

*Center for Quality and Productivity Improvement*  
UNIVERSITY OF WISCONSIN  
610 Walnut Street  
Madison, Wisconsin 53705  
(608) 263-2520  
(608) 263-1425 FAX

*Report No. 65*

## **Sequential Methods in Statistical Process Monitoring**

*Chapter 1: Introduction*

*Chapter 2: Sequential Monitoring of Variances*

George Box and José Ramírez

*May 1991*

---

This work was sponsored by the National Science Foundation Grant DDM-8808138 and by the Vilas Trust of the University of Wisconsin-Madison.

This report is chapters 1 and 2 of a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Statistics) at the University of Wisconsin-Madison (1989). Thesis adviser: George Box.

The Center for Quality and Productivity Improvement cares about your reactions to our reports. Please direct comments (general or specific) to: Report Editor, Center for Quality and Productivity Improvement, 610 Walnut Street, Madison, WI 53705; (608) 263-2520. All comments will be forwarded to the author(s).

## Sequential Methods in Statistical Process Monitoring

George Box

José Ramírez\*

Center for Quality and  
Productivity Improvement

*University of Wisconsin  
Madison, Wisconsin*

### ABSTRACT

Cumulative Sum charts or CUSUM charts are normally used to monitor and detect departures from a process target. Although it is important to be as close as possible to target, this alone will not insure quality. It is also important to monitor the variability around the target. In this report a CUSUM to monitor variability, based on the Wald-Barnard likelihood ratio test, is discussed and a nomogram to aid in the construction of the charts is also presented.

**KEYWORDS:** *ARL, CUSUM, common causes, nomogram, special causes, SPC variation*

---

\* José Ramírez is now a Principal Engineer at Digital Equipment Corporation, 77 Reed Road, HL 02-3/J13, Hudson, Massachusetts 01749-9987.

This work was sponsored by National Science Foundation Grant DDM-8808138 and by the Vilas Trust of the University of Wisconsin-Madison.

Copyright © 1989 by José Ramírez. This report is chapters 1 and 2 of a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Statistics) at the University of Wisconsin-Madison (1989). Thesis adviser: George Box.

## Table of Contents

### Chapter 1: *INTRODUCTION*

1.1 <i>Motivation</i>	1
1.2 <i>CUSUM Methods for Location</i>	5
1.3 <i>CUSUM Methods for Variance</i>	8
1.4 <i>CUSUM Methods for Models</i>	13

### Chapter 2: *SEQUENTIAL MONITORING OF VARIANCES*

2.1 <i>Sequential Methods and CUSUM Tests</i>	15
2.2 <i>Average Run Length</i>	22
2.3 <i>Solutions for <math>P(z)</math> and <math>N(z)</math></i>	25
2.4 <i>Contour Nomogram</i>	33
2.5 <i>Comparison with Shewhart R-charts</i>	36
2.6 <i>The Effect of Non-normality</i>	46
2.6.1 <i>Exponential Power Distribution</i>	50
2.6.2 <i>Sequential Test</i>	51
2.6.3 <i>Average Run Length</i>	54
2.6.4 <i>Relation with t-distribution</i>	57
<i>REFERENCES</i>	67

# CHAPTER 1

## INTRODUCTION

### 1.1 Motivation

In this era of a world competitive economy quality has become the primary differentiator in the market place; never ending quality improvement is necessary to stay in business. Variability is present in everything we do. It is a fact of life that no two objects are exactly alike. Although it is important to recognize the existence of process variation, our goal should be that of reducing variation if we want to be able to provide products and services that have consistent quality. Reducing variation therefore, is the key element in any quality and productivity improvement program.

Fifty years ago Shewhart realized that process variation can be classified as controlled variation, characterized by an stable and consistent pattern over time, and uncontrolled variation in which the pattern of variation changes over time. Deming calls *common causes* of variation those that produce controlled variation and *special causes* of variation those that induce uncontrolled variation. See for example Deming (1986).

Common cause variation is inherent in a process and is comprised of many small sources that affect all the elements in the process. Special variation, on the other hand, is not part of the overall system, its causes are localized in nature and

should be considered as abnormalities.

Wheeler and Chambers (1986), discuss four categories that are useful ways to describe the state of a process.

- a) *Chaos*: the process is producing some nonconforming product and special causes of variation are present.
- b) *Brink of Chaos*: the process produces 100% conforming product but some special causes are still present
- c) *Threshold State*: the process is stable but is producing some nonconforming products. Variation is the product of common causes.
- d) *Ideal State*: in which a process is stable and produces 100% conforming products.

In view of this classification we can say that the main objectives of any *process monitoring* procedure, in an effort to build quality into the process, should be :

- i) Recognize and monitor the common cause system, and
- ii) Identify special causes of variation so they can be removed.

It is important then to be able to quickly detect process shifts or special causes so that an investigation of the process and remedial actions can be undertaken. Control charts, by showing the nature and amount of variation by time, are useful tools for studying variation. They enable the user to judge whether the process is in a state of statistical control, to interpret patterns and to detect changes.

Control charts not only aid in process monitoring, but they also help minimize overadjustment and underadjustment of the process. Overadjustment occurs when changes are made even though the process is affected by common causes only, thus

increasing output variability. Underadjustment, on the other hand, occurs when there is a delay in changing operating conditions even though evidence exists of a change in excess of the natural variability of the process; i.e., special cause of variation.

Various types of control charts have been proposed for various purposes and specifically to indicate deviations from the common cause state. These include the Shewhart control chart and the cumulative sum (CUSUM) control chart.

### *Shewhart Charts*

Control charts for averages and ranges have been used traditionally as a simple statistical tool for understanding the sources of variation in a process. They help to understand and analyze process variables, and to identify and differentiate between the two causes of variation.

Shewhart control charts have the advantage of not being specific for any particular type of alternative; in other words, they do not assume that the special causes of variation are associated with any particular pattern. This provides a sort of safety net whatever type of variation should occur, even if it is something completely unexpected.

This type of control charts, being a sequence plot ordered in time, uses only the information about the process contained in the last plotted point, without taking into account the historical record of the data. Several criteria have been suggested; e.g., warning limits, tests of runs, etc., to use in conjunction with the Shewhart

charts in order to incorporate the information of the whole sample. These additional criteria however, reduce the simplicity and ease of interpretation of the Shewhart chart.

### *CUSUM Charts*

Cumulative sum charts or CUSUM charts, on the other hand, use the information in the entire sample by plotting the cumulative sums of the observations. Starting from a given point, all subsequent plots contain information from all the observations up to, and including, the plotted point.

These charts are more effective than Shewhart charts if we wish to look for a specific pattern associated with a particular kind of special cause, as for example the detection of small shifts in process level. This is a desirable feature because, as Marquardt (1984) points out, it is important to discover small persistent changes quickly before a major problem develops. The best strategy however, see Box (1980), will involve a combination of more general, Shewhart charts, and more specific, CUSUM charts that search for special causes of variation.

Cumulative sum charts are particularly effective for samples of size 1, making them a very good candidate to use as *dynamic* control charts. These charts, see Hahn and Cockrum (1987), utilize on-line measurements in order to provide early warning of significant changes in the measured characteristics.

## 1.2 Cusum Methods for Location

Cumulative sum charts were first proposed by Page (1954), who suggested constructing control charts using sums of observations rather than individual observations. Page's procedure consisted of assigning a score  $d_j$  to the  $j$ th sample, and to plot the sums  $\sum_{j=1}^t d_j$  against  $t$ . For example, if we are interested in the number of defectives in a sample, we can assign a score +19 to each defective occurring in the sample, and -1 for each non-defective item.

For detecting positive deviations from an acceptable value  $\mu_a$ , the scores are chosen so that the path of the cumulative sum is upwards when the process is out of control and downwards when the process is in control. Action is taken whenever  $\sum_{j=1}^t d_j$  exceeds a decision interval  $h$ .

Barnard (1959) suggested subtracting the target value  $\mu_a$  from each sample mean  $\bar{x}_j$ , and plotting the sums  $\sum_{j=1}^t (\bar{x}_j - \mu_a)$ . He also introduced the use of a V-mask as a decision rule to determine whether the process is out of control. This is a two-sided procedure in which a signal is given whenever the V-mask fails to obscure any previously plotted point, when the point located a distance  $d$  from the vertex is placed at the last plotted point

Johnson (1961) emphasized that the application of the V-mask to the control chart, is equivalent to the application of the Wald-Barnard sequential probability

ratio test applied in reverse, and obtained some approximate formulas relating the parameters of the V-mask with those of the sequential test.

The most commonly used scheme consists of subtracting a reference value  $k$  from each sample mean  $\bar{x}_j$  and plotting the cumulative sums  $\sum_{j=1}^t (\bar{x}_j - k)$ . When looking for positive deviations, one only has to plot those points that produce a positive contribution to the cumulative sum. In other words, we should plot those points for which the deviation from the reference value,  $r_t = \bar{x}_t - k$ , is positive. Following this, see also Van Dobben de Bruyn (1968), we can define the cumulative sum procedure as follows: let  $CS_t$  denote the cumulative sum plotted at time  $t$ ; then starting at zero; i.e.,  $CS_0 = 0$ , we have

$$CS_t = \max(0, CS_{t-1} + r_t) \quad (1.1)$$

Action is taken whenever the cumulative sum exceeds the decision interval  $h$ ; i.e., whenever  $CS_t \geq h$ .

The performance of the CUSUM scheme is measured in terms of the Average Run Length (ARL) which is the average number of samples taken at a given process level before a signal is given by the chart. For one sided procedures, the CUSUM charts are designed by choosing the sample size  $n$ , the reference value  $k$  and the decision interval  $h$ , according to some pre-specified ARL at the acceptable quality level  $\mu_a$  and the rejectable quality level  $\mu_r$ .

The reference value  $k$  is usually taken to be  $k = (\mu_a + \mu_r)/2$ . This choice of  $k$  is believed to be optimal, in the sense that it provides maximum discrimination between the two values of  $\mu$ . Although no formal proof has been given of this result, Bagshaw and Johnson (1975) proved, by means of a Wiener process approximation, that the result is true, at least approximately for the normal case.

The values of  $n$  and  $h$  are chosen with the aid of tables, such as those found in Lucas (1976), or the contour nomogram constructed by Goel and Wu (1971), which gives contours for the ARL at the acceptable level  $\mu_a$  and the rejectable  $\mu_r$  in the  $(h\sqrt{n}/\sigma, |\mu - k| \sqrt{n}/\sigma)$  plane.

A frequent criticism of the CUSUM chart, see for example Woodall (1986), is that it does not detect large shifts in the process level as quickly as the Shewhart chart. Several modifications to the CUSUM chart have been proposed; in particular the use of a parabolic V-mask has been emphasized by Lucas (1973); although the idea was mentioned in Barnard (1959).

For initial out of control situations, Lucas and Crosier (1982) suggest the use of the Fast Initial Response (FIR) CUSUM. The idea is to start the CUSUM half way between the lower boundary and the action limit  $h$ ; i.e., to set  $CS_0 = h/2$ , for more rapid response. It is not difficult to see however, that the closer you start to  $h$  the faster you will get an out of control signal. Finally, Lucas (1982) has suggested the addition of Shewhart limits to the CUSUM chart to increase the speed of detecting large shifts. Using this procedure, he says, we combine the best properties of the two charts. This may be the case, but as we have seen the two charts

serve different purposes, CUSUM charts look for particular patterns while Shewhart charts are not specific to any particular pattern.

### 1.3 Cusum Methods for Variance

Although much work has been done for CUSUM charts for detecting location shifts in normally distributed data, the problem of monitoring shifts in process variability, from an acceptable level  $\sigma_a$  to a rejectable level  $\sigma_r$ , using the CUSUM methodology has received less attention.

Johnson and Leone (1962), based on the equivalence between sequential tests and CUSUM charts, suggested the use of sample variances,  $\sum_{i=1}^k (x_i - \bar{X})^2 / (k - 1)$ , to monitor process variance; they do not however, give any suggestion of what the sample size should be. Formulas for the construction of a V-mask and for obtaining approximate ARL values are also given. These are based on the probability  $\alpha$  of producing false alarms. These approximations rest on the assumption that although  $\alpha$  will not be exactly the proportion of false alarms, tests with the same  $\alpha$  will have the same ARL. Van Dobben de Bruyn (1968) has pointed out that we should be careful in drawing conclusions from those approximations because we have no information about the accuracy of the assumption made.

Page (1963) suggests the use of cumulative sums of the sample ranges, and gives a small table of ARL values for the design of such charts. These results were obtained by approximating the integral equations that define the ARL. Although

these results are more reliable than those obtained by the Johnson and Leone, the author points out that his results are of low accuracy and that errors of one unit in the second significant figure can be expected.

Others have suggested, see for example Gunst *et al.* (1989), plotting the cumulative sums of  $\log(s_i^2)$ , which is approximately normal, and using the tables and nomogram available for monitoring shifts in process mean. However this approximation implies the use of relatively large samples. For example, Bartlett and Kendall (1946), suggested that this approximation is good for  $n \geq 10$ , should be used with caution for  $5 \leq n \leq 9$ , and not at all for  $n < 5$ .

More recently Hawkins (1981) has proposed the use of the same CUSUM procedure for monitoring the mean to monitor the variance, by using  $Y_i = |X_i/\sigma|^{1/2}$  which is nearly normal if  $X$  is distributed as  $N(0, \sigma^2)$ . The problem with this approach is that if the mean of the variable  $Y_i$  changes by an amount  $\sqrt{a}$ , which corresponds to a change in the variance of  $X_i$  from  $\sigma^2$  to  $a^2\sigma^2$ , its variance also changes by an amount  $a$ . Some results, Bagshaw and Johnson (1975), suggest that an unstable variance can cause drastic changes in the the ARL values of the scheme. This prevents the use of the available ARL tables or the contour nomogram of Goel and Wu.

### *Normal Distribution*

We follow on the ideas of Johnson and Leone, of equivalence between cumulative sums and sequential tests, to derive cumulative sums for monitoring vari-

ance; and those of Page, of obtaining ARL values by solving the integral equations that define the ARL.

Chapter 2 discusses a cumulative sum based on the Wald-Barnard likelihood ratio which leads to the cumulative sum

$$\sum_{i=1}^k [(x_i - \mu)^2 - s^2] ; \quad (1.2)$$

where  $s^2$  is a reference value. One of the advantages of this chart is that we only need one observation; i.e., subgroups of size 1, as opposed to the use of ranges for which we need at least two observations.

By solving the integral equations that define the ARL, a contour nomogram similar to the one of Goel and Wu (1971), has been developed that helps in the design of the charts. The nomogram has contours of ARL at the acceptable quality level  $\sigma_a$  and at the rejectable quality level  $\sigma_r$ , in the plane  $(h/\sigma_a^2, \sigma_r/\sigma_a)$ .

### *Process Improvement*

In the past control charts were used mainly to insure the process met levels of specification. More recently their use in improving a process has been emphasized; in particular reduction of process variation. Rather than focusing on the detection and avoidance of adverse change, we need never ending improvement striving to implement beneficial changes. Therefore, as Gitlow *et al.* (1989) point out, we must constantly attempt to reduce process variation around desired target values to achieve the degree of uniformity required to produce products and services that

surpass customer needs and expectations.

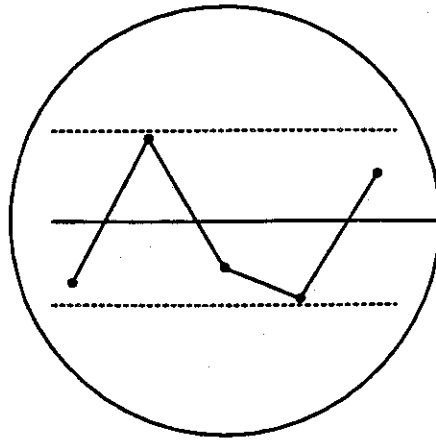
Sequential tests can quickly detect deviations from particular levels of variation. For example, figure 1a shows a Shewhart chart with  $3\sigma$  limits as shown in the logo of the American Society for Quality Control. This process does not show any abnormal pattern, nor any point out of control. Assuming that  $\sigma_a = 1$ , the CUSUM scheme obtained by plotting the cumulative sums given by equation 1.2 with  $s^2 = 1.85$ , can quickly detect, figure 1b, a change in process variance of 100%; i.e.,  $\sigma_r = 2$ . The chart clearly shows an upward trend, which is an indication of possible lack of control.

The methods presented in chapter 3 focus on the design of CUSUM charts for monitoring positive or negative shifts in process variability, as well as some applications for the detection of outliers. Approximate formulas for the computation of ARL values are also given. These formulas can give an idea of what the performance of the chart would be when using a chosen  $s^2$  and  $h$ .

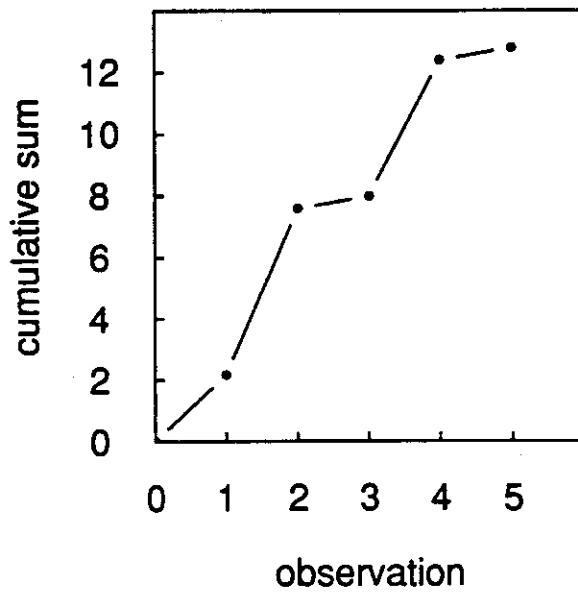
#### *Robustness of the ARL with Respect to Distributional Assumptions*

It has been shown, Kemp (1967), that for CUSUM charts for location the ARL estimates obtained by assuming that a non-normal variable is normal, can be very inaccurate. By assuming that the variate being cumulated is positively skewed, he showed that the ARL values are overestimated when one assumes normality.

**Figure 1** a) ASQC logo showing an  $\bar{X}$ -chart with action limits at  $\pm 3$ . b) CUSUM chart for monitoring an increase in variance from  $\sigma_a = 1$  to  $\sigma_r = 2$ .



ASQC logo



In section 2.6 we study the effect of non-normality in the ARL values by assuming that the observations are distributed according to the exponential power distribution, Box and Tiao (1973). This is a family of symmetric distributions, of which the normal is a member, with different degrees of kurtosis ranging from the rectangular distribution to the double exponential. We show how the assumption of normality can lead to overestimated ARL values, for platykurtic distributions, and underestimated values for leptokurtic members of the family.

#### 1.4 Cusum Methods for Models

Statistical models are very useful for the study of nature. They allow us to express the relationship between a response variable and some explanatory variables. Modelling of the relationship between variables involves several unknown *parameters* that have to be estimated from the data. Once a model has been adequately identified, fitted and checked, it can be used to forecast future values of the response.

Time series models, for example, can produce methods for optimal forecasting of future observations. However, if the parameters in the model were to change; such forecasts would be in error and it would be necessary to update the parameters in our model. These situations need to be detected quickly so appropriate action can be taken in order to minimize possible losses from the errors.

In chapter 4 it is shown how a sensitive test for shifts of the parameters of any model can be derived using a cumulative score technique (CUSCORE), of which

the cumulative sums mentioned in section 2.2 are a special case. As a particular example consider an industrial process in which trouble has been experienced because of a characteristic harmonic cycling at a particular period and phase. Suppose action has been taken which, temporarily anyway, has removed this source of difficulty, but it is feared that it might recur. In such a case a test procedure is needed which is particularly sensitive to this kind of deviation from randomness and so can give the earliest possible warning of recurrence of the trouble.

## CHAPTER 2

### SEQUENTIAL MONITORING OF VARIANCES

In every process it is important that the mean be as close as possible to the target value; but ensuring that the mean is on target is usually not enough to produce products and services with a desired level of quality. If the variability of the process is out of control, no matter how close to the target the mean is, we are always going to produce products that do not satisfy the customer requirements. It is then important to monitor the variability around the target value.

In section 1.3 we described several methods that can be used to monitor process variability. In particular, Shewhart charts for ranges have been used as a complement to the  $\bar{X}$  charts. The purpose of this chapter is to show how sequential tests lead to the use of cumulative sums to monitor the variability of a process.

#### 2.1 Sequential Methods and CUSUM Tests

In the usual setting of point estimation and hypothesis testing the sample size is determined in advance. A sample of  $n$  independent observations is made, and from this sample we obtain point estimates for the parameters of interest and base our final decisions on the calculated value of some test statistic that usually has certain optimal properties. In sequential tests on the other hand, the sample size is not fixed in advance but may depend, in some specified way, on the data as they

become available in the course of the investigation. In other words, the sample size is a random variable determined by the observed data.

Sequential methods take the approach that acceptance or rejection of the null hypothesis should be made as soon as convincing evidence is available either in support of or against the null hypothesis. Sample information is evaluated as it becomes available, rather than at the end of the sampling process. These methods are motivated by economy of experimentation and duration of sampling; i.e., since observations are taken until a decision can be reached, time and money are not wasted by taking more measurements than are needed.

These methods when combined with graphics as in the cumulative sum charts, are particularly useful for process surveillance and improvement since they allow us to monitor processes by plotting observations one at a time until a lack of control is indicated.

Consider a process that produces observations  $x_i$  that are distributed with known mean as  $N(\mu, \sigma)$ . We want to be able to discriminate between the process being at an acceptable quality level with standard deviation  $\sigma_a$  and a rejectable quality level with standard deviation  $\sigma_r$ .

At each stage of sampling  $k$ , the ratio of the likelihood,  $LR_k$ , of obtaining the observations under the two states of the process, in control with variability  $\sigma_a$  or out of control with variability  $\sigma_r$ , is calculated. If this ratio is large, we conclude that the process is out of control; if the ratio is small, we conclude that the process is in control; finally if  $LR_k$  has an intermediate value, the decision is postponed

until further observations produce either a higher or lower value of  $LR_k$ .

The upper and lower limits for the likelihood ratio  $LR_k$  are often chosen to make the probabilities of making wrong decisions, errors of type I and II, small; or in terms of the Average Run Length (ARL) of the test.

Following Wald (1947), the likelihood ratio test,  $LR_k$ , for detecting changes in the variability of the process from a level  $\sigma_a$  to a level  $\sigma_r$ , is the ratio of two normal densities with standard deviations  $\sigma_a$  and  $\sigma_r$ , which simplifies to

$$LR_k = \left( \frac{\sigma_a}{\sigma_r} \right)^k \exp \left[ \frac{1}{2} \left( \frac{1}{\sigma_a^2} - \frac{1}{\sigma_r^2} \right) \sum_{i=1}^k (x_i - \mu)^2 \right] \quad (2.1)$$

If  $A$  and  $B$  denote the lower and upper limits respectively, the sequential probability ratio test is then

$$\begin{aligned} \text{if } LR_k \leq A & \quad \text{process at acceptable variability } \sigma_a, \\ \text{if } LR_k \geq B & \quad \text{process at rejectable variability } \sigma_r, \\ \text{if } A \leq LR_k \leq B & \quad \text{continue taking observations.} \end{aligned} \quad (2.2)$$

After taking logarithms equation (2.1) simplifies to

$$\log A + k \frac{2 \log \left( \frac{\sigma_r}{\sigma_a} \right)}{\frac{1}{\sigma_a^2} - \frac{1}{\sigma_r^2}} \leq \sum_{i=1}^k (x_i - \mu)^2 \leq \log B + k \frac{2 \log \left( \frac{\sigma_r}{\sigma_a} \right)}{\frac{1}{\sigma_a^2} - \frac{1}{\sigma_r^2}} \quad (2.3)$$

The right and left hand sides of equation (2.3) define the boundary lines of the sequential test. Note that the slope depends on the two values of the variability of the process. Crossing any of these two lines will give rise to a decision that either

the process variability has changed or has remained constant. Sampling will continue as long as we stay within the boundary lines.

The sequential test leads to the use of the cumulative sums of  $(x_i - \mu)^2$ . Johnson (1961) illustrated that the application of the V-mask to the cumulative sum control chart can be regarded as equivalent to the application of a sequential probability ratio test in reverse (see also Khan 1984). The difference between the two procedures is that the CUSUM chart makes no provision for the definite acceptance of no lack of control. In other words, even if the cumulative sum hits the lower boundary we will continue taking observations until the upper boundary is crossed.

A convenient CUSUM chart can be obtained by choosing a "reference" value  $s^2$ , and plotting the cumulative sum of

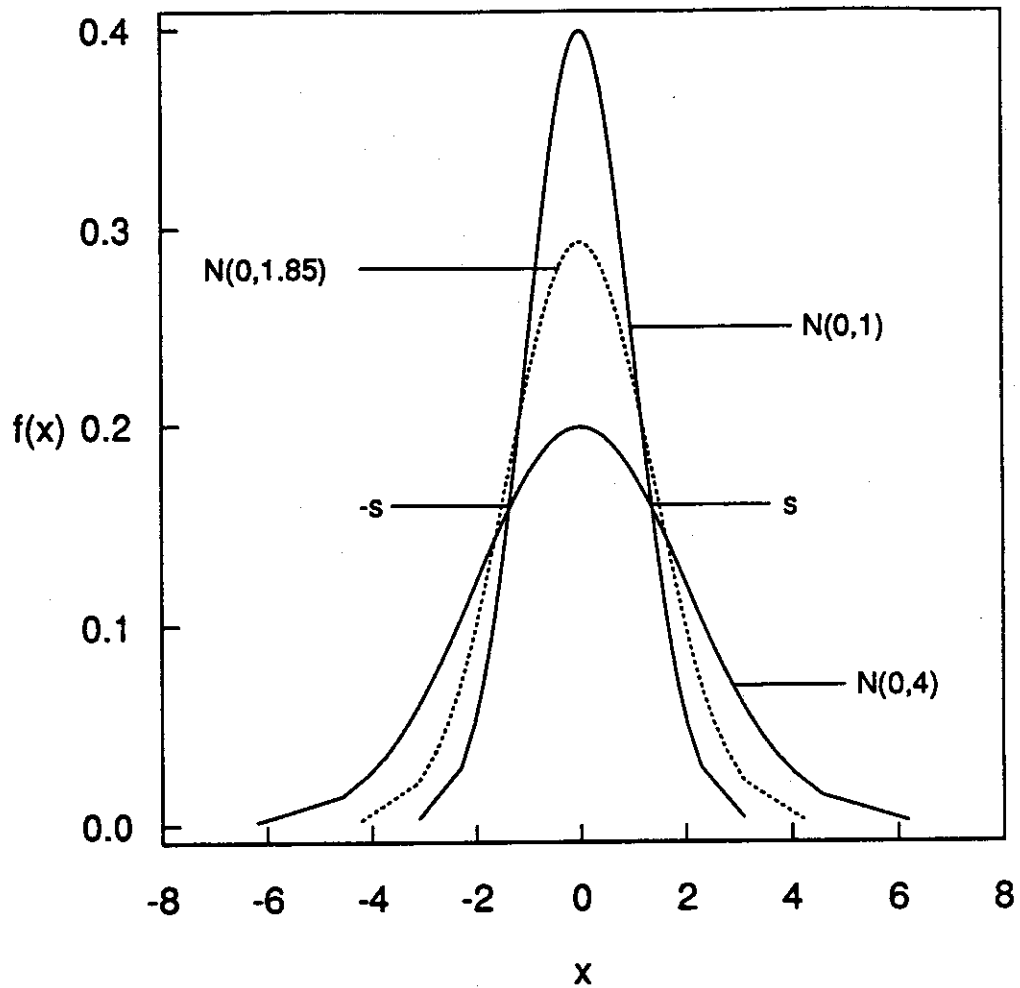
$$S_k = \sum_{i=1}^k [(x_i - \mu)^2 - s^2] \quad (2.4)$$

versus the number of observations  $k$ . If the population variance is equal to this value, the resulting figure will be horizontal; otherwise, the figure will slope downwards if the population variance is less than  $s^2$  and upwards if it is greater than  $s^2$ .

By looking at equation (2.3), we see that we can choose  $s^2$  to be the slope of the sequential test; i.e.

$$s^2 = \frac{\log\left(\frac{\sigma_r^2}{\sigma_a^2}\right)}{\frac{1}{\sigma_a^2} - \frac{1}{\sigma_r^2}} \quad (2.5)$$

Figure 2.1 Normal distributions with mean 0 and variance 1 and 4 (solid lines).  
Normal distribution with variance  $s^2 = 1.85$  (dotted line).

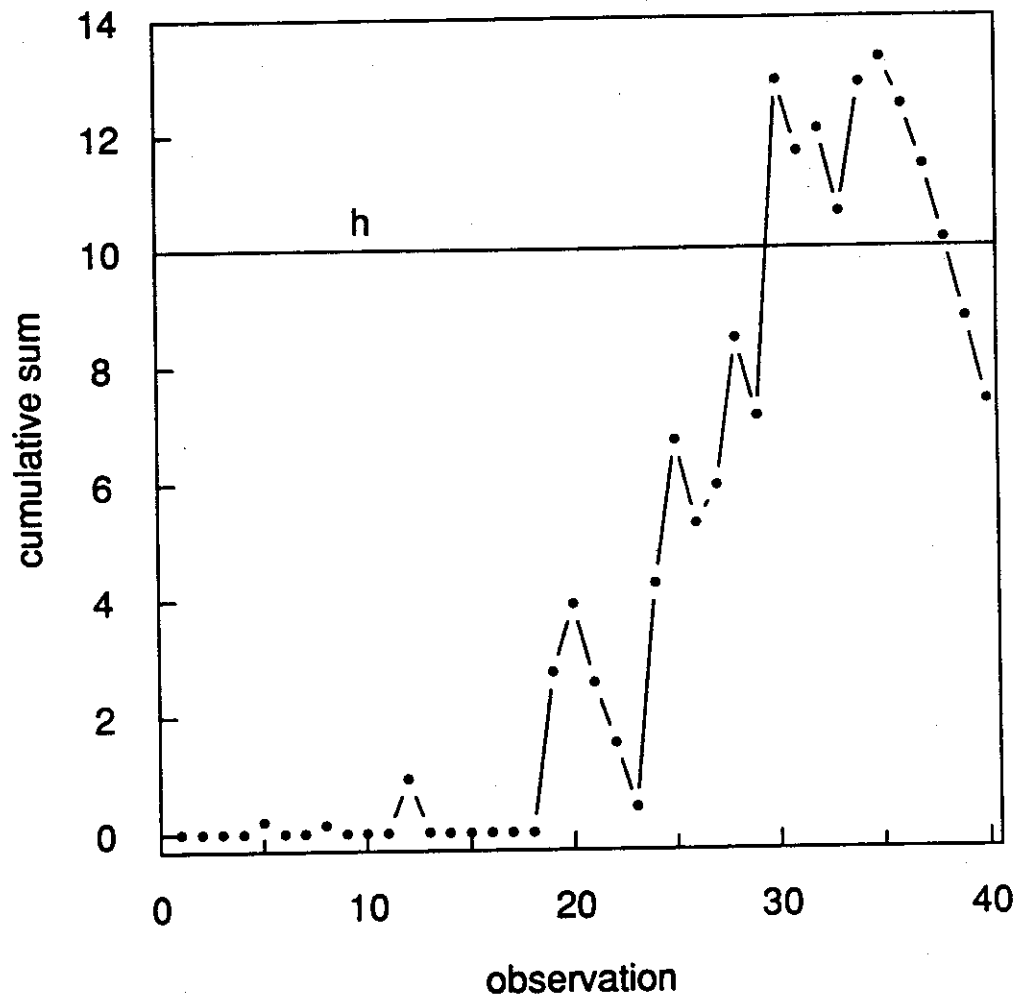


This choice of  $s^2$  corresponds to the value at which the likelihood of an observation coming from each of the two stages of the process or populations is the same, providing maximum discrimination between  $\sigma_a$  and  $\sigma_r$ . This can also be seen as a "mean" value between  $\sigma_a^2$  and  $\sigma_r^2$ , in the sense that two normal curves with variances  $\sigma_a^2$  and  $\sigma_r^2$  will intersect at the points  $\pm s$ . This is illustrated in figure 2.1 which shows two normal distributions with mean 0 and variances 1 and 4 respectively; the dotted line corresponds to a normal distribution with mean 0 and "mean" variance  $s^2$  which in this example is equal to 1.85.

If we consider departures from  $\sigma_a$  in one direction only; i.e., one-sided tests, we could plot the cumulative sums in (2.4) only when they are relevant towards taking a decision that, for example, the process variance has increased. This means that starting from zero, values of  $S_k$  less than zero need not be plotted and the process can be considered satisfactory. However, as soon as a result exceeds zero, a cumulative chart is started. If the cumulative sum subsequently reverts to zero, it is concluded that the process is satisfactory, but whenever  $S_k$  reaches or crosses a stated amount  $h$ , called the "decision interval", the process variance is said to have increased. See figure 2.2 which shows a chart with decision interval  $h = 10$  and out-of-control signal given at observation 30.

The control scheme just described corresponds to a sequence of Wald sequential tests with horizontal boundaries  $(0, h)$  and initial score zero. The test is reapplied when the previous test ends at the lower boundary and action is taken when a test ends at the upper boundary  $h$ . The efficiency of this control scheme will

**Figure 2.2** Cusum chart for detecting an increase in variance. The decision interval is  $h = 10$  and a signal is given at observation 30.



depend on the choice of the parameters  $h$  and  $s^2$ . This choice, as in the case of cumulative sum charts for the mean, see for example Goel and Wu (1971), will be based on some predefined Average Run Lengths.

## 2.2 Average Run Length

Following Page (1954), we defined the average run length (ARL) of a control scheme as the expected number of items sampled at a given quality level before a lack of control is indicated. In other words, if we focus on monitoring increases in process variability from an acceptable quality level  $\sigma_a$  to a rejectable quality level  $\sigma_r$ , the ARL is the expected number of items sampled before concluding that an increase in standard deviation has occurred.

The average run length depends on the values of  $h$ ,  $s^2$ , the true standard deviation of the process and the distribution of the observations. The values of  $h$  and  $s^2$  are chosen to yield large values of ARL when the process variability is at an acceptable level, and low values when is at a rejectable level. In other words, we want to have as few as possible false positives and we want to be able to detect changes quickly.

For a sequential test with boundaries  $(0, h)$  starting at a distance  $z$  from the lower boundary, the average sample number  $N(z)$  is the expected number of observations until the test either returns to zero or exceeds the decision interval  $h$ . The operating characteristic  $P(z)$  is the probability that a return to zero occurs.

The functions  $P(z)$  and  $N(z)$  are the solutions to Fredholm integral equations of the second kind

$$P(z) = \int_{-s}^{-z} f(y)dy + \int_0^h P(x)f(x-z)dx, \quad 0 \leq z \leq h \quad (2.6)$$

$$N(z) = 1 + \int_0^h N(x)f(x-z)dx, \quad 0 \leq z \leq h. \quad (2.7)$$

Where  $f(y)$  is the density function of the increments of the cumulative sum; i.e.,  $f(y)$  is the density of  $Y = [(X - \mu)^2 - s^2]$ ; where  $X$  is  $N(\mu, \sigma^2)$

$$f(y) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} (y + s^2)^{-1/2} \exp\left[-\frac{(y + s^2)}{2\sigma^2}\right] & \text{if } y \geq -s^2, \\ 0 & \text{if } y < -s^2. \end{cases} \quad (2.8)$$

This corresponds to a chi-squared distribution with one degree of freedom translated by an amount  $s^2$ .

An interpretation of equations (2.6) and (2.7) is as follows. The probability that a sequential test with boundaries  $(0, h)$  and initial value  $z$ , equation (2.6), ends in the lower boundary, equals the probability that the next observation  $x$  is such that  $z + [(x - \mu)^2 - s^2] \leq 0$  plus the integral over the probabilities that the cumulative sum lands somewhere between 0 and  $h$  multiplied by the respective probability  $P(x)$ . In a similar way, the average sampling number, equation (2.7), equals one (the next observation is always there for  $0 \leq z \leq h$ ), plus the integral over the

probabilities that the cumulative sum lands somewhere between 0 and  $h$  multiplied by the respective average sampling number  $N(x)$ .

The average run length, being the expected number of observations of a succession of sequential tests which all but the last return to zero, is therefore the average sample number  $N(0)$  times the expected number of tests before one exceeds the decision interval  $h$ . The probability that one particular test exceeds the upper boundary  $h$  is  $1 - P(0)$  and the average number of tests is  $\frac{1}{1 - P(0)}$ . The average run length is given by

$$\text{ARL} = \frac{N(0)}{1 - P(0)}. \quad (2.9)$$

For known values of  $P(0)$  and  $N(0)$ , the value of the average run length can be obtained from the above expression. These values however have to be obtained by solving the integral equations satisfied by  $P(z)$  and  $N(z)$ .

Since explicit solutions of equations (2.6) and (2.7) are not possible, values of  $P(z)$  and  $N(z)$  have to be obtained by either approximate expressions or by numerical methods. Kemp (1958) has given some approximate formulae for calculating the values of  $P(z)$  and  $N(z)$ . In the case of the operating characteristic for example, the value of  $P(z)$  is approximated by  $A + Be^{\omega z}$ , where  $\omega$  is the non-zero root of the equation  $M_Y(t) = 1$ , where  $M_Y(t)$  is the moment generating function of the density given by equation (2.8). The method fails in this case because  $M_Y(t) = (1 - 2t)^{-1/2} = 1$  does not have a non-zero root. Numerical methods that

replace the integral by a system of linear equations are presented in the next section.

### 2.3 Solutions for $P(z)$ and $N(z)$

The integral equations (2.6) and (2.7) can be replaced by a system of linear equations and solved for the unknown variables. The system employed is similar to the one given by Kantorovich and Krylov (1964), in which a partition of the interval of integration is constructed and the integrals are approximated in each of the resulting subintervals.

Let  $\{a_0, a_1, \dots, a_n\}$  be a partition of the interval  $[0, h]$  where  $a_0 = 0$  and  $a_n = h$ . Different partitions are obtained by using different quadratures. When  $h$  and  $n$  are small, Gaussian quadrature is recommended; however, when the value of  $h$  gets large it is necessary to use a larger partition of the interval and it is better, for computational purposes, to use a simpler and more convenient quadrature, for example a tangential one in which the distance between consecutive  $a_i$ 's is constant and equal to  $h/n$ . The equation (2.6) can then be written as

$$P(z) = \int_{-s}^{-z} f(y)dy + \sum_{j=1}^n \int_{a_{j-1}}^{a_j} P(x)f(x-z)dx \quad (2.10)$$

Since the function  $P(z)$  is smooth and decreasing in the interval  $(0, h)$ , and assuming that the subintervals  $(a_{j-1}, a_j)$  are sufficiently small so that  $P(x)$  is constant and equal to  $P(z_j)$ , where  $z_j = \frac{a_j + a_{j-1}}{2}$ , equation (2.10) can be approxi-

ated by

$$\tilde{P}(z) = \int_{-s}^{-z} f(y)dy + \sum_{j=1}^n \tilde{P}(z_j) \int_{a_{j-1}}^{a_j} f(x-z)dx \quad (2.11)$$

Substituting  $z$  for the values  $z_i, i = 1, \dots, n$ , and letting  $\mathbf{b}$  and  $\mathbf{C}$  be the vector and matrix

$$\mathbf{b} = (b_1, b_2, \dots, b_n), \quad \text{where } b_i = \int_{-s}^{-z_i} f(y)dy \quad (2.12)$$

$$\mathbf{C} = (c_{ij}), \quad \text{where } c_{ij} = \int_{a_{j-1}}^{a_j} f(x-z_i)dx \quad (2.13)$$

The vector  $\tilde{\mathbf{P}} = (\tilde{P}(z_1), \dots, \tilde{P}(z_n))$  can be computed as

$$\tilde{\mathbf{P}} = (\mathbf{I} - \mathbf{C})^{-1} \mathbf{b} \quad (2.14)$$

where  $\mathbf{I}$  is the  $n \times n$  identity matrix. Similarly a numerical approximation for the function  $N(z)$  is given by

$$\tilde{\mathbf{N}} = (\mathbf{I} - \mathbf{C})^{-1} \mathbf{e}, \quad \text{where } \mathbf{e} = (1, 1, \dots, 1). \quad (2.15)$$

The value of  $\tilde{P}(0)$  is obtained by setting  $z = 0$  in equation (2.11)

$$\tilde{P}(0) = \int_{-s}^0 f(y)dy + \sum_{j=1}^n \tilde{P}(z_j) \int_{a_{j-1}}^{a_j} f(x)dx \quad (2.16)$$

Similarly for  $\tilde{N}(0)$  we have

$$\tilde{N}(0) = 1 + \sum_{j=1}^n \tilde{N}(z_j) \int_{a_{j-1}}^{a_j} f(x) dx \quad (2.17)$$

The solution of the integral equations is then obtained by solving a system of linear equations and the definite integrals given by the  $b_i$ 's and the  $c_{ij}$ 's, which are the area under the chi-squared density with one degree of freedom.

A computer program using FORTRAN 77, was written to compute  $N(0)$ ,  $P(0)$ , and ARL. The program was run on a VAX 11/780 computer running VMS. The IMSL subroutines, *LEQIF* and *MDCH*, were used to solve the system of linear equations and to compute the values of the chi-squared distribution (2.8) respectively.

The accuracy of the approximation was evaluated by obtaining values of ARL from equation (2.9), and checking the convergence of the solutions for an increasing number of points in the quadrature. For the ranges of  $s^2$  and  $h$  under study, the computed ARL converged to a fixed value when the number of points in the quadrature was less than 500.

The rate of convergence largely depends on the closeness of  $P(0)$  to unity, which is determined by the values of  $h$  and  $s^2$  (see figure 2.3a). This is particularly the case for large values of  $h$ , so the accuracy of the approximation depends on the number of points in the quadrature. The effect of  $P(0)$  being close to unity is to produce large values of ARL; since values larger than  $10^6$  are not of much practical interest (Goel and Wu 1971), we can regard the approximations as exact.

Figure 2.3a Value of the operating characteristic,  $P(0)$ , for  $h = 5, 10, 15, 20$ .

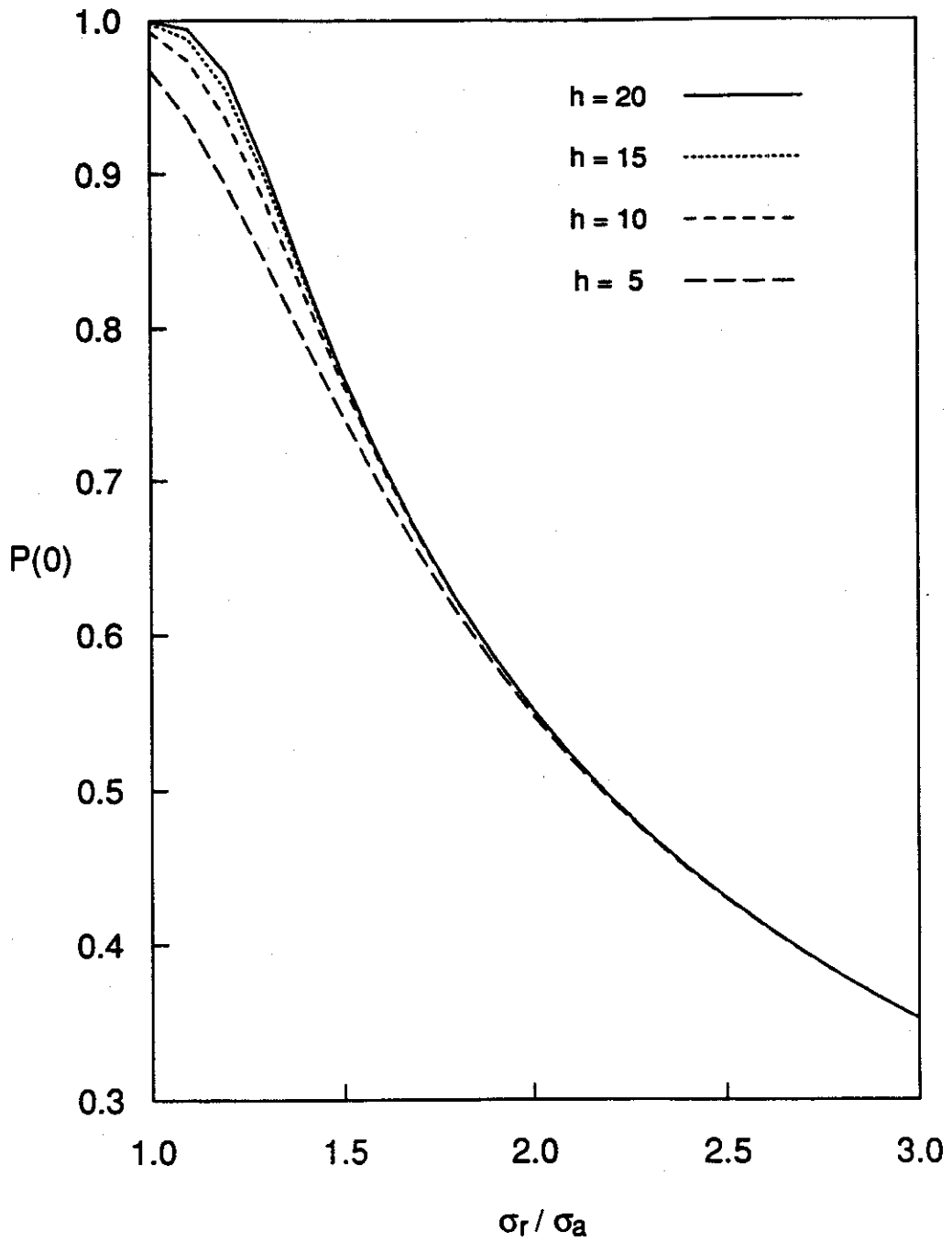
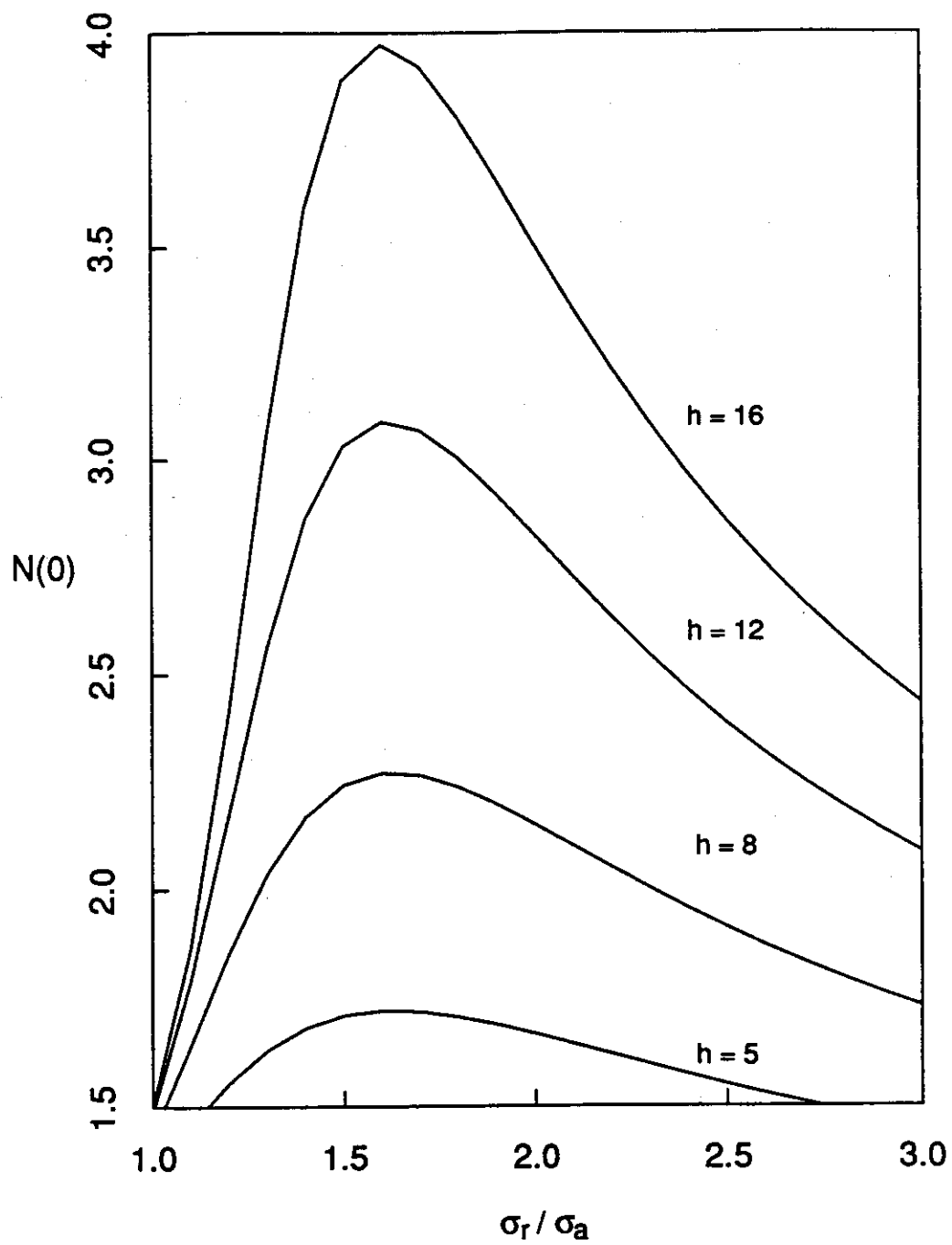


Figure 2.3b Value of the average sample number,  $N(0)$ , for  $h = 5, 8, 12, 16$ .



The figures 2.3a and 2.3b show the graphs of  $P(0)$  and  $N(0)$  for different values of  $h$ . Note how  $P(0)$ , the probability of a sequential test with boundaries  $(0, h)$  ends on or below the lower boundary, is a decreasing function of  $\sigma_r/\sigma_a$  and increases with  $h$ , for fixed values of  $\sigma_r/\sigma_a$ . The average sample number,  $N(0)$ , is a unimodal function of  $\sigma_r/\sigma_a$  for fixed  $h$ , and increases with  $h$  for fixed values of  $\sigma_r/\sigma_a$ .

Figure 2.4a shows the values of ARL, in logarithmic scale, for four different values of  $h = 5, 10, 15, 20$  and  $s^2 = 1.46$ . The figure was constructed to give values of the ARL for changes in the standard deviation from 1 to 3. Subsequently, we assume that the variable under study has been scaled to give  $\sigma_a = 1$ .

Note that the ARL is a decreasing function of  $\sigma$ , and that large values of ARL correspond to large values of  $h$ . Other graphs for values of  $h$  in the interval  $(5, 20)$  can be obtained by logarithmic interpolation.

An example will help to clarify how this graphs can be used to obtain desired ARL values corresponding to some predefined values of  $h$  and  $s^2$ . Let us suppose that a CUSUM chart, like figure 2.2, has been constructed with  $h = 10$  and  $s^2 = 1.46$ , to detect changes in the standard deviation of a continuous process, and that the variable representing this process has been scaled to give  $\sigma_a = 1$ . Figure 2.4a shows that even if the process variability is maintained at an acceptable level, i.e.  $\sigma_a = 1$ , on average we are going to have a false alarm every 250 observations.

On the other hand, if the variability of the process increases to a rejectable level  $\sigma_r = 2$ , on the average we expect the CUSUM chart to detect the change after

Figure 2.4a Average run length (logarithmic scale) for  $h = 5, 10, 15, 20$  and  $s^2 = 1.46$ .

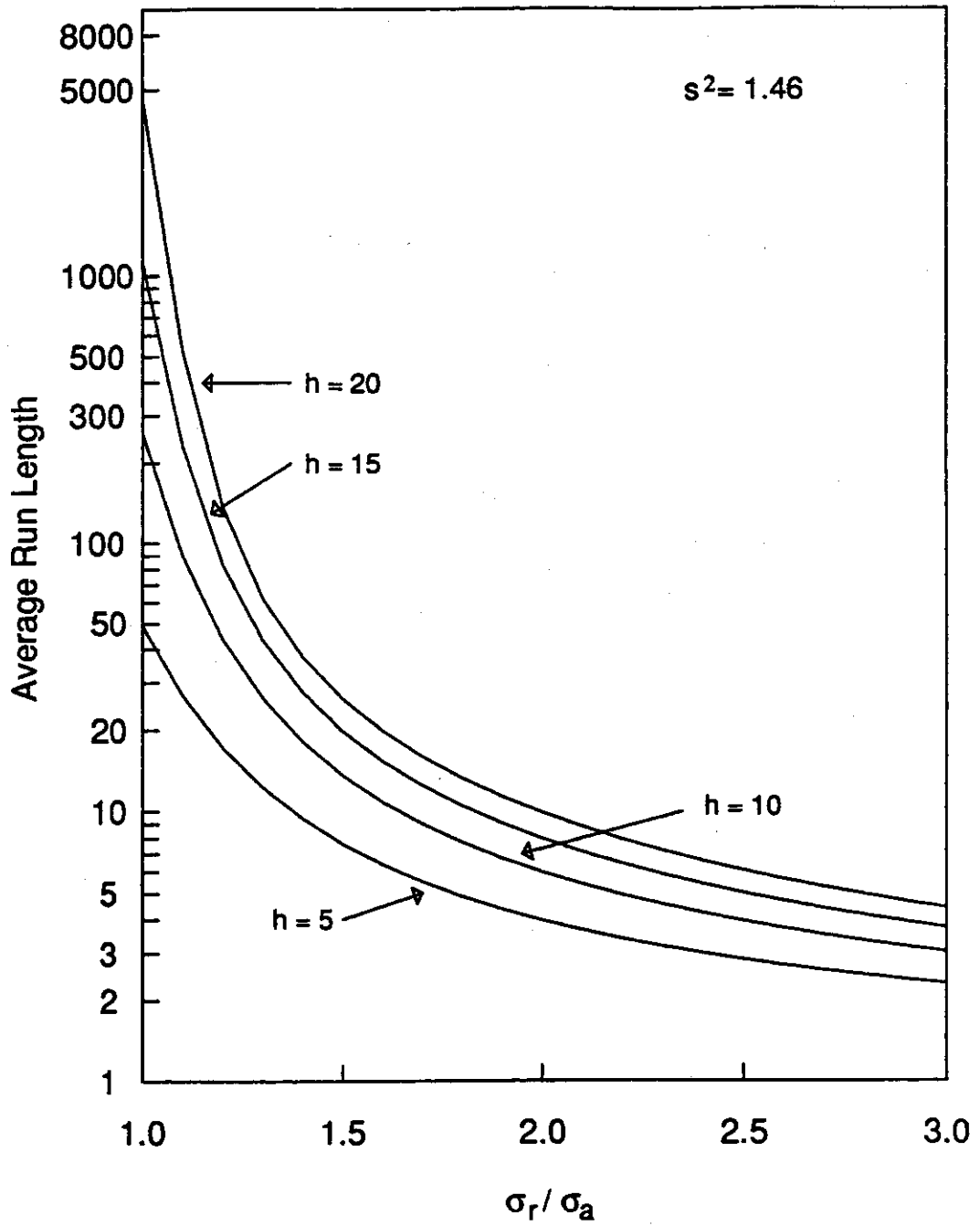
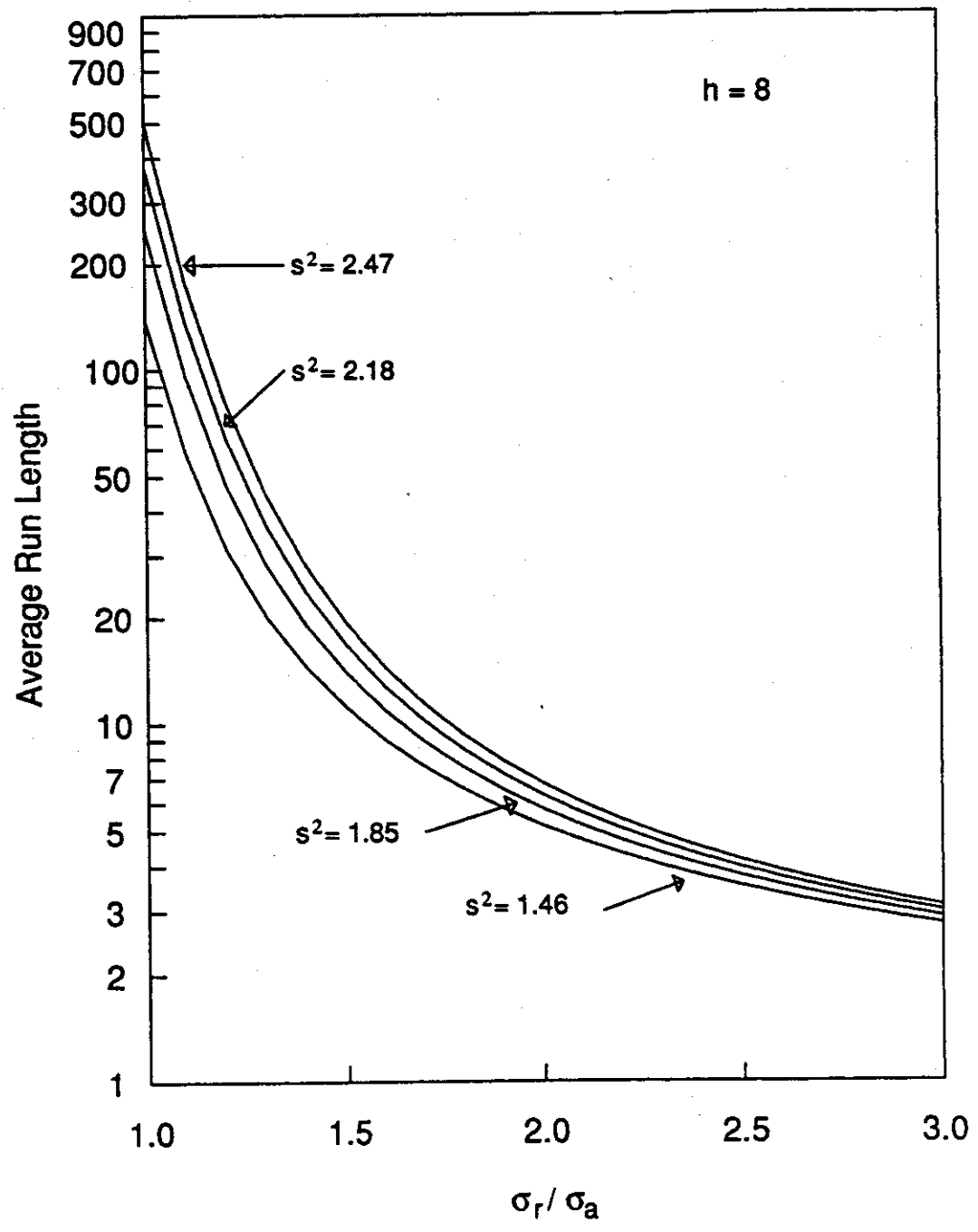


Figure 2.4b Average run length (logarithmic scale) for  $s^2 = 1.46, 1.85, 2.18, 2.47$  and  $h = 8$ .



6 observations. If we choose the decision interval to be  $h = 20$ , we would have "false alarms" once in every 5000 observations, and we would be able to detect a change of 100% ( $\sigma_r = 2$ ), in approximately 10 observations. Figure 2.4b was constructed in a similar way with  $h$  fixed and equal to 8, and different values of  $s^2$ .

The figures 2.4a and 2.4b summarize the behaviour of CUSUM charts designed to detect changes in the standard deviation up to 200%, and can be used to select the optimum values of  $h$  and  $s^2$ . This procedure however, is like the one-factor at a time approach in which we fixed  $h$  and compute the ARL values for different values of  $s^2$ . It is better to produce a diagram in which both  $h$  and  $s^2$  can be varied simultaneously. In the next section a contour nomogram, similar to the one given by Goel and Wu (1971) for monitoring means, is introduced that simplifies the design of the CUSUM charts for variances by giving a wider range of values for  $h$  and  $s^2$ .

#### 2.4 Contour Nomogram

At the two specified quality levels of the process variability, the acceptable quality level  $\sigma_a$  and the rejectable quality level  $\sigma_r$ , we denote the A.R.L by  $L_a$  and  $L_r$  respectively. As we have seen, the ARL depends on the values of  $h$  and  $s^2$  and is defined as the average number of samples taken at a given quality level before a lack of control is indicated. Therefore, it is desirable to have large values of  $L_a$  (as few false alarms as possible), and small values of  $L_r$  (quick detection of changes).

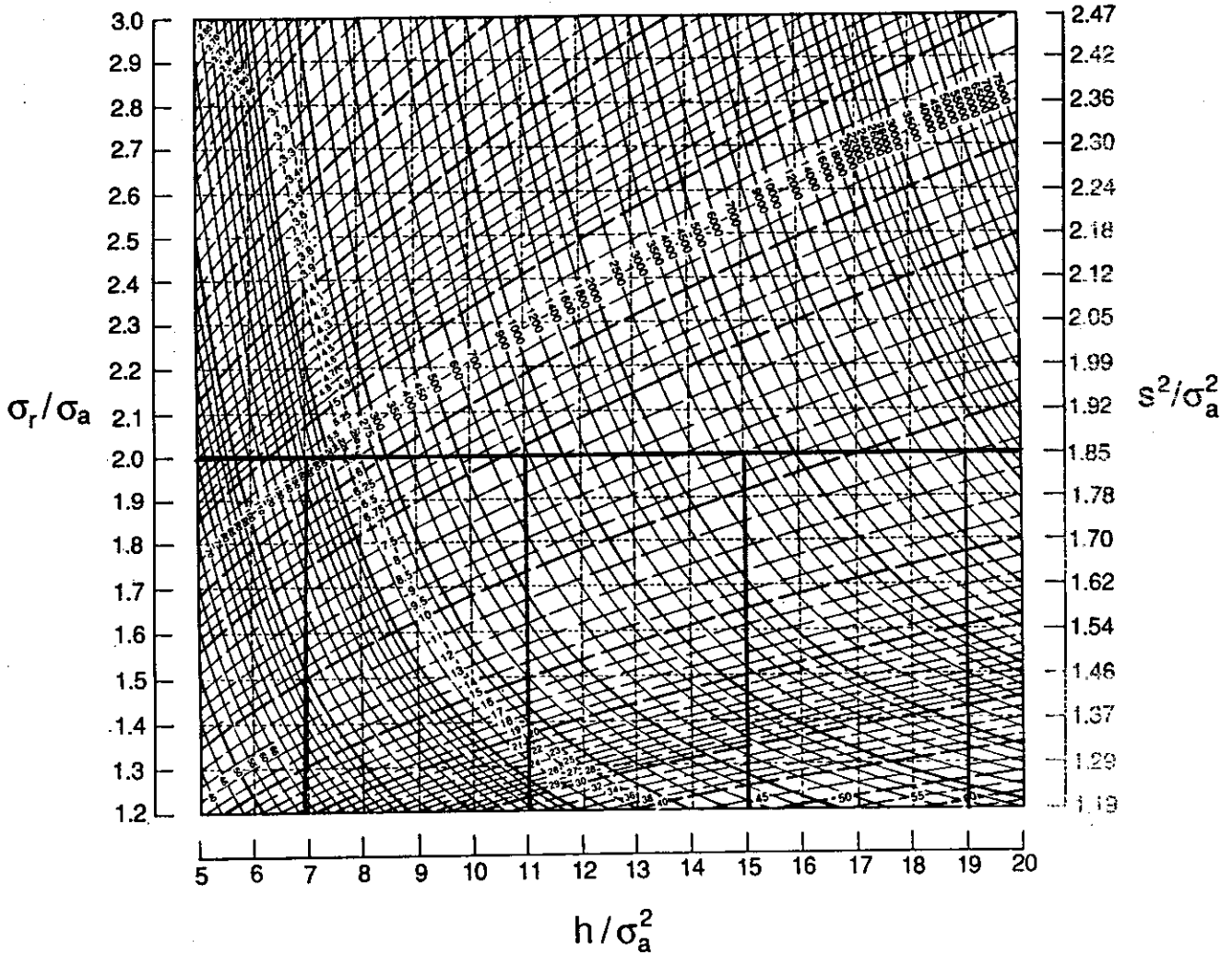
The design of a cumulative sum chart consists of finding the values of  $h$  and  $s^2$  that approximately yield predetermined values of the average run length  $L_a$  and  $L_r$ . To help in the design of the charts a contour nomogram with contour lines of constant  $L_a$  and  $L_r$  as a function of  $h$  and  $s^2$ , provides a flexible way of selecting the CUSUM parameters  $h$  and  $s^2$ .

The nomogram was constructed using a two dimensional grid with values of  $\sigma_r/\sigma_a$  ranging from 1.2 to 3 in steps of 0.1; i.e., displacements of the process variability from 20% to 200%, and values of  $h/\sigma_a^2$  ranging from 5 to 20 in steps of 1. The values of  $L_a$  and  $L_r$  at the grid points were obtained by solving the integral equations using the methods described in section 2.3. Using these ARL values the contours of  $L_a$  from 40 to 75000, and the contours of  $L_r$  from 2.5 to 65, were computed using bivariate interpolation (Akima 1978), as implemented in the statistical software S.

Figure 2.5 shows the contours of  $L_a$  and  $L_r$  superimposed and the values of  $\sigma_r/\sigma_a$ ,  $h/\sigma_a^2$  and  $s^2/\sigma_a^2$ . Since the  $s^2$  scale is nonlinear, is more accurate to use formula (2.5) to obtain  $s^2$  than trying to interpolate in that scale.

In the next chapter we will see how the contour nomogram can be used to design cumulative sum charts aimed to monitor increases or decreases in the process variability.

**Figure 2.5** Contour nomogram with constant lines of  $L_a$  (solid line), and  $L_r$  (dotted line). Axis represent the decision interval  $h$ , the reference value  $s^2$  and the shift in standard deviation  $\sigma_r / \sigma_a$ .



## 2.5 Comparison with Shewhart R and S Charts

Shewhart range charts, see for example Duncan (1965), are designed to show variations in the ranges of samples with the aim of detecting changes in process variability. The range of samples of  $n$  items is calculated and plotted on the R-chart. If the sample range does not fall outside the  $3\hat{\sigma}_R$  control limits the process variability is said to be in control. If a point falls outside the limits the process variability is said to be out of control.

Range charts, along with  $\bar{X}$ -charts, have been the most widely used control scheme for the last sixty years. They are simple to use and in general will detect large shifts in the variance. Cumulative sum charts for monitoring the mean, see for example Barnard (1959), Woodward and Goldsmith (1964), or Lucas (1976), work better than Shewhart charts for detecting small shifts of the process level from the target value. We compare the R-charts with the cumulative sum charts for variance in terms of the average run length or the expected number of observations that need to be taken before a lack of control is indicated.

A comparison based on the average run length values begins by matching the two charts so both have the same ARL value at the acceptable quality level. The ARL values are then compared for various shifts,  $\sigma_r = k\sigma_a$ , of the process variability.

Since most often samples of size 4 or 5 are used when designing Shewhart control charts, we base our comparisons using this two sample sizes. For the

comparisons we assume that false alarms occur approximately every 200 samples. This implies that for the sample size 5 the ARL at  $\sigma_a$  is set at approximately 1000 while for the sample size 4 it is set at approximately 800. Appropriate CUSUM charts with these two  $L_a$  values were selected, with the help of the nomogram, giving the following parameters

Sample Size	$L_a$	$h$	$s^2$
4	800	11.75	1.62
5	1000	11.60	1.85

**Table 2.1** Parameters for the CUSUM charts with  $L_a = 800$  and 1000

The values of  $s^2 = 1.62$  and  $s^2 = 1.85$  corresponds to CUSUM charts designed to detect increases in variability of 70% and 100% respectively. In other words, by putting  $\sigma_r = 1.7$  and  $\sigma_r = 2$  and applying formula 2.5 with  $\sigma_a = 1$ , we obtain these values of  $s^2$ .

The cumulative sum chart is compared with two R-charts. The traditional R-chart with action limits set at  $B_1 \sigma$ , and the same chart with warning limits at  $B_2 \sigma$ ; where  $B_2 \leq B_1$ . The results indicate that the addition of warning limits do not improve the performance of the chart. In section 2.5.3 the CUSUM chart is compared with an S-chart with action and warning limits.

### 2.5.1 R-chart with $B_1 \sigma$ action limits

In order to compute the ARL for the Shewhart chart we need to compute the probability that the range, computed from a sample of size 4 or 5, will be outside of the upper control limit  $B_1\sigma$ . This is equivalent to computing the probability that the standardized normal range  $W = \frac{X_{\max} - X_{\min}}{\sigma}$  is greater than or equal than the upper control limit divided by the process standard deviation.

If  $d_2$  and  $d_3$  denote the mean and standard deviation of  $W$  respectively, the upper control limit (UCL) of the R-chart is given by  $(1 + 3\frac{d_3}{d_2})\bar{R}$ . Dividing this quantity by  $\sigma_r = k\sigma_a = k\frac{\bar{R}}{d_2}$ , we have  $B = \frac{d_2 + 3d_3}{k}$ . For the sample sizes 4 and 5, the values of  $d_2$  and  $d_3$  were obtained from table D1 of Duncan (1965), and are given in table 2.2.

Sample Size	$d_2$	$d_3$
4	2.059	0.8798
5	2.326	0.8641

**Table 2.2** Values of the parameters  $d_2$  and  $d_3$  for the UCL of the R-chart

The average run length of the Shewhart R-chart for a sample of size  $N$  is given by

$$ARL \text{ (Shewhart Chart)} = \frac{1}{1 - P} \times N, \quad (2.18a)$$

where  $P = Pr(W \leq B)$ .

For these values of  $d_1$  and  $d_2$  we obtain values of  $B_1$  equal to 4.698 and 4.918 for  $n = 4$  and  $n = 5$  respectively. The choice of  $B_1 = 4.918$  for  $n = 5$  gives an  $L_a \approx 1085$ . For the comparisons we use  $B_1 = 4.886$  which gives  $L_a \approx 1001$ .

Barnard (1978) gives an algorithm to compute the probability integral of the normal range given the sample size. Using this algorithm a small Fortran program was written to compute the ARL values for the R-chart at various shifts of the process variability by assigning different values to  $k$  the ratio of the new process variability,  $\sigma_r$ , to the old process variability,  $\sigma_a$ .

The average run length values for the cumulative sum chart were obtained by solving the integral equations (2.6) and (2.7) with the values of  $h$  and  $s^2$  given in table 2.1. These values were calculated for shifts in standard deviation from 10% to 200% and are shown in table 2.3.

### 2.5.2 R-chart with warning limits

Several improvements to the Shewhart charts have been suggested; in particular the addition of warning limits at  $B_2\sigma$  where  $B_2 < B_1$ . The idea is to monitor the number of consecutive points that fall in the warning region between  $B_2\sigma$  and  $B_1\sigma$  and declare that the process is out of control whenever we observe  $M$  of such points. In practice  $B_1 = 3$  and  $B_2 = 2$ , the usual  $3\sigma$  and  $2\sigma$  limits, and we look for  $M = 2$  consecutive points falling in the warning region.

**Table 2.3** Values of  $L_r$  for Cusum Charts and R-Charts with action limit at  $B_1 \sigma_a$ ,  
for a fixed  $L_a$  and an increases in variability equal to  $\sigma_r = k \sigma_a$ .

Average Run Length at $\sigma_r$				
Increase in Variability	$L_a = 800$		$L_a = 1000$	
	Cusum Chart	R-Chart	Cusum Chart	R-Chart
$\frac{\sigma_r}{\sigma_a} = k$	$h = 11.75$ $s^2 = 1.62$	$B_1 = 4.698$	$h = 11.75$ $s^2 = 1.85$	$B_1 = 4.886$
1.0	809.04	809.75	1022.06	1001.08
1.1	186.47	297.26	264.83	343.74
1.2	74.18	139.09	100.67	153.61
1.3	39.40	77.27	50.37	82.72
1.4	25.05	48.64	30.41	51.01
1.5	17.87	33.59	20.83	34.79
1.6	13.73	24.89	15.53	25.60
1.7	11.10	19.47	12.27	19.96
1.8	9.30	15.88	10.11	16.29
1.9	8.00	13.40	8.59	13.78
2.0	7.03	11.62	7.47	11.99
2.1	6.28	10.29	6.61	10.68
2.2	5.68	9.28	5.94	9.68
2.3	5.19	8.49	5.40	8.91
2.4	4.79	7.86	4.96	8.31
2.5	4.45	7.35	4.59	7.82
2.6	4.16	6.93	4.28	7.43
2.7	3.92	6.59	4.02	7.10
2.8	3.70	6.30	3.79	6.84
2.9	3.52	6.05	3.59	6.61
3.0	3.35	5.84	3.42	6.42

If  $p_1 = Pr(W \leq B_2)$  and  $p_2 = Pr(B_2 \leq W \leq B_1)$ , denote the probabilities that the range falls below the warning limit and in the warning region respectively, calculated using the algorithm given by Barnard, the average run length for the R-chart with action limits at  $B_1\sigma$  and warning limits at  $B_2\sigma$ , see Page (1963), is given by

$$ARL = \frac{1 - p_2^M}{1 - p_2 - p_1(1 - p_2^M)} \times \text{sample size} . \quad (2.18b)$$

Average run length values for the R-chart were computed using  $M = 2$  and values of  $B_1$  and  $B_2$  and different shifts in the process variability as shown in table 2.4.

Figures 2.6a and 2.6b show the ARL values, in logarithmic scale, for the CUSUM chart and both the R-chart with  $B_1\sigma$  action limits and the R-chart with action and warning limits, for samples of size 4 and 5 respectively. Contrary to what is expected, the R-chart with warning limits does not perform better than the standard R-chart. There is only a slight improvement in the region  $1.3 \leq \sigma_r/\sigma_a \leq 2$ , and for large shifts,  $\sigma_r/\sigma_a > 2$ , the two charts are equivalent.

The ARL curve for the cumulative sum chart is below the curve for both R-charts, indicating that the CUSUM chart not only detects small shifts faster but, as opposed to the chart for means, also performs better than the Shewhart charts in detecting large shifts.

Some people may argue that the sensitivity of the R-chart to detect small changes in the process variability can be improved, by adding some other warning limits and rules e.g.  $1\sigma$  limits, as to make as sensitive as the CUSUM chart. This may be the case but after adding three warning limits and five extra rules the simplicity of the R-chart is lost, while the CUSUM remains simple to use with only one upper limit.

In comparison with the R-charts, the CUSUM procedure performs better over the whole range of process variability  $1.0 \leq \frac{\sigma_r}{\sigma_a} \leq 3.0$ . These results agree with the ones given by Page (1963), although he used the cumulative sum of ranges instead of the deviations  $(x_i - \mu)^2$ . This contrasts with the results obtained by comparing  $\bar{X}$ -charts and CUSUM for the mean, in which the CUSUM performs better than the  $\bar{X}$ -chart but only for small displacements of the mean, usually for displacements less than 2 standard deviations.

In the next section we compare the CUSUM procedure with the Shewhart chart for standard deviations, S-chart, that uses action limits at  $B_1\sigma$ , warning limits at  $B_2\sigma$ , and a rule of 2 consecutive points falling in the warning region.

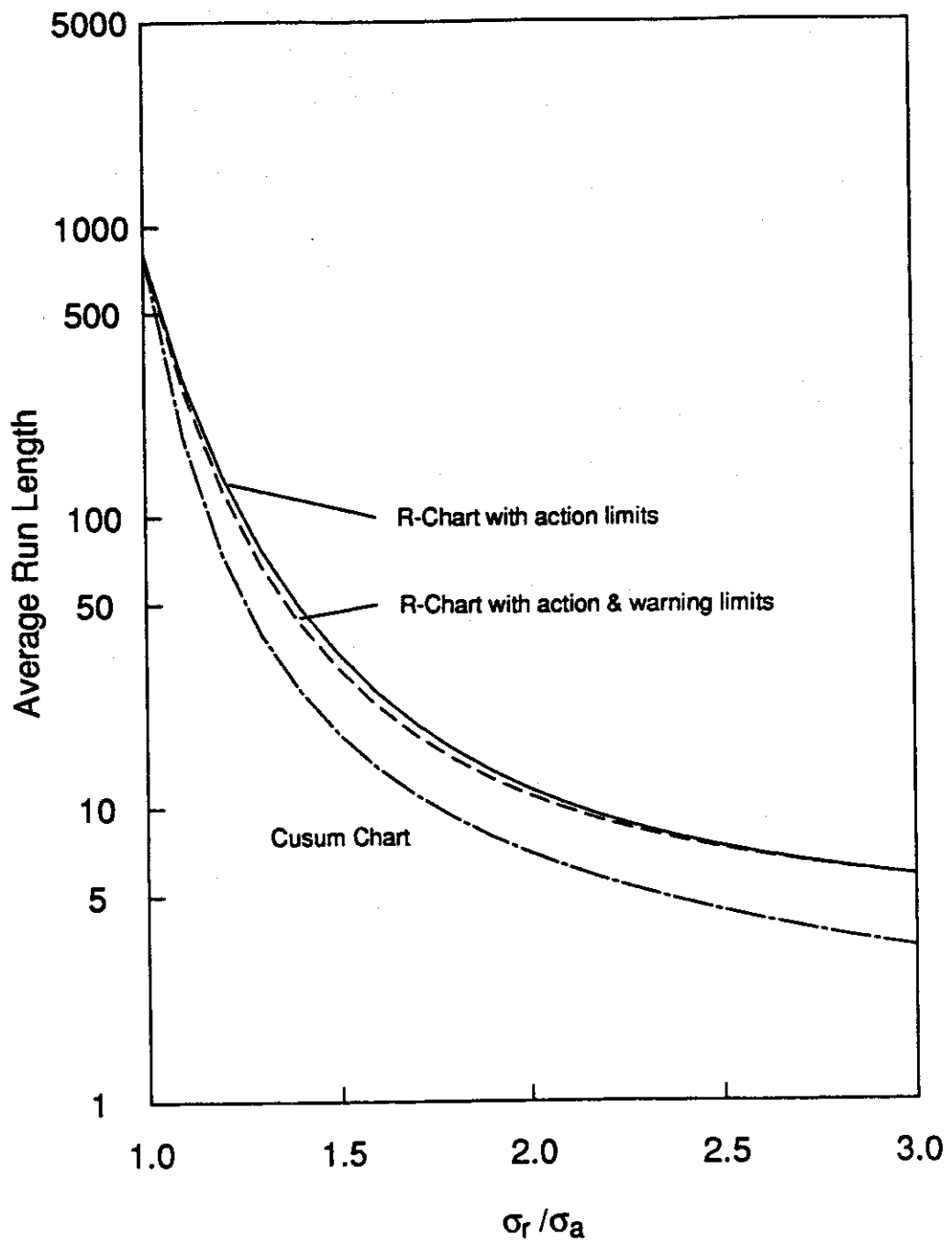
### 2.5.3 S-chart with warning limits

In the S-chart the standard deviation of samples of  $n$  observations are plotted and monitored until they cross the  $B_1\sigma$  action limits which indicates a lack of control. As with R-charts, a  $B_2\sigma$  warning limit can be added and consecutive points

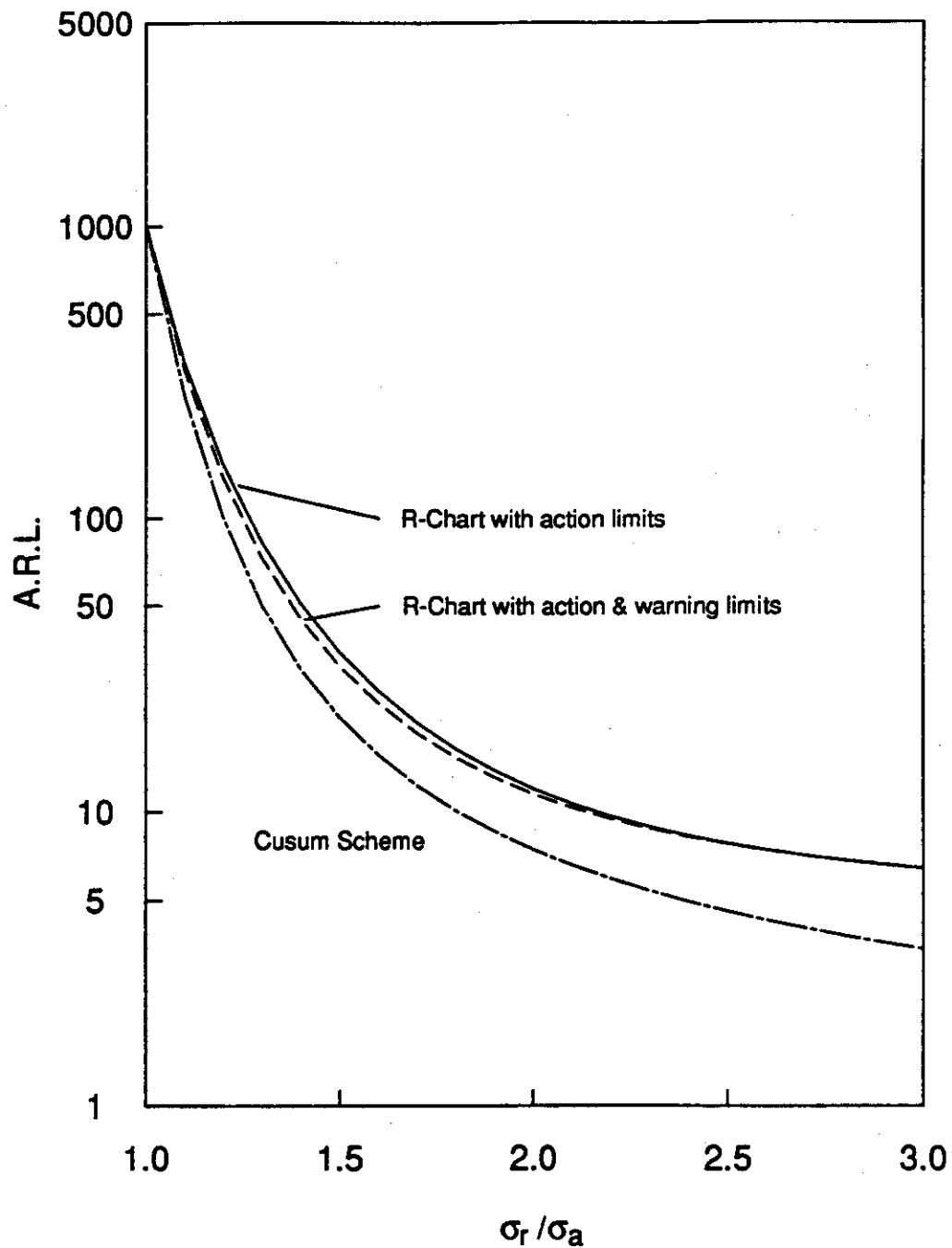
**Table 2.4** Values of  $L_r$  for Cusum Charts and R-Charts with action limits at  $B_1\sigma_a$  and warning limits at  $B_2\sigma_a$ , for a fixed  $L_a$  and an increases in variability equal to  $\sigma_r = k\sigma_a$ .

Average Run Length at $\sigma_r$				
Increase in Variability	$L_a \approx 800$		$L_a \approx 1000$	
	Cusum Chart	R-Chart	Cusum Chart	R-Chart
$\frac{\sigma_r}{\sigma_a} = k$	$h = 11.75$ $s^2 = 1.62$	$B_1 = 4.843$ $B_2 = 3.713$	$h = 11.75$ $s^2 = 1.85$	$B_1 = 5.01$ $B_2 = 3.98$
1.0	809.04	808.14	1022.06	1028.86
1.1	186.47	275.31	264.83	324.30
1.2	74.18	124.02	100.67	138.49
1.3	39.40	67.97	50.37	73.30
1.4	25.05	42.83	30.41	45.21
1.5	17.87	29.85	20.83	31.14
1.6	13.73	22.41	15.53	23.24
1.7	11.10	17.78	12.27	18.41
1.8	9.30	14.71	10.11	15.25
1.9	8.00	12.57	8.59	13.07
2.0	7.03	11.03	7.47	11.51
2.1	6.28	9.87	6.61	10.35
2.2	5.68	8.97	5.94	9.47
2.3	5.19	8.27	5.40	8.77
2.4	4.79	7.70	4.96	8.22
2.5	4.45	7.24	4.59	7.78
2.6	4.16	6.86	4.28	7.41
2.7	3.92	6.54	4.02	7.10
2.8	3.70	6.27	3.79	6.85
2.9	3.52	6.04	3.59	6.63
3.0	3.35	5.85	3.42	6.45

Figure 2.6a Comparison of A.R.L. curves for Cusum Chart and Shewhart R-Charts when the subgroup size is equal to 4.



**Figure 2.6b** Comparison of A.R.L. curves for Cusum Chart and Shewhart R-Charts when the subgroup size is equal to 5.



falling in this region are used as rule for lack of control.

As before  $p_1$  and  $p_2$  denote the probabilities that the standard deviation falls below the warning limit  $B_2\sigma$ , and in the warning region between  $B_2\sigma$  and  $B_1\sigma$ . The average run length values are calculated by means of equation 2.18b using the chi-squared distribution to compute these probabilities.

The results for subgroups of size 4 and 5 are given in table 2.5. Comparing these results with the ones in table 2.4 we see that the S-chart with warning limits has a similar rate of detection as the R-chart with warning limits. This is not surprising since for samples of size 4 and 5 the range is 98% and 95% efficient respectively.

A plot of the ARL curves is given in figures 2.7a and 2.7b. Note that as with the R-chart the CUSUM outperforms the S-chart in the entire range of shifts in process variabilities  $1.0 \leq \frac{\sigma_r}{\sigma_a} \leq 3.0$ . The main conclusion that can be drawn from these comparisons is that when monitoring process variability the CUSUM is better than both R and S charts.

## 2.6 The Effect of Non-normality

In section 2.1 we focused on the sequential monitoring of variances under the assumption that the observations coming from the process under study are normally distributed. Although this is approximately true when the errors arise from a variety of independent sources of which none are dominant, there are cases in

**Table 2.5** Values of  $L_r$  for Cusum Charts and S-Charts with action limits at  $B_1\sigma_a$  and warning limits at  $B_2\sigma_a$ , for a fixed  $L_a$  and an increase in variability equal to  $\sigma_r = k\sigma_a$ .

Average Run Length at $\sigma_r$				
Increase in Variability	$L_a = 800$		$L_a = 1000$	
	Cusum Chart	S-Chart	Cusum Chart	S-Chart
$\frac{\sigma_r}{\sigma_a} = k$	$h = 11.75$ $s^2 = 1.62$	$B_1 = 1.815$ $B_2 = 1.485$	$h = 11.75$ $s^2 = 1.85$	$B_1 = 1.75$ $B_2 = 1.45$
1.0	809.04	799.08	1022.06	1023.24
1.1	186.47	270.29	264.83	310.73
1.2	74.18	121.07	100.67	130.17
1.3	39.40	66.05	50.37	68.29
1.4	25.05	41.48	30.41	41.99
1.5	17.87	28.84	20.83	28.92
1.6	13.73	21.61	15.53	21.62
1.7	11.10	17.13	12.27	17.17
1.8	9.30	14.17	10.11	14.27
1.9	8.00	12.11	8.59	12.28
2.0	7.03	10.62	7.47	10.85
2.1	6.28	9.51	6.61	9.79
2.2	5.68	8.66	5.94	8.99
2.3	5.19	7.99	5.40	8.36
2.4	4.79	7.45	4.96	7.86
2.5	4.45	7.01	4.59	7.45
2.6	4.16	6.65	4.28	7.12
2.7	3.92	6.35	4.02	6.85
2.8	3.70	6.10	3.79	6.62
2.9	3.52	5.88	3.59	6.43
3.0	3.35	5.69	3.42	6.26

Figure 2.7a Comparison of A.R.L. curves for Cusum Chart and Shewhart S-Chart with warning limits when the subgroup size is equal to 4.

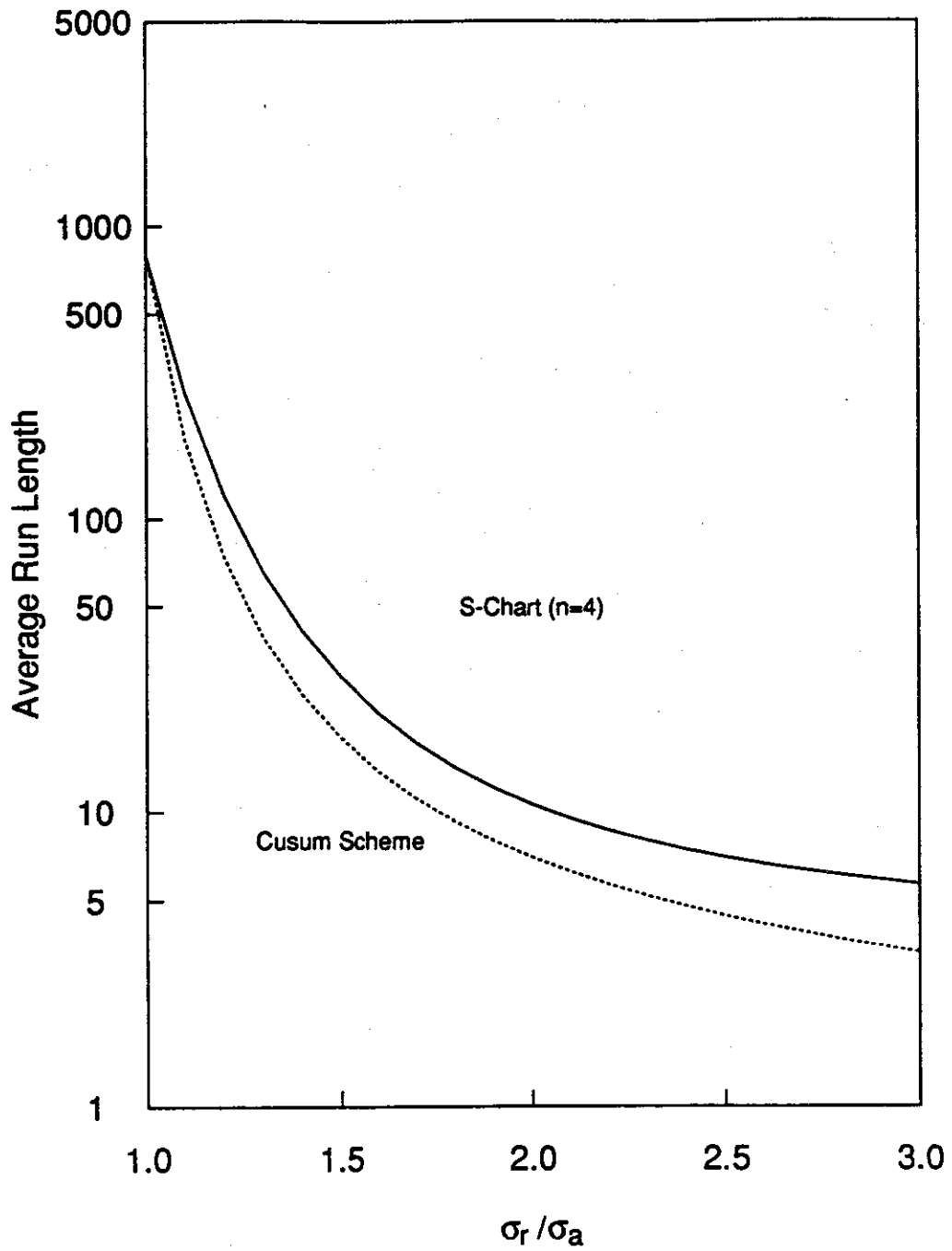
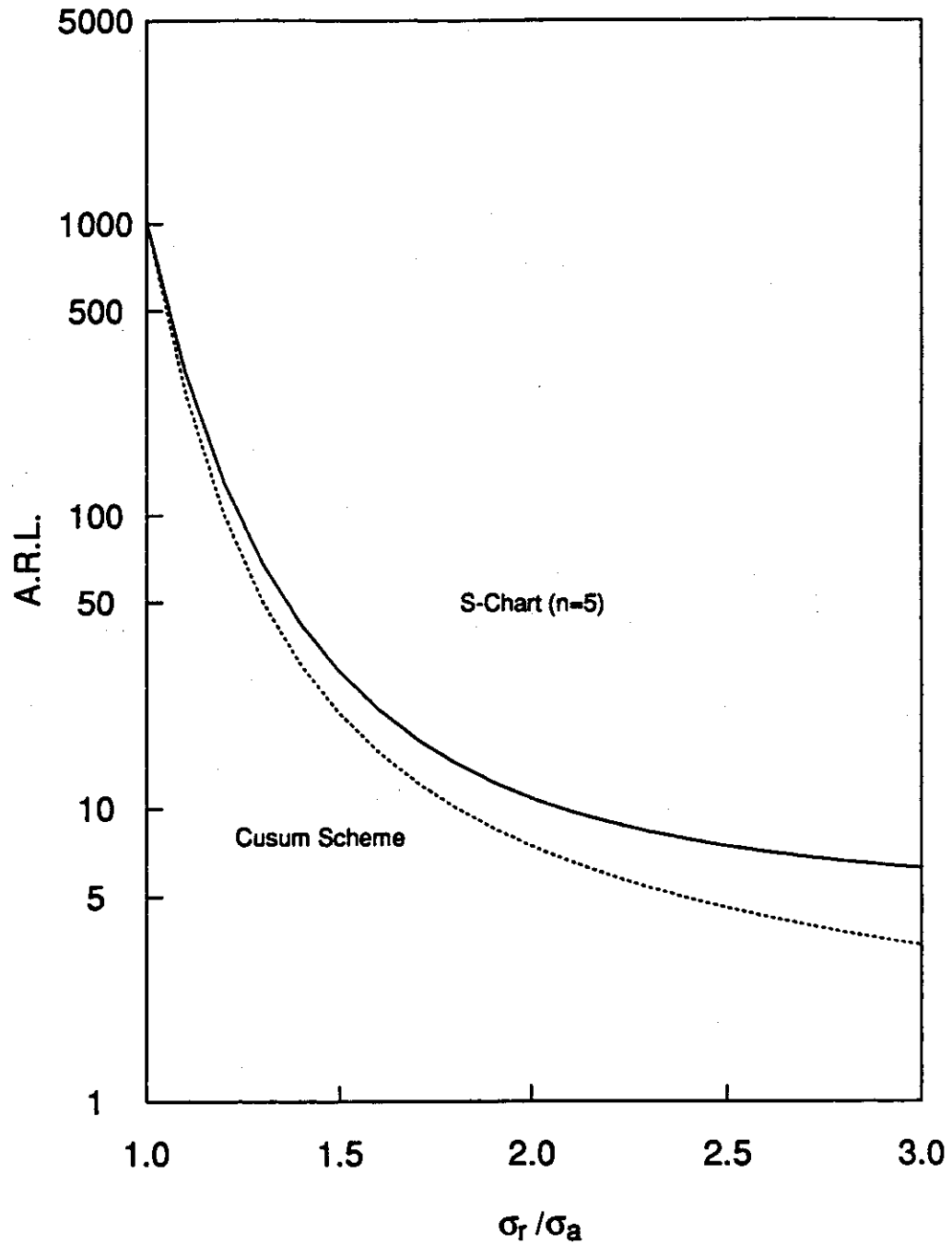


Figure 2.7b Comparison of A.R.L. curves for Cusum Chart and Shewhart S-Chart with warning limits when the subgroup size is equal to 5.



which the measurements are not normally distributed.

By considering distributions other than the normal we can evaluate the extent to which the average run length depends upon the normality assumption. A convenient class of distributions that are center around the normal is the class of exponential power distributions.

### 2.6.1 Exponential Power Distribution

The exponential family of distributions, see Box and Tiao (1973), Turner (1960), is class of symmetric distributions that includes the Normal, together with other distributions that are more leptokurtic; e.g. the double exponential, and distributions that are more platykurtic; e.g. the rectangular distribution.

The members of this class are three-parameter distributions with density given by

$$f(y|\mu, \sigma, \gamma) = \frac{\omega(\gamma)}{\sigma} \exp \left[ -c(\gamma) \left| \frac{y - \mu}{\sigma} \right|^{\frac{2}{1+\gamma}} \right], \quad -\infty < y < \infty, \quad (2.19)$$

where

$$\omega(\gamma) = \frac{\Gamma[\frac{3}{2}(1+\gamma)]^{1/2}}{(1+\gamma) \Gamma[\frac{1}{2}(1+\gamma)]^{3/2}}$$

and

$$c(\gamma) = \left[ \frac{\Gamma[\frac{3}{2}(1+\gamma)]}{\Gamma[\frac{1}{2}(1+\gamma)]} \right]^{\frac{1}{1+\gamma}} \quad \sigma > 0, \quad -\infty < \mu < \infty, \quad -1 < \gamma \leq 1.$$

The parameters  $\mu$  and  $\sigma$  are the mean and standard deviation of the population, and the parameter  $-1 < \gamma \leq 1$  can be regarded as a measure of kurtosis indicating the extent of "non-normality", see figure 2.8, of the population. In particular, when  $\gamma=0$ ,  $c(0) = 1/2$ ,  $\omega(0) = 1/\sqrt{2}\pi$  corresponding to the normal distribution, when  $\gamma=1$  the distribution is the double exponential

$$f(y|\mu, \sigma, \gamma) = \frac{1}{\sqrt{2}\sigma} \exp\left[-\sqrt{2} \left| \frac{y-\mu}{\sigma} \right|^2\right], \quad -\infty < y < \infty, \quad (2.20a)$$

and when  $\gamma$  tends to  $-1$  the distribution tends to the rectangular distribution

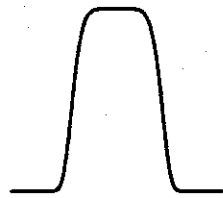
$$\lim_{\gamma \rightarrow -1} f(y|\mu, \sigma, \gamma) = \frac{1}{2\sqrt{3}\sigma}, \quad \mu - \sqrt{3}\sigma < y < \mu + \sqrt{3}\sigma. \quad (2.20b)$$

Although no explicit solutions exist for the likelihood equations, maximum likelihood estimators for  $(\mu, \sigma, \gamma)$ , can be obtained using minimization techniques in which the variables are subject to simple bounds.

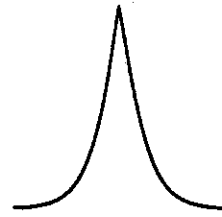
### 2.6.2 Sequential Test

For fixed values of  $\mu$  and  $\gamma$ , the sequential probability ratio test for testing  $\sigma^2 = \sigma_a^2$  versus  $\sigma^2 = \sigma_r^2$  simplifies to

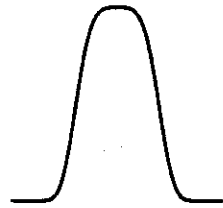
**Figure 2.8** Exponential power distributions with common standard deviation, ( $\sigma = 1$ ), for different values of  $\gamma$ .



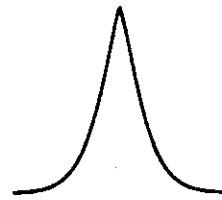
$\gamma = -0.7$



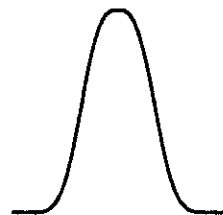
$\gamma = 0.7$



$\gamma = -0.5$



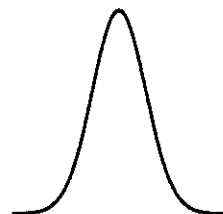
$\gamma = 0.5$



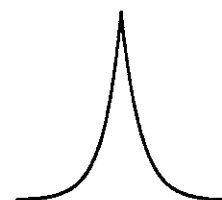
$\gamma = -0.3$



$\gamma = 0.3$



$\gamma = 0$   
normal



$\gamma = 1$   
double exponential

$$LR_k = \left[ \frac{\sigma_a}{\sigma_r} \right]^n \exp \left[ c(\gamma) \left[ \frac{1}{\sigma_a^{1+\gamma}} - \frac{1}{\sigma_r^{1+\gamma}} \right] \sum_{i=1}^k |x_i - \mu|^{\frac{2}{1+\gamma}} \right] \quad (2.21)$$

After taking logarithms and letting  $h_0$  and  $h_1$  denote the lower and upper limits of the sequential test we have that the test leads to the cumulative sum

$$h_0 + k s^2(\gamma) \leq \sum_{i=1}^k |x_i - \mu|^{\frac{2}{1+\gamma}} \leq h_1 + k s^2(\gamma) , \quad (2.22)$$

where

$$s^2(\gamma) = \frac{\ln\left(\frac{\sigma_r}{\sigma_a}\right)}{c(\gamma) \left[ \frac{1}{\sigma_a^{1+\gamma}} - \frac{1}{\sigma_r^{1+\gamma}} \right]} . \quad (2.23)$$

When  $\gamma = 0$  (normal case), equations (2.22) and (2.23) coincide with equations (2.3) and (2.5) respectively. Note that the slope  $s^2(\gamma)$  depends not only on the two values of the variability of the process,  $\sigma_a$  and  $\sigma_r$ , but also on the non-normality parameter  $\gamma$  of which it is a decreasing function.

As we have seen, the cumulative sum scheme is equivalent to a sequence of Wald sequential tests with boundaries  $(0, h)$ . For the case in which the interest centers around monitoring increases in variability, i.e.  $\sigma_a \leq \sigma_r$ , the cumulative sum procedure consists of plotting or charting, for each observation  $x_k$ , the quantity

$$\begin{aligned}
CS_k^+ &= \max(0, CS_{k-1}^+ + r_k) , & CS_0^+ &= 0 , \\
r_k &= |x_i - \mu| \frac{2}{1+\gamma} - s^2(\gamma) ;
\end{aligned}
\tag{2.24}$$

until it either exceeds the decision interval  $h$  or reaches zero again.

As we showed in section 2.2, the choice of the reference value  $s^2(\gamma)$  and the decision interval  $h$  would depend on some predetermined average run lengths.

### 2.6.3 Average Run Length

Recall from section 2.2 that the average run length for a sequential test starting at  $z$  is given by

$$\text{ARL} = \frac{N(0)}{1 - P(0)} ,$$

where  $P(z)$  is the probability that a return to zero occurs, and  $N(z)$  the average sample number of the test.

Average run length values for different values of the non-normality parameter  $\gamma$  were calculated using the methods described in section 2.3, in which the integral equations for  $P(z)$  and  $N(z)$  are replaced by a system of linear equations.

As was pointed out in section 2.2 the equations for  $P(z)$  and  $N(z)$  depend on the density of the increments  $r_k = |x_i - \mu| \frac{2}{1+\gamma} - s^2(\gamma)$ , in the cumulative sum  $CS_k^+$ . Under the class of exponential power distributions this density becomes:

$$f(y) = (1 + \gamma) \frac{\omega(\gamma)}{\sigma} \left[ y + s^2(\gamma) \right]^{\frac{\gamma-1}{2}} \exp \left[ \frac{-c(\gamma)(y + s^2(\gamma))}{\frac{2}{\sigma^{1+\gamma}}} \right], \quad (2.25)$$

for  $y \geq -s^2(\gamma)$  and

$$f(y) = 0 \quad \text{for} \quad y < -s^2(\gamma).$$

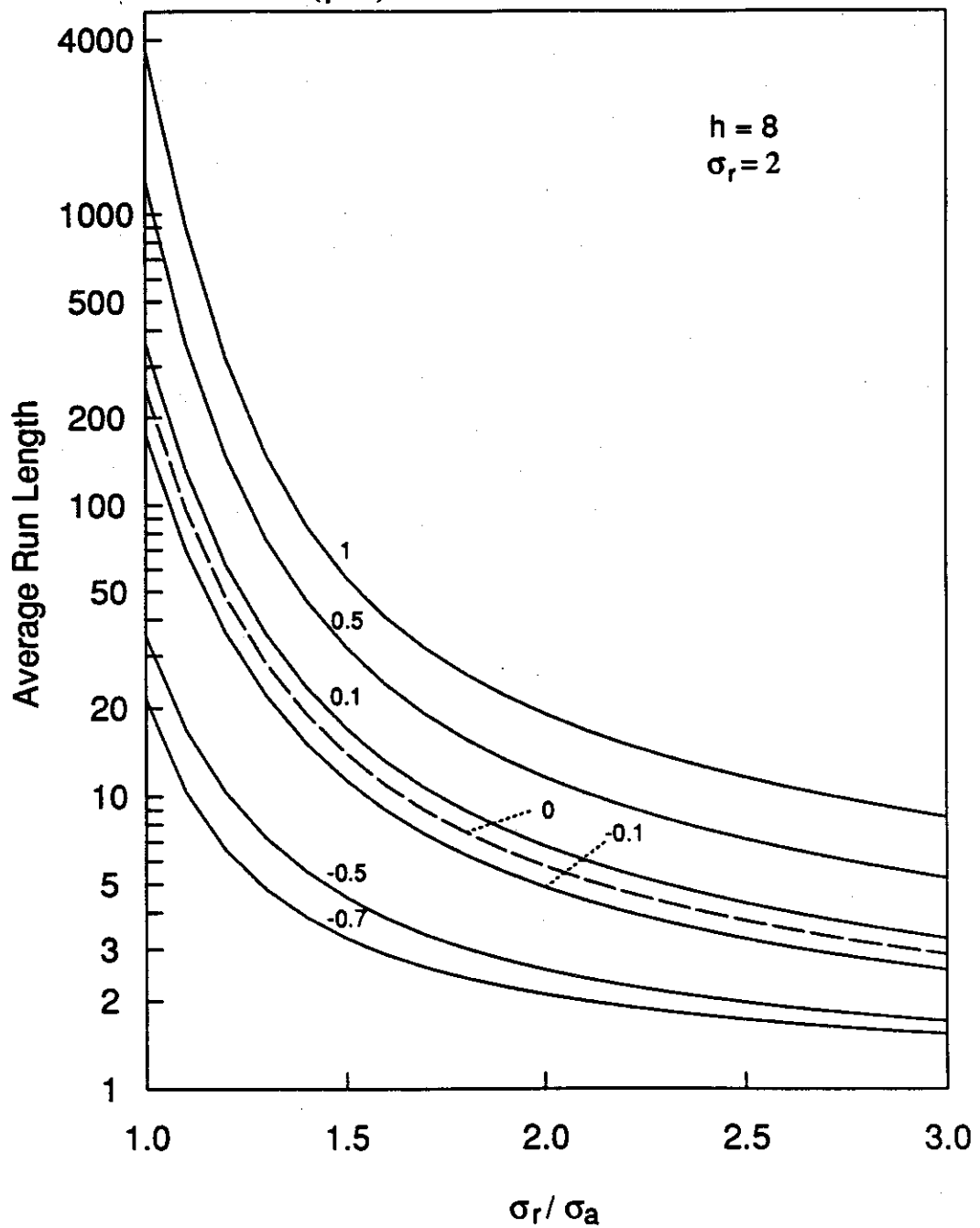
Fortran code was written to obtain values of  $P(0)$ ,  $N(0)$  and the corresponding ARL values, using the IMSL subroutines DLSARG, to solve the system of linear equations, and DQAG to integrate numerically the density (2.25), and was runned in a DEC VAXstation 3100.

Tables of average run length at the acceptable quality level,  $L_a$ , and at the rejectable quality level,  $L_r$ , for values of  $\gamma$  starting at  $-0.7$  in increments of  $0.2$  were constructed and are given in the appendix. These tables can be used to design cumulative sum charts for monitoring variances when the observations are distributed according to a member of the class of exponential power distributions.

In chapter 3 will discuss how to use these tables to design cumulative sum charts and give two formulas that give approximate ARL values as a function of the decision interval  $h$ , the ratio of variances  $\frac{\sigma_r}{\sigma_a}$  and the non-normality parameter  $\gamma$ .

Figure 2.9 shows the average run length values for a CUSUM chart designed to detect increases in variability of at least 100%; i.e.  $\sigma_r/\sigma_a = 2$ , with a decision interval  $h = 8$ , and values of  $\gamma = -0.7, -0.5, -0.1, 0, 0.1, 0.5, 0.7$  and  $1$ . It is clear

Figure 2.9 Average Run Length for a cusum chart with parameters  $h = 8$ ,  $\sigma_r = 2$  and different values of the non-normality parameter  $\gamma$ . The broken line corresponds to the normal distribution ( $\gamma = 0$ ).



from the plot that the assumption of normality can lead to underestimated ARL values for  $\gamma \geq 0.1$ , and overestimated ARL values for  $\gamma \leq -0.1$ . For example, if we assume that the observations are distributed as normal when in fact their distribution is leptokurtic with  $\gamma = 0.5$ , say; then we will be underestimating the  $L_a$  value, since for  $\gamma = 0$  we have that  $L_a \approx 250$ , while for  $\gamma = 0.5$  the  $L_a$  is approximately 1200. On the other hand, if the distribution is platykurtic as given by  $\gamma = -0.5$ , the ARL would be overestimated since  $L_r \approx 4$  for  $\gamma = 0$ , while for the normal distribution  $L_r \approx 2$ .

#### 2.6.4 Student $t$ distribution

In section 2.6.1 we used the exponential power family of distributions to evaluate the effect that kurtosis has in the average run length of the CUSUM procedure for monitoring changes in the variability of a process. The main features of this family of distributions are its symmetry around a mean value and that the normal distribution is a member of the family.

We can think of the  $t$  distributions as forming a family of distributions indexed by the degrees of freedom. This is a family of bell-shape symmetric distributions of which the normal, when the number of degrees of freedom approaches infinity, is a member.

The purpose of this section is to derive a CUSUM procedure for monitoring variances assuming that the observations are distributed according to the  $t$  distribution, and to show how the exponential power family of distributions can be used to

approximate the results obtained by using the  $t$  distribution.

The p.d.f. of a  $t$  distribution with  $v$  degrees of freedom, location parameter  $\mu$  and scale parameter  $\sigma^2$  is given by

$$f(x|\mu, \sigma^2, v) = \frac{\Gamma(\frac{v+1}{2})}{\sigma\sqrt{\pi v}\Gamma(\frac{v}{2})} \left[ 1 + \frac{(x-\mu)^2}{v\sigma^2} \right]^{-\frac{v+1}{2}}; \quad (2.26)$$

where

$$-\infty \leq \mu \leq \infty, \quad \sigma^2 \geq 0, \quad \text{and} \quad v > 0.$$

The sequential probability ratio test for testing an increase in variance from  $\sigma_a^2$  to  $\sigma_r^2$ , leads to the likelihood ratio

$$LR_k = \left[ \sigma_a / \sigma_r \right]^k \prod_{i=1}^k \left[ \left[ 1 + \frac{(x_i - \mu)^2}{v\sigma_a^2} \right] / \left[ 1 + \frac{(x_i - \mu)^2}{v\sigma_r^2} \right] \right]^{\frac{v+1}{2}}. \quad (2.27)$$

After taking logarithms, the likelihood ratio simplifies to the cumulative sum

$$\sum_{i=1}^k \log \left[ \left[ 1 + \frac{(x_i - \mu)^2}{v\sigma_a^2} \right] / \left[ 1 + \frac{(x_i - \mu)^2}{v\sigma_r^2} \right] \right] - ks^2(v). \quad (2.28)$$

Where  $s^2(v)$  is a reference value, as described in section 2.1, that depends on the degrees of freedom  $v$  and is given by

$$s^2(v) = \frac{\log(\sigma_r^2 / \sigma_a^2)}{v+1}. \quad (2.29)$$

To graphically monitor increases in the variability, we can plot the cumulative sums in 2.28 and declare that a change has occurred whenever this sum is greater or equal than a decision interval  $h$  equal to

$$h = \frac{2 \log\left(\frac{1-\beta}{\alpha}\right)}{\nu+1} \quad (2.30)$$

Where  $\alpha$  and  $\beta$  are the probabilities of making errors of type I and II respectively.

When the number of degrees of freedom tends to infinity, the cumulative sum procedure given by 2.28, 2.29 and 2.30, tends to the one given in section 2.1 (equation 2.3). To see this, note that for  $\epsilon$  small,  $\log(1+\epsilon) = \epsilon$ . Therefore when  $\nu \rightarrow \infty$

the expression  $\frac{(x_i - \mu)^2}{\nu \sigma_a^2} \rightarrow 0$ , and we have

$$\log \left[ \frac{\left[ 1 + \frac{(x_i - \mu)^2}{\nu \sigma_a^2} \right]}{\left[ 1 + \frac{(x_i - \mu)^2}{\nu \sigma_r^2} \right]} \right] = \frac{(x_i - \mu)^2}{\nu} \left[ \frac{1}{\sigma_a^2} - \frac{1}{\sigma_r^2} \right]$$

Now the cumulative sum procedure with decision interval  $h$ , given by equation 2.30, can be written as

$$\left[ \frac{1}{\sigma_a^2} - \frac{1}{\sigma_r^2} \right] \sum_{i=1}^k (x_i - \mu)^2 - \frac{\nu k \log(\sigma_r^2 / \sigma_a^2)}{\nu + 1} \leq \frac{2\nu \log\left(\frac{1-\beta}{\alpha}\right)}{\nu + 1}$$

Which gives the right hand side of equation 2.3 when  $\nu \rightarrow \infty$ .

In section 2.6.1 we indicated that for the exponential power distribution the parameter  $\gamma$  can be considered as a measure of kurtosis. In fact, the kurtosis  $\gamma_2$  can

be written as a function of the non-normality parameter  $\gamma$

$$\gamma_2 = \frac{\Gamma[\frac{5}{2}(1+\gamma)]\Gamma[\frac{1}{2}(1+\gamma)]}{\Gamma[\frac{3}{2}(1+\gamma)]^2} - 3 \quad (2.31)$$

On the other hand for the  $t$  distribution, the kurtosis, here denoted by  $\gamma_t$ , is a function of the degrees of freedom  $v$

$$\gamma_t = \frac{6}{v-4}, \quad v > 4 \quad (2.32)$$

Equating 2.31 and 2.32 we can write the degrees of freedom  $v$  as a function of the kurtosis  $\gamma_2$  of the exponential power family of distributions.

$$v = \frac{6}{\gamma_2} + 4 \quad (2.32)$$

Equation 2.32 allow us to approximate a  $t$  distribution with  $v$  degrees of freedom, location parameter  $\mu$  and variance  $\sigma^2$ , by a member of the exponential power family of distributions with the same mean and variance and that has the same kurtosis. The following table shows the relation between the non-normality parameter  $\gamma$ , the kurtosis  $\gamma_2$  and the degrees of freedom  $v$ .

$\gamma$	0	0.1	0.2	0.3	0.4	0.5	0.7	0.9	1
$\gamma_2$	0	0.21	0.43	0.68	0.94	1.22	1.86	2.59	3
$v$	$\infty$	33	18	13	10	9	7	6	6

Table 2.6 Values of  $v$  and  $\gamma$  of distributions with the same kurtosis.

To compare the performance of both procedures in detecting changes in variability, random samples of 20 observations from a  $t$  distribution with  $\mu = 0$ ,  $\sigma = 1$  and degrees of freedom 33, 13, 9, 7 and 6, were generated and the standard deviation of the last ten observations was doubled. Cumulative sums were then computed assuming that the observations 1) were distributed according to a  $t_v$  and 2) were distributed as an exponential power distribution with parameter  $\gamma$  given by table 2.6.

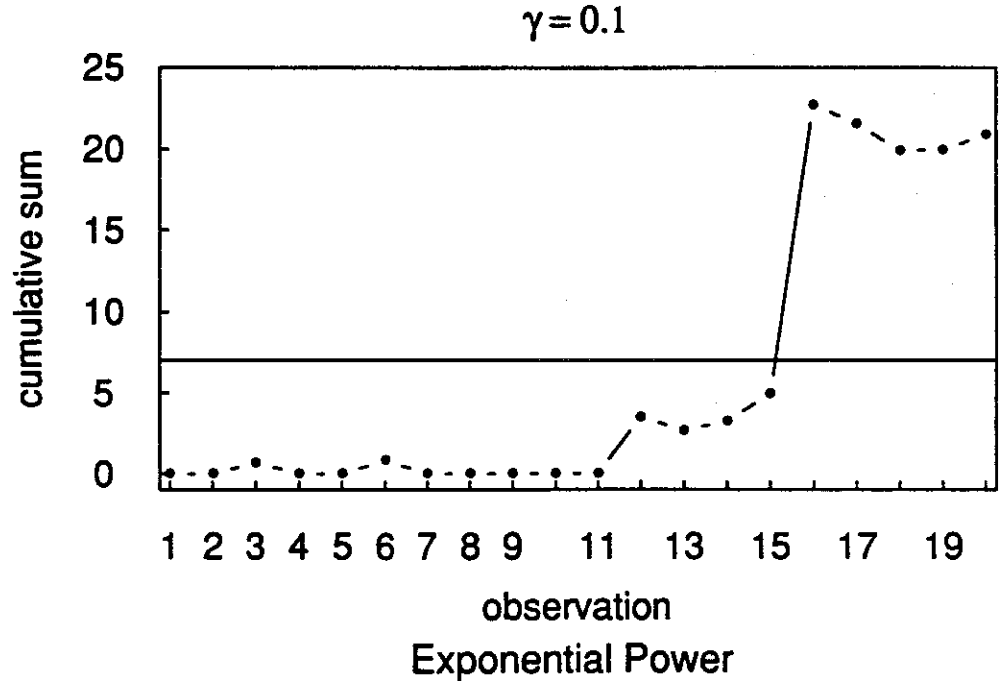
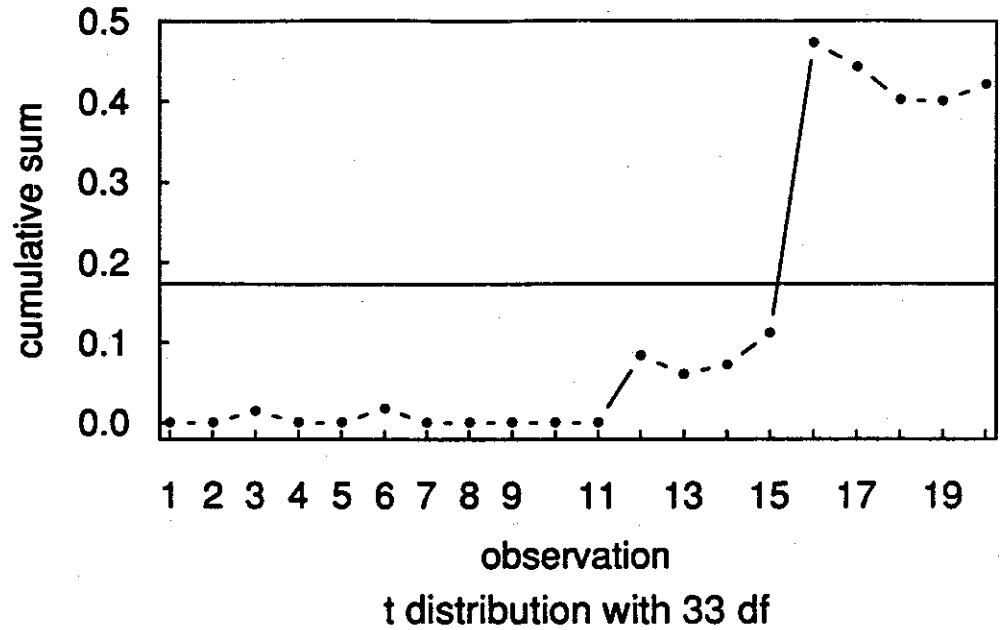
Cumulative sums for both distributions are shown in figures 2.10a to 2.10e, along with the decision interval  $h$ . For the  $t$  distribution  $h$  is calculated using equation 2.30, while for the exponential power distribution  $h$  is given by

$$h = \frac{\log\left(\frac{1-\beta}{\alpha}\right)}{c(\gamma) \left[ \frac{1}{\sigma_a^{1+\gamma}} - \frac{1}{\sigma_r^{1+\gamma}} \right]}$$

Where  $\alpha$  and  $\beta$ , the probabilities of type I and II errors, are set equal to 0.05.

From these figures we see that both CUSUM procedures have essentially the same behaviour; especially for values of  $\gamma < 0.7$ , and that they signal changes at approximately the same time. Therefore CUSUM charts based on the  $t$  distribution can be replaced by a CUSUM scheme using the exponential power family of distributions.

**Figure 2.10a** Cumulative sums for  $t$  distribution with 33 degrees of freedom and the corresponding exponential power distribution with parameter  $\gamma = 0.1$ .



**Figure 2.10b** Cumulative sums for  $t$  distribution with 13 degrees of freedom and the corresponding exponential power distribution with parameter  $\gamma = 0.3$ .

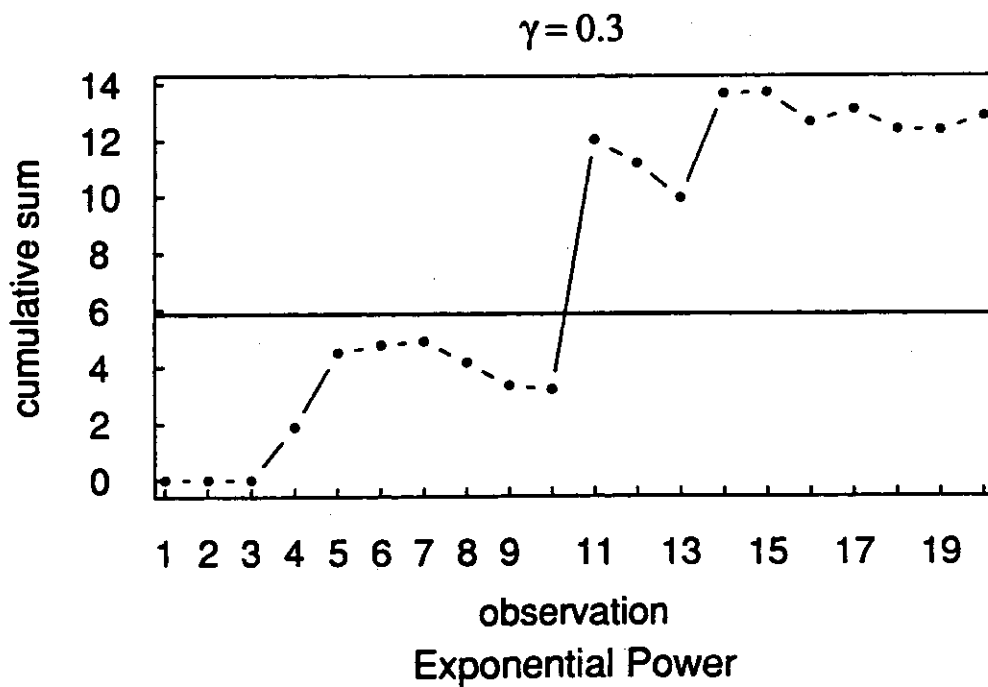
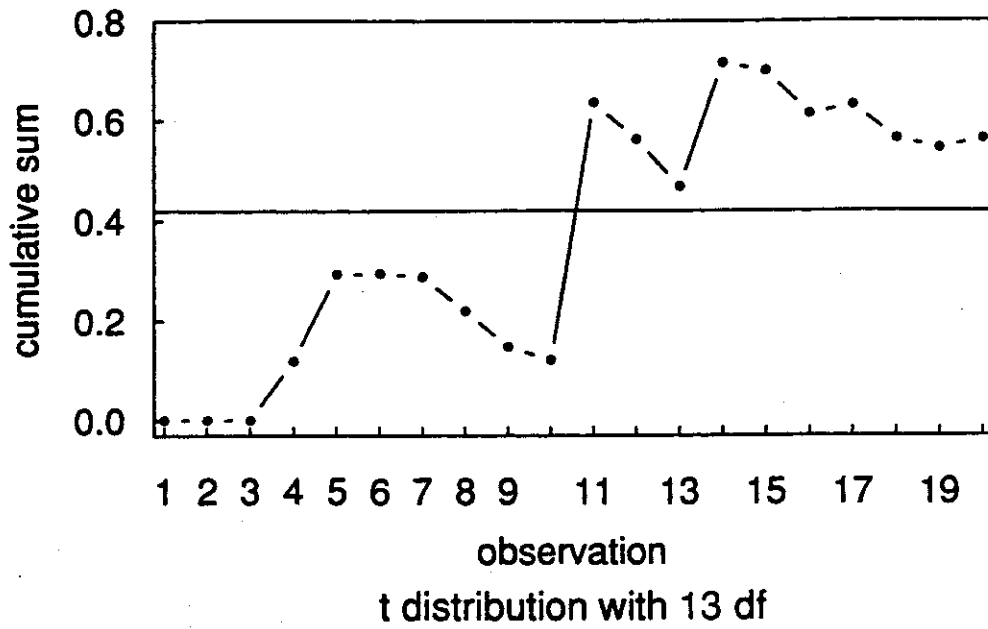


Figure 2.10c Cumulative sums for  $t$  distribution with 9 degrees of freedom and the corresponding exponential power distribution with parameter  $\gamma = 0.5$ .

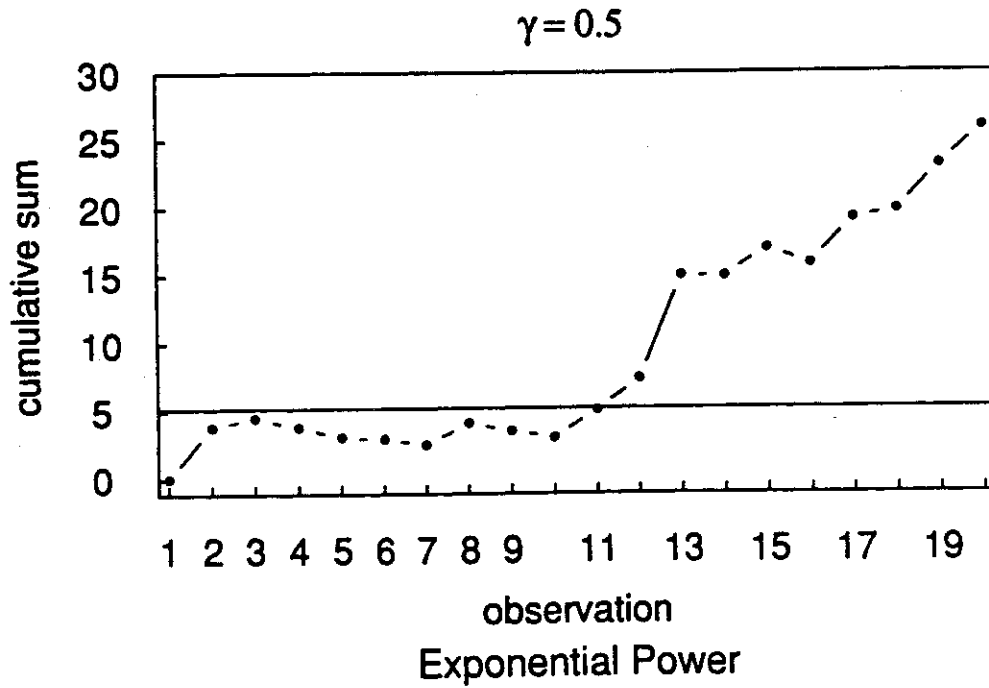
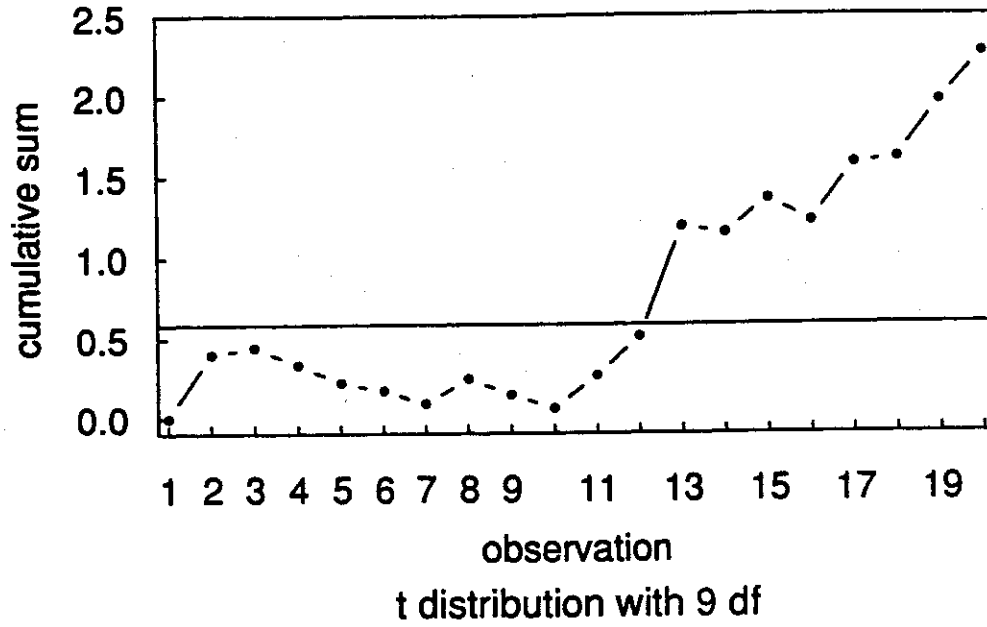


Figure 2.10d Cumulative sums for  $t$  distribution with 7 degrees of freedom and the corresponding exponential power distribution with parameter  $\gamma = 0.7$ .

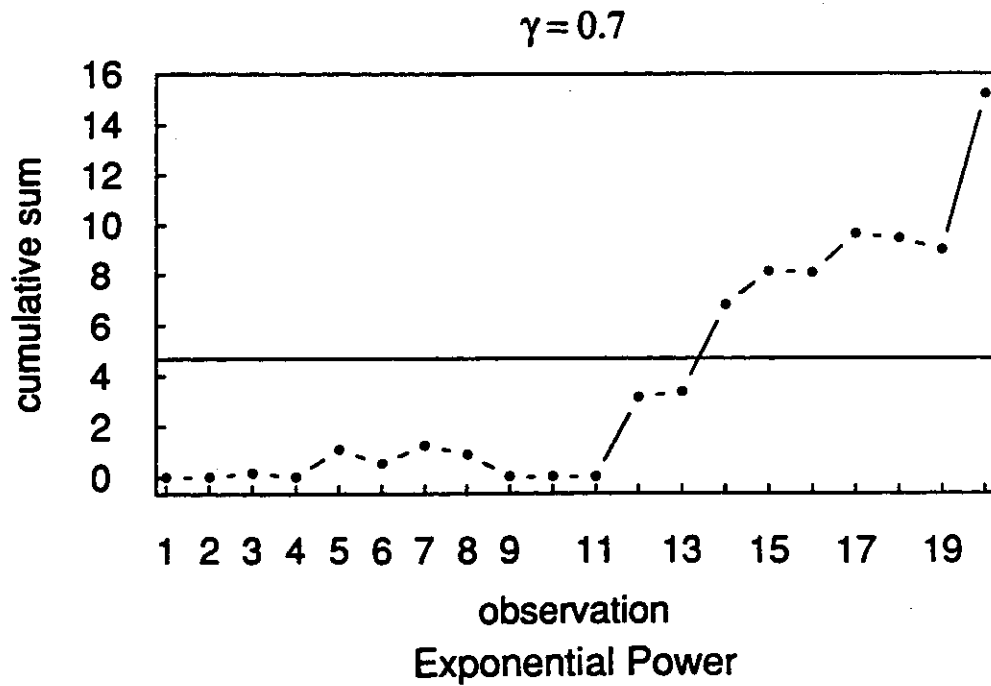
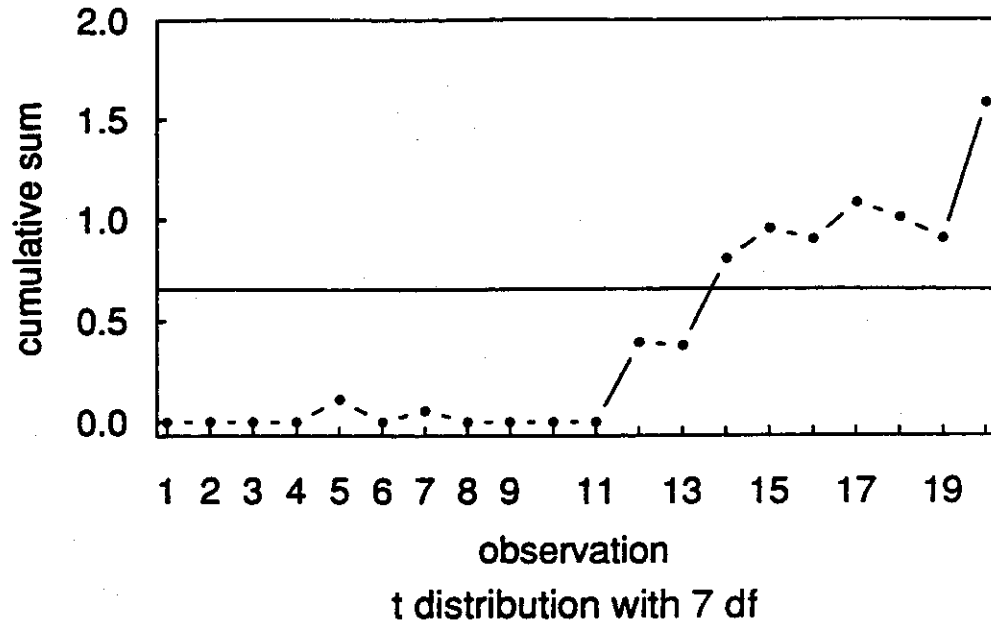
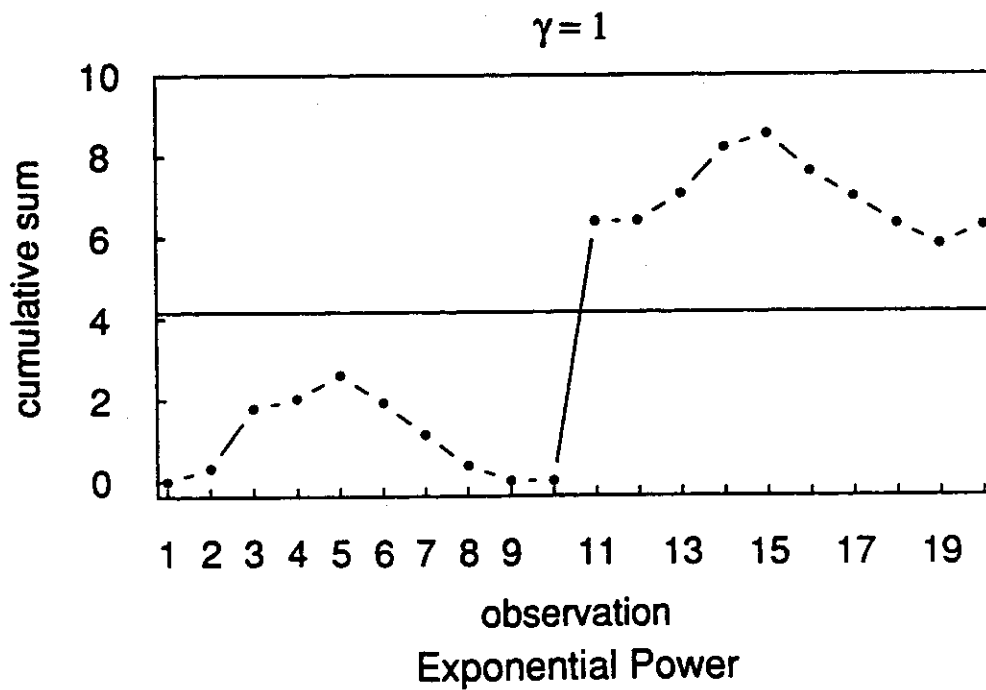
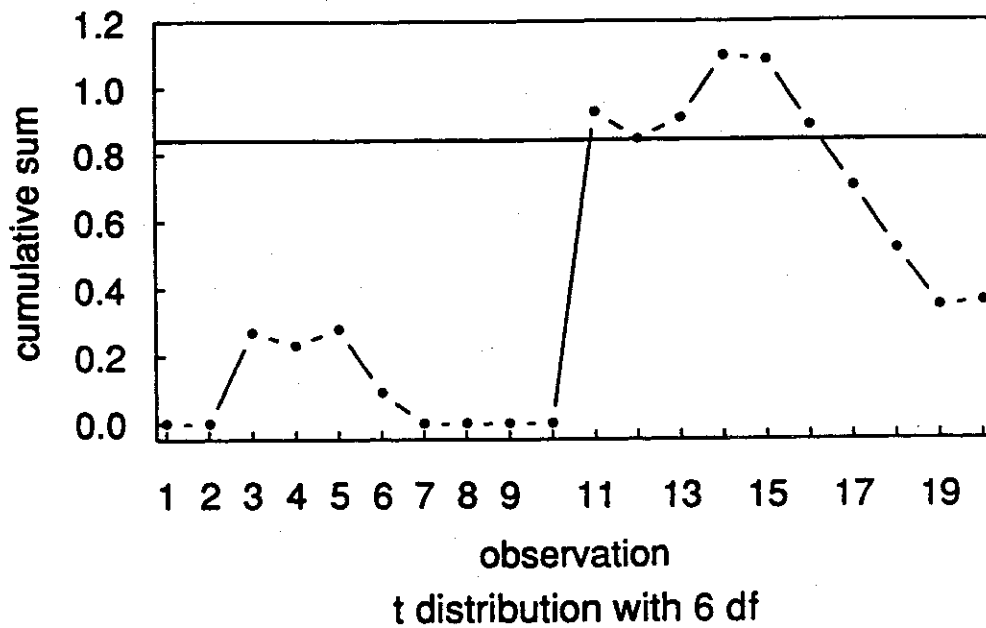


Figure 2.10e Cumulative sums for  $t$  distribution with 6 degrees of freedom and the corresponding exponential power distribution with parameter  $\gamma = 1$ .



## REFERENCES

- Akima, H. (1978), "A Method of Bivariate Interpolation and Smooth Surface Fitting for Irregularly Distributed Data Points," *ACM Transactions on Mathematical Software*, 4, 148-164.
- Bagshaw, M., and Johnson, R. A. (1975), "The Influence of Reference Values and Estimated Variance on the ARL of Cusum Tests," *Journal of the Royal Statistical Society B*, 37, 413-420.
- Bagshaw, M., and Johnson, R. A. (1977), "Sequential Procedures for Detecting Parameter Changes in a Time-Series Model," *Journal of the American Statistical Association*, 72, 593-597.
- Balasoorya, U., Tse, Y. K., and Liew, Y. S. (1987), "An Empirical Comparison of Some Statistics for Identifying Outliers and Influential Observations in Linear Regression Models," *Journal of Applied Statistics*, 14, 177-184.
- Barnard, G. A. (1946), "Sequential Tests in Industrial Statistics," *Supplement to the Journal of the Royal Statistical Society*, VIII, 1-21.
- Barnard, G. A. (1959), "Control Charts and Stochastic Processes," *Journal of the Royal Statistical Society B*, 21, 239-271.
- Barnard, J. (1978), "Probability Integral of the Normal Range," *Applied Statistics*, 27, 197-198. (Algorithm AS 126)
- Barnett, V., and Lewis, T. (1984), *Outliers in Statistical Data*, New York: NY: John Wiley & Sons.
- Bartlett, M. S., and Kendall, D. G. (1946), "The Statistical Analysis of Variance-Heterogeneity and the Logarithmic Transformation," *Journal of the Royal Statistical Society B*, 8, 128-150.
- Becker, R. A., and Chambers, J. M. (1984), *S An Interactive Environment for Data Analysis and Graphics*, Belmont, CA: Wadsworth.
- Bissell, A. F. (1969), "Cusum Techniques for Quality Control," *Applied Statistics*, 18, 1-30.
- Bissell, A. F. (1984), "The Performance of Control Charts and Cusums Under Linear Trend," *Applied Statistics*, 33, 145-151. (Corrigendum, Applied

Statistics 35.)

- Box, G. E. P., and Tiao, G. C. (1962), "A Further Look at Robustness Via Bayes' Theorem," *Biometrika*, 49, 419-432.
- Box, G. E. P., and Jenkins, G. M. (1966) "Models for Prediction and Control: VI. Diagnostic Checking." *Technical Report No. 99*, University of Wisconsin-Madison.
- Box, G. E. P., and Tiao, G. C. (1968), "A Bayesian Approach to Some Outlier Problem," *Biometrika*, 55, 119-129.
- Box, G. E. P., and Jenkins, G. M. (1970), *TIME SERIES ANALYSIS Forecasting and Control*, San Francisco: Holden-Day.
- Box, G. E. P., and Tiao, G. C. (1973), *Bayesian Inference in Statistical Analysis*, Reading, Massachusetts: Addison-Wesley.
- Box, G. E. P. (1980), "Sampling and Bayes' Inference in Scientific Modelling and Robustness," *Journal of the Royal Statistical Society A*, 143, 383-430.
- Brown, R. L., Durbin, J., and Evans, J. M. (1975), "Techniques for Testing the Constancy of Regression Relationships Over Time," *Journal of the Royal Statistical Society B*, 37, 149-192.
- Cho, S. (1984), "Robust Model-Free Prediction and Control". (Unpublished Ph. D. dissertation)
- Deming, W. E. (1986), *Out of the Crisis*, Cambridge, MA: MIT, Center for Advanced Engineering Study.
- Draper, N. R., and John, J. A. (1981), "Influential Observations and Outliers in Regression," *Technometrics*, 23, 21-26.
- Duncan, A. J. (1965), *Quality Control and Industrial Statistics*, Homewood, IL: Richard Irwin, INC..
- Ewan, W. D., and Kemp, K. W. (1960), "Sampling Inspection of Continuous Processes with No Autocorrelation Between Successive Samples," *Biometrika*, 47, 363-380.
- Ewan, W. D. (1963), "When and How to Use Control Charts," *Technometrics*, 5, 5-22.

- Galpin, J.S., and Hawkins, D.M. (1984), "The Use of Recursive Residuals in Checking Model Fit in Linear Regression," *The American Statistician*, 38, 94-105.
- Garbade, K. (1977), "Two Methods for Examining the Stability of Regression Coefficients," *Journal of the American Statistical Association*, 72, 54-63.
- Girshick, M. A. (1946), "Contributions to the Theory of Sequential Analysis. I," *Annals of Mathematical Statistics*, 17, 123-143.
- Gitlow, H., Gitlow, S., Oppenheim, A., and Oppenheim, R. (1989), *Tools and Methods for the Improvement of Quality*, Homewood, IL: Irwin.
- Goel, A.L., and Wu, S.M. (1971), "Determination of A.R.L. and a Contour Nomogram for Cusum Charts to Control Normal Mean," *Technometrics*, 13, 221-230.
- Goel, A.L. (1981), "Cumulative Sum Control Charts," *Encyclopedia of Statistical Sciences*, 2, 233-241, John Wiley. (eds. S. Kotz and N. L. Johnson)
- Goldsmith, P.L., and Whitfield, H. (1961), "Average Run Lengths in Cumulative Chart Quality Control Schemes," *Technometrics*, 3, 11-20.
- Grubbs, F.E. (1969), "Procedures for Detecting Outlying Observations in Samples," *Technometrics*, 11, 1-21.
- Hahn, G.J., and Cockrum, M. B. (1987), "Adapting Control Charts to Meet Practical Needs: A Chemical Processing Application," *Journal of Applied Statistics*, 14, 35-52.
- Hawkins, D.M. (1981), "A Cusum for a Scale Parameter," *Journal of Quality Technology*, 13, 228-231.
- IMSL (1979), *The IMSL Library*, Houston, TX: International Mathematics and Statistics Library.
- IMSL (1987), *The IMSL Library*, Houston, TX: International Mathematics and Statistics Library.
- John, J.A., and Draper, N. R. (1978), "On Testing for Two Outliers or One Outlier in Two-Way Tables," *Technometrics*, 20, 69-78.
- Johnson, N.L. (1961), "A Simple Theoretical Approach to Cumulative Sum Charts," *Journal of the American Statistical Association*, 56, 835-840.

- Johnson, N.L., and Leone, F.C. (1962), "Cumulative Sum Control Charts: Mathematical Principles Applied to Their Construction and Use," *Industrial Quality Control*, 18, june 15-21, july 29-36, august 22-28..
- Johnson, N.L., and Leone, F.C. (1964), *Statistics and Experimental Design In Engineering and the Physical Sciences*, New York: NY: John Wiley & Sons.
- Kantorovich, L. V., and Krylov, V.I. (1964), *Approximate Methods of Higher Analysis*, New York, NY: John Wiley & Sons.
- Kemp, K. W. (1958), "Average Sample Number of Some Sequential Tests," *Journal of the Royal Statistical Society B*, 20, 379-386.
- Kemp, K. W. (1961), "The Average Run Length of the Cumulative Sum Chart When a V-Mask Is Used," *Journal of the Royal Statistical Society B*, 23, 149-153.
- Khan, R. A. (1984), "On Cumulative Sum Procedures and the SPRT with Applications," *Journal of the Royal Statistical Society*, 46, 77-85.
- Lucas, J. M. (1973), "A Modified "V" Mask Control Scheme," *Technometrics*, 15, 833-847.
- Lucas, J. M. (1976), "The Design and Use of V-Mask Control Schemes," *Journal of Quality Technology*, 8, 1-12.
- Lucas, J. M., and Crosier, R. B. (1982), "Fast Initial Response for CUSUM Quality Control Schemes: Give Your CUSUM a Head Start," *Technometrics*, 24, 199-205.
- Lund, R. E. (1975), "Tables for An Approximate Test for Outliers in Linear Models," *Technometrics*, 17, 473-476.
- Marquardt, D. W. (1984), "New Technical and Educational Directions for Managing Product Quality," *The American Statistician*, 38, 8-14.
- Mason, R. L., Gunst, R. F., and Hess, J. L. (1989), *Statistical Design and Analysis of Experiments with Applications to Engineering and Science*, New York, NY: John Wiley & Sons..
- Page, E. S. (1954), "Continuous Inspection Schemes," *Biometrika*, 41, 100-114.
- Page, E. S. (1955), "Control Charts with Warning Limits," *Biometrika*, 42, 243-257.

- Page, E. S. (1957), "On Problems in Which a Change in a Parameter Occurs at an Unknown Point," *Biometrika*, 44, 248-252.
- Page, E. S. (1961), "Cumulative Sum Charts," *Technometrics*, 3, 1-9.
- Page, E. S. (1963), "Controlling the Standard Deviation by Cusums and Warning Limits," *Technometrics*, 5, 307-310.
- Ramírez, J. G., and Juan, J. (1989) "A Contour Nomogram for Designing Cusum Charts for Variance." Report # 33, Center for Quality and Productivity Improvement, University of Wisconsin-Madison.
- Rao, C. R. (1973), *Linear Statistical Inference and Its Applications*, New York, N.Y.: John Wiley & Sons.
- Reynolds, M. R. (1975), "Approximations to the Average Run Length in Cumulative Sum Control Charts," *Technometrics*, 17, 65-71.
- Schweder, T. (1976), "Some "Optimal" Methods to Detect Structural Shift or Outliers in Regression," *Journal of the American Statistical Association*, 71, 491-501.
- Shewhart, W. A. (1986), *Statistical Method from the Viewpoint of Quality Control*, Mineola, NY: Dover Publications. (first published in 1939)
- Snedecor, G. W., and Cochran, W. G. (1967), *Statistical Methods*, Ames: IA: The Iowa State University Press. (6th edition)
- Snedecor, G. W., and Cochran, W. G. (1980), *Statistical Methods*, Ames: IA: The Iowa State University Press. (7th edition)
- Truax, H. M. (1961), "Cumulative Sum Charts and Their Application in the Chemical Industry," *Industrial Quality Control*, 17, 18-25.
- Turner, M. E. (1960), "On Heuristic Estimation Methods," *Biometrics*, 16, 299-301.
- Van Dobben de Bruyn, C. S. (1968), *Cumulative Sum Tests: Theory and Practice*, New York, NY: Hafner Publishing Co.
- Wald, A. (1947), *Sequential Analysis*, New York, NY: John Wiley & Sons.
- Weisberg, S. (1985), *Applied Linear Regression*, New York: John Wiley & Sons.

Wheeler, D.J., and Chambers, D.S. (1986), *Understanding Statistical Process Control*, Knoxville, TN: Statistical Process Control, Inc..

Woodall, W.H. (1986), "The Design of CUSUM Quality Control Charts," *Journal of Quality Technology*, 18, 99-102.

Woodward, R. H., and Goldsmith, P.L. (1964), *Cumulative Sum Techniques*, London: Oliver & Boyd. (I.C.I. Monograph No. 3)