) the kind of signal you are looking for, b) the kind of noise in which it is buried, these authors showed how to determine the most sensitive statistic for signal detection by using the efficient score statistic of R.A. Fisher[6]. As an illustration, Figure 3 shows three different kinds of signals that might be of interest a) a "spike", b) a step function, and c) an exponential increase. If such signals are concealed in white noise then the plots of the appropriate Cuscore statistics can be shown to be, respectively, a) the Shewhart chart, b) Cusum chart, and c) the exponentially weighted moving average (EWMA) chart[7].

The Cuscore principle can be used quite generally for finding the most efficient detector for virtually any kind of signal in any kind of noise. In particular it may be used[7] to detect a signal in a process which is subject to feedback control.

References

1. Page, E.S., "Continuous inspection schemes," *Biometrika*, **41**, 100-114, 1954.
2. Barnard, G.A., "Control charts and stochastic processes," *Journal of the Royal Statistical Society, Series B*, **21**, 239-271, 1959.
3. Goldsmith, C.H. and Whitfield H.W., "Average Run Lengths in Cumulative Chart Quality Control Schemes," *Technometrics*, **III**, 11-20, 1961.
4. Lucas, J.M. and Crosier, R.B., "Fast initial response for cusum quality-control schemes: give your Cusum a head start," *Technometrics*, **24**, 196-206, 1982.
5. Box, G.E.P. and Ramirez, J., "Cumulative Score Charts," *Quality and Reliability Engineering International*, **8**, 17-27, 1992.
6. Fisher, R.A., "Theory of statistical estimation," *Proceedings of the Cambridge Philosophical Society*, **22**, 700-725, 1925.
7. Box, G.E.P. and Luceño, A., "*Statistical Control by Monitoring and Feedback Adjustment*," Wiley, New York, 1997.