TECHNICAL REPORT

ISO/TR 7871

First edition
1997-02-15

Cumulative sum charts — Guidance on quality control and data analysis using CUSUM techniques

Cartes des sommes cumulées — Lignes directrices pour le contrôle de la qualité et l'analyse des données utilisant les procédures CUSUM

Reference number
ISO/TR 7871:1997(E)

Contents

International Organization for Standardization
Case postale 56 • CH-1211 Genève 20 • Switzerland
Internet central@iso.ch
X.400 c=ch; a=400net; p=iso; o=isocs; s=central

Printed in Switzerland

Foreword

ISO (the International Organisation for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental or non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The main task of technical committees is to prepare International Standards. In exceptional circumstances a technical committee may propose the publication of a Technical Report of one of the following types:

— type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;

— type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;

— type 3, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example).

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

ISO/TR 7871, which is a Technical Report of type 3, was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 4, *Statistical process control*.

0 Introduction

0.1 Basis of cusum chart

The cumulative sum chart (hereafter referred to by the generally accepted contraction "cusum chart") is a highly informative graphical presentation of data which are ordered in a logical sequence. Frequently this sequence corresponds to the order of observation on a time scale.

A reference value, T, is subtracted from each observation. This reference value is generally a constant but may be a prediction from a forecasting model or a target which may vary. The cumulative sums of the deviations from T are formed, and these cusums (C) are plotted against the serial numbers of the observations.

In a cusum chart intended to check a process for departure from a mean value equal to the reference value, that value is also known as the target value or aim. Without more advanced cusum procedures the two concepts, target value or aim and reference value must be distinguished. The former refers to the actual or intended process average, the latter to the reference values used in the cusum procedure. The intuitive appeal of the term target value is strong, however, and for most of this standard, clauses 0 - 6, the common value of target value and reference value is referred to as target value when this does not create ambiguity. In clause 6 upper and lower reference values are created and these must be distinguished from target values or aims.

The cusum method of plotting results is the representation of average by the local slope of the chart. When the local average corresponds to the target value, the path of the cusum lies roughly parallel to the sequence axis. When the local average of the series is greater than the target value, the cusum slopes upwards ; conversely, when the local average is less than the target value, the cusum slopes downwards. The greater the discrepancy between the local average and the target value, the steeper the slope of the cusum path.

The result of plotting the cusum is that changes in average level over different subdivisions of the total sequence of observations are clearly indicated by changes in slope of the chart. The local averages in such subdivisions can be readily estimated, either from the numerical values of the cusum from which the chart is plotted or directly, from the chart itself.

A second effect of using cumulative sum procedures is that there is an inherent serial dependence between the successive cumulative sums. Decisions regarding acceptable departures from the sequence axis require the use of the method of stochastic processes.

0.2 Simple example of cusum chart

The above principles are best appreciated from a simple example. The calculations and plotting procedure will, at this point, be developed without mathematical symbolism.

It is supposed that the following individual observations have been obtained, over a time sequence in order shown, and that a reference value of 15 is appropriate.

Table 1 : Data for cusum plotting

Observation number	Observed value	Deviation from reference value (= 15)	Cumulative sum of deviations
1	12	- 3	- 3
2	17	+ 2	- 1
3	14	- 1	- 2
4	14	- 1	- 3
5	17	+ 2	- 1
6	16	+ 1	0
7	14	- 1	- 1
8	11	- 4	- 5
9	13	- 2	- 7
10	14	- 1	- 8
11	15	0	- 8
12	11	- 4	- 12
13	14	- 1	- 13
14	16	+ 1	- 12
15	13	- 2	- 14
16	14	- 1	- 15
17	11	- 4	- 19
18	12	- 3	- 22
19	13	- 2	- 24
20	16	+ 1	- 23
21	12	- 3	- 26
22	18	+ 3	- 23
23	18	+ 3	- 20
24	17	+ 2	- 18
25	20	+ 5	- 13
26	15	0	- 13
27	14	- 1	- 14
28	18	+ 3	- 11
29	20	+ 5	- 6
30	16	+ 1	- 5
31	18	+ 3	- 2
32	14	- 1	- 3
33	16	+ 1	- 2

For a conventional control chart, as in figure 1, the observed values are plotted against their corresponding observation numbers. There is some indication that the last dozen values appear to be clustered around a different mean level from the first 20 or so.

Plotting in the cusum mode give a much clearer display than the conventional chart. The cusum (column 4 of table 1) is plotted against the observation number using the y ("vertical") axis for the cusum and the x ("horizontal") axis for the observation number, figure 2.

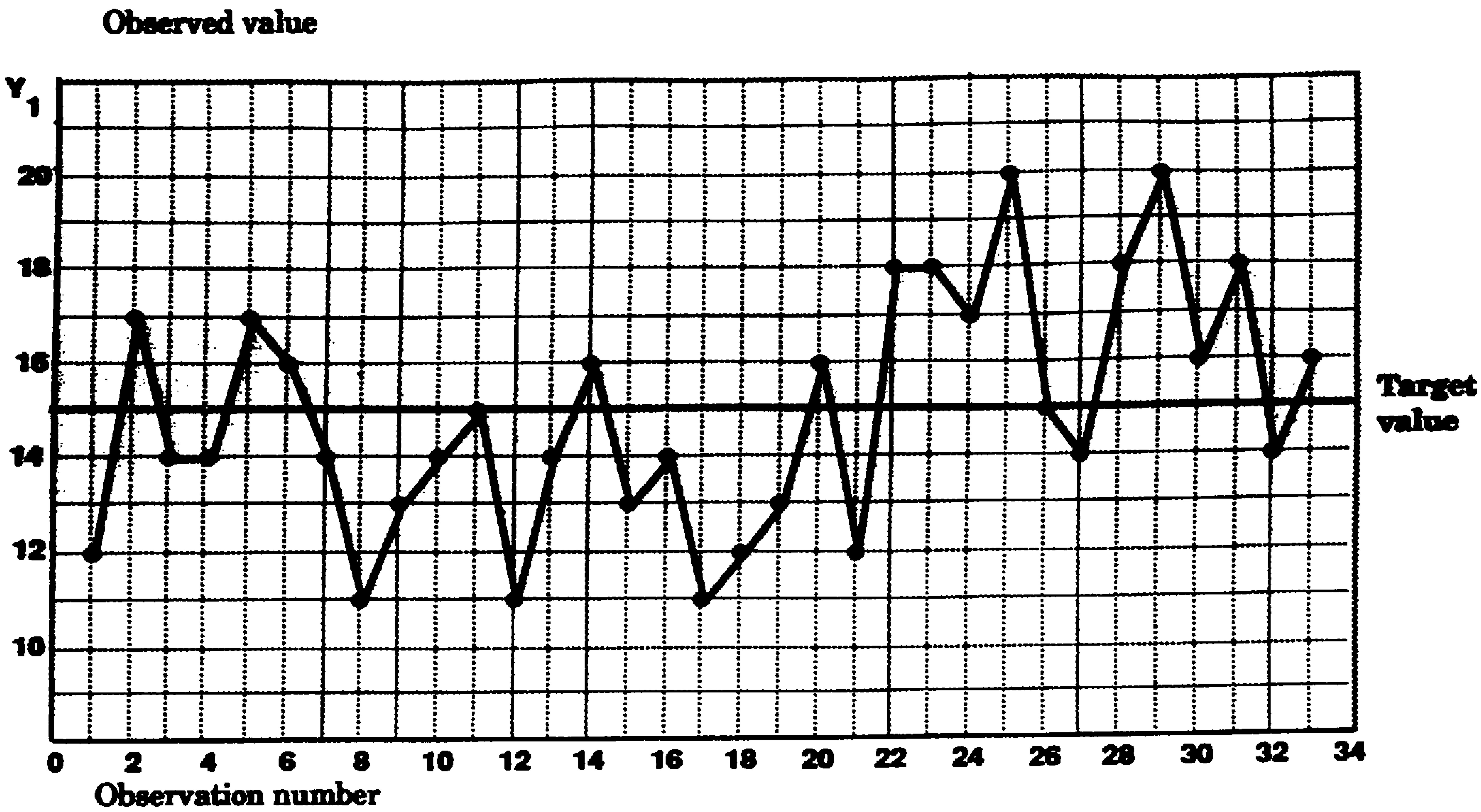


Figure 1 : Conventional chart of data from table 1

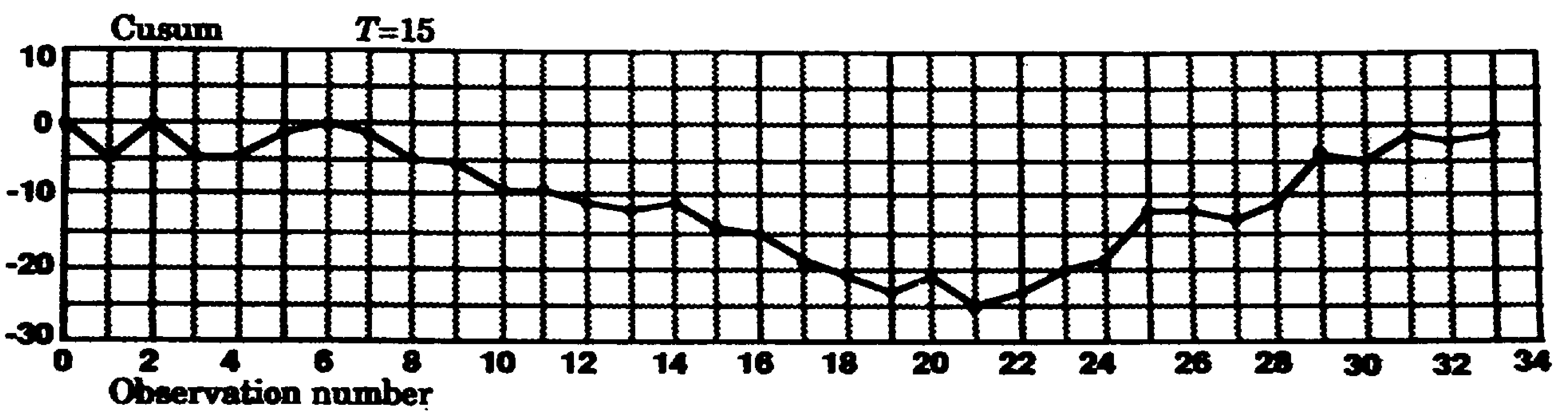


Figure 2 : Cusum chart of data from table 1

The cusum chart clearly separates into three segments. From observations numbered 1 to 7 (inclusive) the cusum path is generally parallel to the observation number axis, i.e. the path is roughly horizontal. From observations 8 to 21 inclusive the path is downward (despite local irregularities such as at observations 14, 20). From observations 22 to 33, the path is upward (again with local irregularities).

Thus it could tentatively be inferred that :

a) observations 1 to 7 constitute a sample from a "population" whose mean is at or near the target value (15) ;

b) observations 8 to 21 appear to have been sampled from a population whose mean is below 15 ;

c) observations 22 onward appear to come from a population whose mean is greater than 15.

There are now a number of questions that might be asked :

1) in the light of the underlying variability (as indicated, for example, by the irregularities in the cusum path) can it be concluded that the changes in slope represent real shifts in average rather than merely lucky or unlucky runs of samples from a stable population ?

2) if the changes are real, how should the data be used to estimate local averages ?

3) to what extent might the inferences or estimates be affected by the choice of the reference value or the cusum scale factor ? Thus figures 3 and 4 show the same series plotted first with the same cusum scale but with a target of 12 ; and second with a target of 15 but a compressed cusum scale.

In figures 3 and 4, the change in slope around observation number 8 is less apparent. The change around number 21 is still visible, but it is less easy to “pinpoint” in figure 4. Thus the choice of reference value and scale factor need careful attention, to avoid either the suppression of useful information or, conversely, the exaggeration of spurious effects. It is also clear from figure 3 that use of a an inappropriate target value may result in the chart running off the upper of lower edge of the graph paper, although this problem may also be minimized by replotting from a new zero at any point in the sequence.

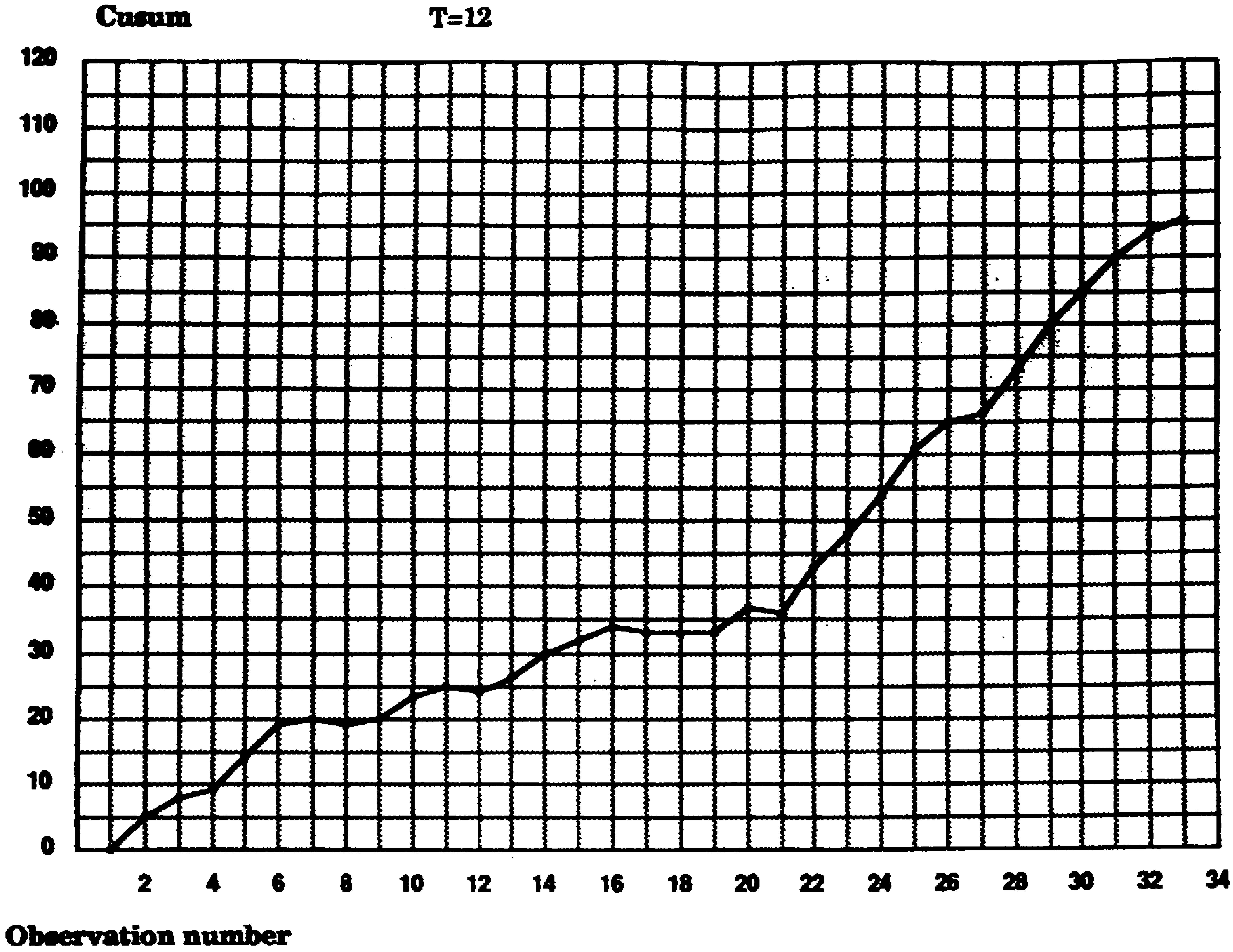


Figure 3 : Cusum chart of data from table 1, with reference value 12

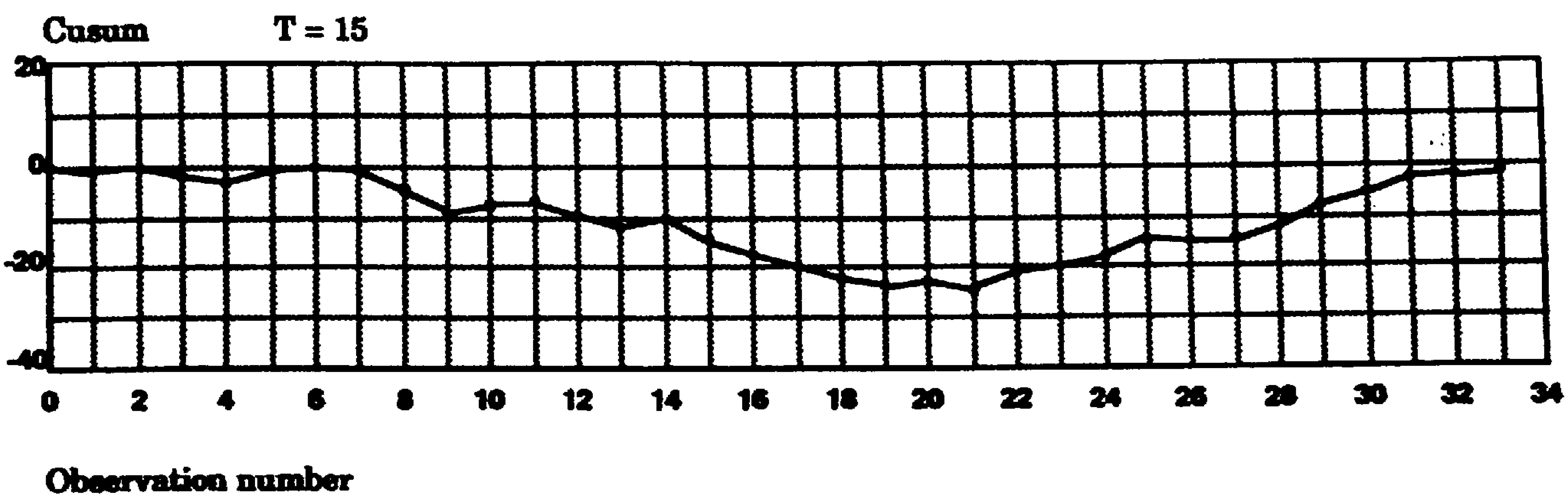


Figure 4 : Cusum chart of data from table 1, with reference value 15 but compressed cusum scale

Cumulative sum charts — Guide to quality control and data analysis using CUSUM techniques

1 Scope and general principles

1.1 General

This standard introduces the principles of cusum charting and includes guidance on the preparation and interpretation of cusum charts using basic decision rules.

1.2 Fundamental requirements

The fundamental requirements from cusum charting are as follows :

a) the observations should be at least on an interval scale of measurement ;

b) there should be logical grounds for the sequence for plotting. This arises naturally in process control.

These requirements are taken in order. The interval property requires any given numerical difference between two observations to have the same interpretation throughout the range of the variable. Thus a difference of 0,1 mm between the lengths of two objects has the same meaning whether the objects are woodscrews of length 10,1 mm and 10,0 mm or steel girders of length 10 000,1 mm and 10 000,0 mm although the latter difference may be unimportant. Many arbitrary scales do not have this property : ratings are an example, where perhaps a serious nonconformity scores 10 points, a moderate nonconformity 5 and minor nonconformity 1. We cannot then interpret this to mean that the following items are necessarily equally undesirable, although their score differences are zero :

- item A one serious nonconformity	Score = 10 ;
- item B two moderate nonconformities	Score = 10 ;
- item C one moderate, five minor nonconformities	Score = 10 ;
- item D ten minor nonconformities	Score = 10.

Interpretation of "average" score could be misleading if the balance of serious, moderate and minor nonconformities, rather than merely their overall frequency, changes.

The logical sequence property may arise in numerous ways. The observations may occur in a time or length sequence, thus forming a natural progression. Monitoring for quality or process control provides many cases of this kind.

Observations may be ordered according to the value of some auxiliary variable measured on the items. The cusum then provides a means of presenting or investigating relationships between variables, or augmenting a regression or correlation analysis. Any kind of ordering or grouping that uses some structural feature of the observations or the background from which they are taken may provide the basis for the cusum sequences.

1.3 Types of data amenable to cusum charting

Many types of data satisfy the fundamental requirements a) and b) of 1.2 ; Perhaps the most frequent applications of cusum charts have been in quality control, where observations such as sample means or ranges are plotted in sequence to assess the state of a process. When using a cusum chart as a device for effective data presentation, it is not necessary to specify a distribution, nor to require independence between successive observations. These conditions are important for decision rules, but not for data presentation. Indeed, the cusum chart may assist in the identification of distributional features such as serial correlation or cyclic behavior.

Thus data involving ranges or sample estimates of standard deviation may be plotted on cusum charts, as well as sample averages. Counts of nonconformities are also encountered in quality control, and may be monitored by cusum charts.

1.4 Monitoring or retrospective analysis

In prescribing decision rules and statistical tests, two distinct situations should be recognized :

a) the object of charting may be to monitor the behavior of a series of observations against some specified or standard reference value such as in quality control operations. Decision rules for monitoring are presented in clause 4.

b) the object may be to examine historical data, or observations grouped in some logical manner, so as to detect any differences between segments. No formal standard or reference value exists. This situation is close to that of testing the significance of apparent differences between groups of observations, but it differs in that the grouping may be effected on the basis of a preliminary inspection of the cusum chart. Statistical tests for retrospective analysis are presented in clause 5.

2 Preparations for cusum charting

2.1 Notation

A reference value will be denoted by T. As each observation in the sequence is encountered, the difference $(Y_r - T)$ is formed. These differences are summed, so that by the time Y_i, is reached, the cusum is formed :

$$C_i = \sum_{r=1}^{i}(y_r - T) \qquad (1)$$

The cusum, C_i, is plotted as ordinate ("vertical" axis) against the abscissa ("horizontal" axis). Assuming that i takes successive integer values 0 (at the origin), 1, 2, ..., the scale factor on the vertical axis will be denoted by A. This is interpreted as meaning that the distance which, on the horizontal scale, corresponds to one plotting interval, represents A on the cusum scale. This scale factor may often be expressed as mutiples of the standard error of the plotted values (σ_e), and this standardized scale factor will carry the a (thus $A = \alpha\sigma_e$). The meaning and estimation of σ_e is detailed in appendix A.

It will frequently be useful to calculate a local average for the sequence of points from i to j, or from i to j - 1, or perhaps from j to j + r, etc. These will be indicated by :

$\bar{y}_{i,j}$; $\bar{y}_{i,j-1}$; $\bar{y}_{j,j+r}$; etc.

Other notation will be defined as it is introduced.

2.2 Choice of reference value (T)

Choice of a suitable target value is one of the two most important steps in the preparations. An unsuitable target value will cause the cusum to slope persistently up or down, making changes more difficult to observe, and necessitating frequent replotting when the chart runs off the top or bottom of the graph (see Appendix D).

2.3 In many cases, T is a specified target or aim level of the quality measure. It is wise to have assurance that the process can produce this quality, otherwise the cusum chart will merely be a persistent reminder of failure, and in such circumstances any control system tends to be ignored or fall into disuse.

There may not always be an aim value for T. Sometimes the mean level of the quality measure over a recent stable series of data may be used.

2.4 Where a cusum chart is to be used for retrospective examination of a series of historical data, or residuals form an experiment, the natural target becomes the arithmetic mean of the complete series. Apart from any minor discrepancy arising from rounding of the mean, this choice for T will result in the cusum being and ending at the same ordinate value. Combined with suitable scaling (see appendix B) the cusum plot may be contained within the limits of the graph grid used for plotting.

2.5 Binary data, coded as a sequence of 0's and 1's, require an appropriate estimate of the proportion of responses which are scored as 1's. This proportion is then used as a target value. For quality control applications, a proportion of nonconformities or Acceptable Quality Level (AQL) may be specified by contract, or a feasibility study may provide a suitable value. In experiments involving binary response, either the overall proportion for the complete experiment may be used or, in the case of sequential experiments, the target may be set at the proportion observed in the first complete segment of the experiment, modifying this later if it proves unsuitable.

2.6 Types of variation

In order to scale the chart effectively, and also to provide a basis for significance tests, a measure of the underlying short-term variation in the series is required. In engineering terms, the noise should be measured in order to scale the system for detection of signals.

The fundamental statistical measure of variation is the standard deviation. It may be estimated from a sample of n values by :

$$s(\text{estimate of } \sigma) = \sqrt{\frac{1}{n-1}\sum(y_i - \bar{y})^2} \qquad (2)$$

where :

$$\bar{y} = \frac{1}{n}\Sigma y_i$$

The summation extending over the n observation in the sample.

Frequently the values to be plotted are some function of a group of observations, a statistic such as their mean, range, proportion nonconforming, etc. The appropriate measure of variation then becomes the standard error of the plotted sample statistic (or an estimate thereof if this is unknown). The simplest cases are those of the sample mean, $\bar{x}$, or proportion non conforming, p, in a sample of n items. Not that n is the size of each of the samples, that is, the number of observations per sample, not the number of samples to be plotted.

In these two simple cases, assuming an in-control process :

Standard error of $\bar{y}$, $\sigma_e = \sigma / \sqrt{n}$, often being replaced by its estimate.

When the observations are counts or proportions of items with a specified attribute in samples of size n, if the process is under statistical control with p_0 the probability an item has the attribute, then the binomial distribution is appropriate giving :

$$\text{Standard error of } p,\ \sigma_p = \sqrt{\frac{p_0(1-p_0)}{n}} \qquad (3)$$

$$\text{often written as, } \sigma_p = \sqrt{\frac{p_0 q_0}{n}},\ q_0 = 1 - p_0$$

The standard deviation of number of items per sample having the required attribute is :

$$\sigma = \sqrt{n p_0 q_0} \qquad (4)$$

This is one instance where the standard deviation itself may be useful for diagnostic purposes in connection with a plot of raw sample data. Another is the number of nonconformities, faults, defects, or other occurrences in some quantity of product or material, or observed in some time segment. Examples are faults per meter (or square meter) of cloth, and accidents per week in a factory. Here, if the conditions appropriate to the Poisson distribution are assumed :

$$\sigma = \sqrt{m}$$

where m is the average level of occurrences per sample. In particular, if the probability of a nonconformity in a very small volume of product is very small and is proportional to the volume of the product, and if the nonconformities occur independently of one another, then a mathematical consequence is that the nonconformities can be expected to follow the Poisson distribution.

The requirement of "statistical control" implies that, during any period when no "assignable" change, or cause of variation, occurs, all the items sampled may be regarded as simple random samples from the whole process (or population, or time segment, etc). In this case, the short-term variation as observed between items within samples forms a suitable basis (via the standard error of the chosen summary statistic) for estimating the expected variation in the sequence as a whole. Any variation greater than this is assumed to arise from assignable causes, indicating a shift in the mean of the series or a change in the nature or magnitude of the variation.

There are many cases where the simple use of overall estimates of the standard deviation is inappropriate. Some circumstances resulting in its breakdown are as follows :

a) In making observations on a continuous process, there may be small but unimportant variations in the average level ; it is against these variations, rather than the extremely short-term variation, that systematic or sustained changes should be judged. As examples, an industrial process may be controlled by a thermostat or other automatic control device ; quality of raw material input may be subject to minor variations although never violating a specification. In monitoring a patient's response to treatment, there may be minor metabolic changes connected with meals, hospital or domestic routine, etc., but any effect of treatment should be judged against the overall typical variation ;

b) The method of sampling may itself induce effects like those in a). Often samples comprise items taken close together from a production line, on the grounds that a true random sample of all items manufactured is inconvenient. The items then constitute a "cluster" sample, and may tend to be too similar to each other to form a basis for assessing overall variation ;

c) Samples may comprise output or observations from several sources (machines, operators, administrative areas). As such, there may be too much local variation to provide a realistic basis for assessing whether meaningful changes have occurred.

Because of this, data arising from a combination of such sources should be treated with caution as any local peculiarities within each contributing source may be overlooked ; moreover, variation between the sources may mask any changes occurring over the whole system as time progresses.

d) Serial correlation may be present in the observations, that is, one observation is correlated with others nearby. For example, if moving averages are used, the overlap between the data values used in one such average and the next produces a positive serial correlation. In estimating use of a bulk material from differences between successive gauge or dipstick readings, an overestimate on one occasion will tend to produce an underestimate on the next, giving negative correlation. The possible presence of one or other of these effects needs to be recognized. Positive serial correlation is especially likely in some industrial processes where one batch of material may partially mix with preceding and succeeding batches producing what is sometimes termed as "heel" effect. Successive additions of fuel to the tank of a vehicle is an everyday example, each new addition being made before the tank is exhausted.

It is thus necessary to consider other measures of variation in the series or sequences of data, and the circumstances to which they are appropriate. Such measures of variation include treating the differences between successive sample values $\left(\delta_j = y_j - y_{j-1}\right)$ as the appropriate type of variation and treating all the sample values, y_j, as though they were drawn from a single population. These measures are discussed in appendix A.

2.7 Measures of variation

See appendix A.

2.8 Scaling the chart

See appendix B.

2.9 Check list of cusum preliminaries

As a reminder of (but not a substitute for) the detailed description of preliminary steps set out in this chapter, the following check list may be useful.

Choose an appropriate target value. The possibilities include :

a) a specification value ;

b) a satisfactory level of performance (for a process) which has a reasonable chance of being achieved ;

c) an average level of performance over a recent and typical period or segment ;

d) the average level of a complete set of observations, where retrospective analysis is involved.

Select a suitable measure of variation, taking the points in 2.7 into account.

Decide on the scaling convention to be adopted. The method of B.1 is generally the simplest for both preparation and interpretation, but for some special purposes one of the other two conventions may be preferred or special forms may be prepared for routine use.

Ensure that staff involved in the preparation or interpretation of the charts are familiar with the procedure.

3 Presentation

3.1 In clause 2 preparations for plotting the chart were detailed, including selection of the target value, defining and calculating a measure of variation, and the choice of a scaling factor.

Some practical points of labelling also deserve attention. The minimum information presented on the chart should include the following :

a) Target value (which may be accompanied by a brief indication of the reason for its selection, e.g. specified mean value, mean of past data) ;

b) Standard error of the observations (which may be accompanied by a note on the method used to estimate it) ;

c) Nature of the observations (original values, sample means, nonconformity counts, etc.) ;

d) Title indicating the purpose of the chart (e.g. "Cusum chart for control of .....", or "Retrospective cusum for data from .....") ;

e) Clear labelling of the i-scale (sample intervals) and cusum scale.

3.2 To assist in interpreting observed cusum chart pattern changes in conditions that are known to have occurred can be noted at the appropriate point on the i-scale. Examples are new deliveries of raw materials for the manufacturing process, or changes of personnel or of methods of operation.

In examining residual errors from an experiment, the points where changes in levels of the experimental factors occurred may be noted. For data collected over time, occasional date marks should be highlighted on the i-scale.

The likely inclusion of information of this kind may affect the choice of sample interval scale ; if appreciable annotation is envisaged, a more generous scale may be preferred than would otherwise be the case, to avoid a chart excessively cluttered with auxiliary information.

3.3 Choice of chart origin

For most applications, the origin for plotting the cusum scale will be zero, with provision for positive and negative cumulative sums to be plotted. However, the subsequent visual interpretation of the chart is not affected by the actual origin adopted, and if non-negative values are preferred, the cusum may be commenced at a suitable positive value. For example, when a chart is a continuation of some previous plot, or when the cusum path runs off the top or bottom of the graph, it may be useful to commence the new chart (or replotted segment) at a value corresponding to the general level near the conclusion of the previous chart (or segment).

3.4 Calculation of local averages

See appendix C.

3.5 Replotting the cusum chart

See appendix D.

4 Decision rules for monitoring and control

4.1 Introduction

4.1.1 In this clause and in clause 5, simple decision rules are presented for determining whether apparent changes in cusum slope (and hence of the average level of the plotted variable) are real, or whether they may be regarded as part of the underlying variability in the data.

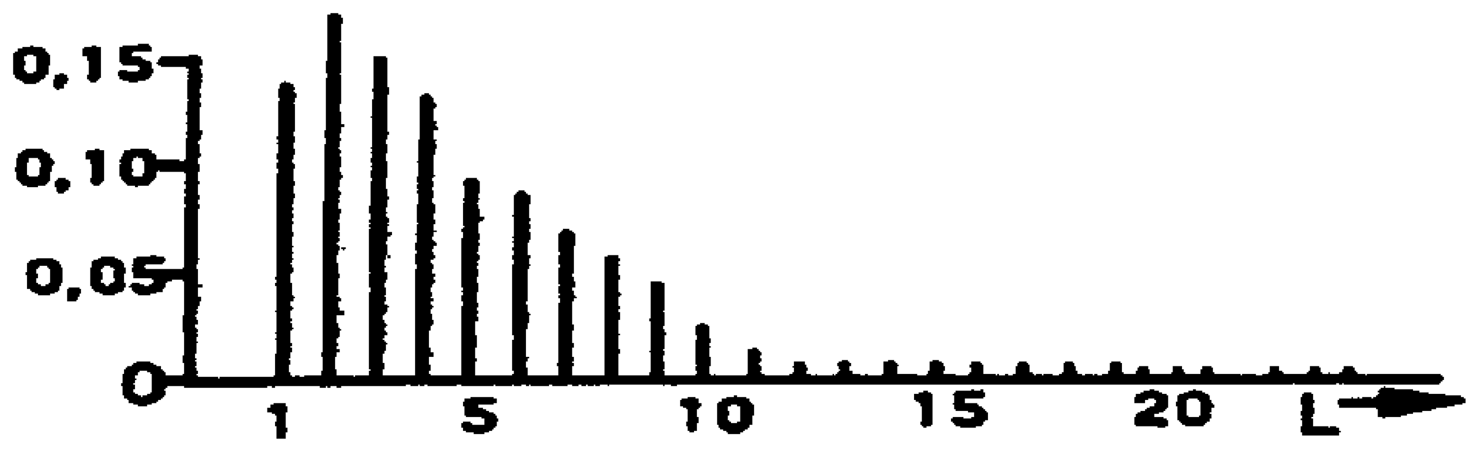
The present clause is concerned with monitoring and control applications : the plotting of data points in sequence as the observations are obtained, to detect as rapidly as possible any shift from a target level. Clause 5 deals with decision rules for the retrospective analysis of existing data to establish whether particular segments differ in average level from the preceding or succeeding segments.

4.1.2 The use of decision rules may not always be necessary, especially in cases where cusum charts are used only as a means of effective presentation of series data. However, where decisions and possibly resulting action need to be taken, the cusum method often yields more efficient use of the available data than other traditional techniques. In this context, efficiency implies either that decisions about changes in level are reached more quickly ; or that false alarms (erroneous decisions that a shift in mean has occurred when in fact there has been no change) are less frequent ; or that both advantages may be obtained together.

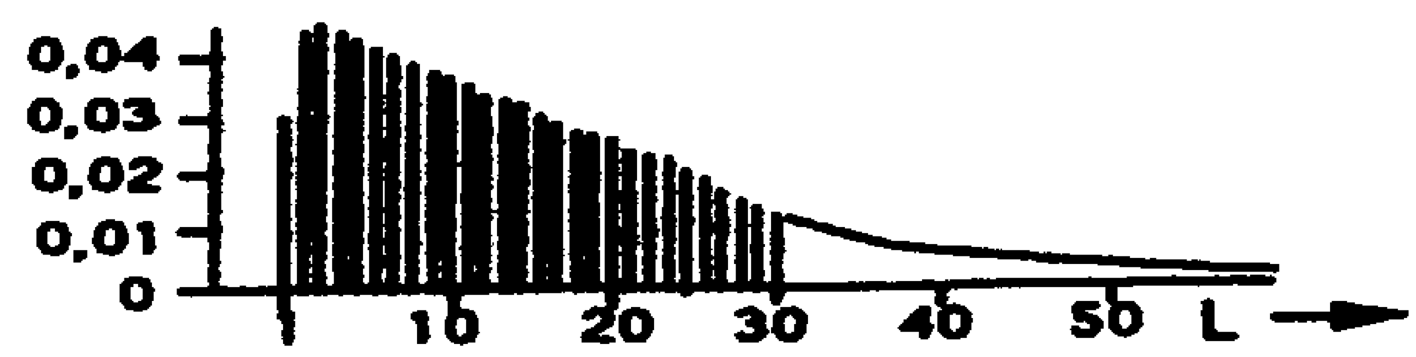
4.1.3 Cusum procedures when used for monitoring and control have the dual objectives of detecting significant changes and the points at which they occur. In order to measure the performance of such procedures, it is relevant to consider the average number of samples taken until an apparent change is signalled by the chosen decision rule. When no real change has occurred, this average run length should be long -perhaps hundreds of samples - as any signals will then be false alarms. When a large shift from target occurs, a short average run length is desirable so as to detect off-standard conditions rapidly. The response of a decision rule to various magnitudes of shift from target can be represented by an average run length curve (ARL curve), which provides a means of comparing decision rules in a manner analogous to that of the operating characteristic (OC curve) for comparing acceptance sampling procedures, or the power curve for comparing statistical hypothesis tests.

It should be noted that the run length itself is subject to statistical variation. Sometimes one may be "lucky" in obtaining no false alarms over a long run, or detecting a change very quickly. Sometimes an "unfortunate" run of samples may generate false alarms, or mask a real change so that it does not yield a signal. The average run length (ARL) is used as a summary measure for comparison of schemes and decision rules, but other aspects of run length (mode or median run length, or the complete run length distribution) may sometimes deserve attention. Typical run length distributions are illustrated in Figure 5.

a) ARL = 5
σ_L = 4
Median = 3,5
Mode = 2
2,5 % of runs ≥ 16
1 % of runs ≥ 19, etc.



b) ARL = 20
σ_L = 18,9
Median = 14
Mode = 3
12,5 % of runs ≥ 41
7,5 % of runs ≥ 50, etc.



c) ARL = 96,5
σ_L = 95,6
Median = 67
Mode = 4
12,5 % of runs ≥ 200
7,5 % of runs ≥ 250, etc.

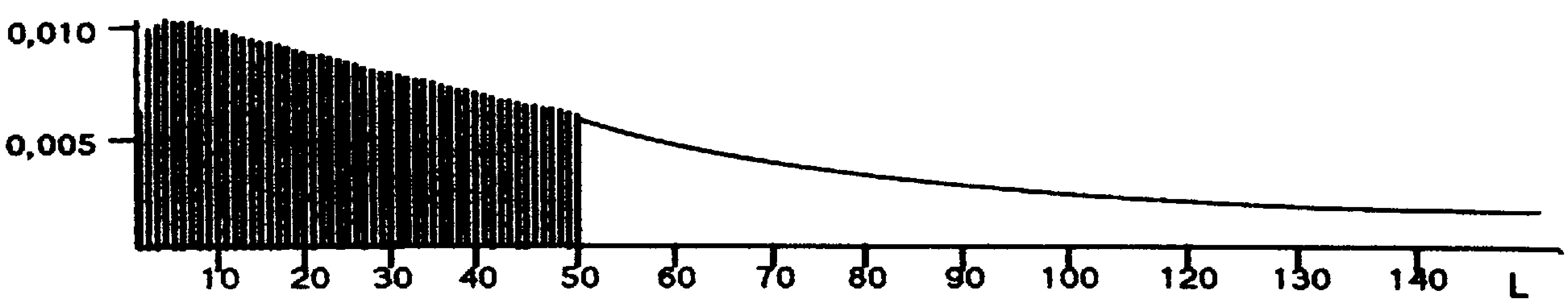


Figure 5 : Distributions of characteristic run lengthes

4.1.4 The essential features of the cusum chart relevant to the detection of changes in the underlying average level are the steepness of the cusum slope and the number of samples over which the slope persists. For monitoring and control, the object is to detect shifts from a target level. When the process is running at the target level, because of sampling variation, inevitably there will be occasions when the cusum path apparently diverges from the horizontal.The decision rule should not give false alarms that indicate that such sequences represent significant departure from target. On the other hand, when the process moves to an unsatisfactory condition, the decision rule should give as rapid a response as possible. The desiderata are thus :

True process condition	Cusum response
At or near target	Long ARL (few false alarms)
Substantial departure from target	Short ARL (Rapid detection)

The detailed specification of cusum procedures to give prescribed ARLs at particular quality levels is deferred to other standards. Here are presented simple rules that cover a wide range of applications. Their run length properties are related to quality levels generally defined as departures from target, measured in standard errors of the plotted variable. The standard error is assumed to be known from some preliminary evaluation, as described in Appendix A of this standard.

4.2 CUSUM V-masks

4.2.1 General

The simplest decision rules for use in conjunction with cusum charts are embodied in V-masks. There are three slightly differing forms, but all are identical in principle and effect. Each form may be applied with any of the three scaling conventions ; they will be applied to the more generally used version, as detailed in B.1 of Appendix B of this standard.

If the cusums are implemented by computing, the decision rules are also identical in principle and effect, but are different in form and are described in clause 6.

4.2.2 The truncated V-mask

Although a complete V-mask may be used, the truncated form is convenient in practice. The mask geometry is referenced to a point, indicated by A in figure 6, which may be identified by a small notch in the vertical section BC. Two sloping arms run from B to D and C to E, but they may be extended indefinitely beyond D or E if required. To correspond with later sections of this standard the vertical half-distances AB, AC will be termed the decision interval, and the lines BD, CE as decision lines. The slope of these decision lines, measured in scale units per sample plotting interval, will correspond to the reference values introduced in 6.1.3.

The construction of the mask is detailed in figure 6. Note that the distances $5\sigma_e$, $10\sigma_e$ may be transferred from the cusum scale if desired, or calculated from σ_e and the scale factor, if preferred.

The mask is used by placing the geometric reference point A, over any plotted point on the chart : this will often be the most recent point plotted, or the last point in some segment of particular interest. The AF axis is laid parallel to the sample number axis of the chart. If any preceding cusum point is outside the sloping arms (or their extensions beyond D or E), a significant departure from the target value is signalled. However, if the entire cusum path remains inside the arms, no significant shift is indicated.

Figures 7 and 8 show the truncated mask applied at two points on the cusum chart of the data from 0.2 (table 1 and figure 2). The value of σ_e is assumed to be 2,0. With the mask applied at observation 16, the cusum remains within the sloping arms, and the segment from samples 8 to 16 does not differ significantly from the target value of 15. However, with the mask applied at sample 18, the cusum path is seen to touch the upper decision line, indicating that, by the time the 18th value is obtained, sufficient evidence has accumulated to signal a downward shift from the target value.

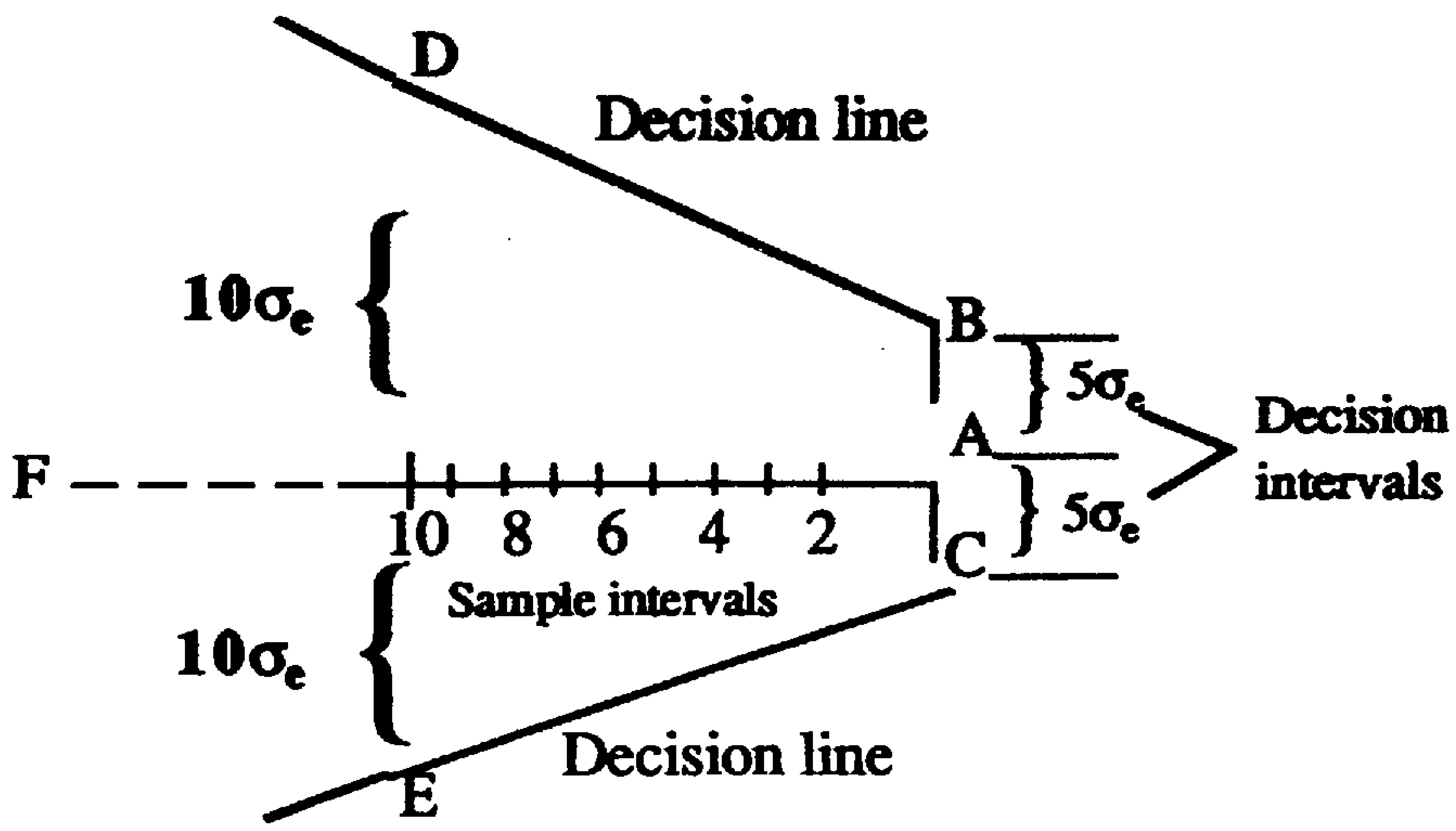


Figure 6 - General-purpose truncated V-mask

Several points are worth noting here :

a) the upper decision line is touched by sample points 6 and 7. If the cusum touches or crosses the decision line it is deemed to be a violation, yielding a signal ;

b) the point at which the decision line is violated requires extending the arm of the mask, but the inference of a change in mean is valid ;

c) if the mask is applied at any point before 18, no decision is signalled. If it is applied at any point between 18 and 21 (the end of the downward segment) a significant downward shift is indicated. Repeated (even unnecessary) use of the mask does not invalidate or modify the signals already obtained ;

d) the signal occurs at sample 18. At this point, the evidence is sufficient to support a conclusion that the mean has changed, but the change is likely to have set in at some earlier stage. Inspection of the chart suggests samples 7 to 8 as the first of a sequence in which the average is displaced from the target value ;

e) similar use of the V-mask later in this series would reveal a violation of the lower arm, indicating an upward shift by the 29th sample point.

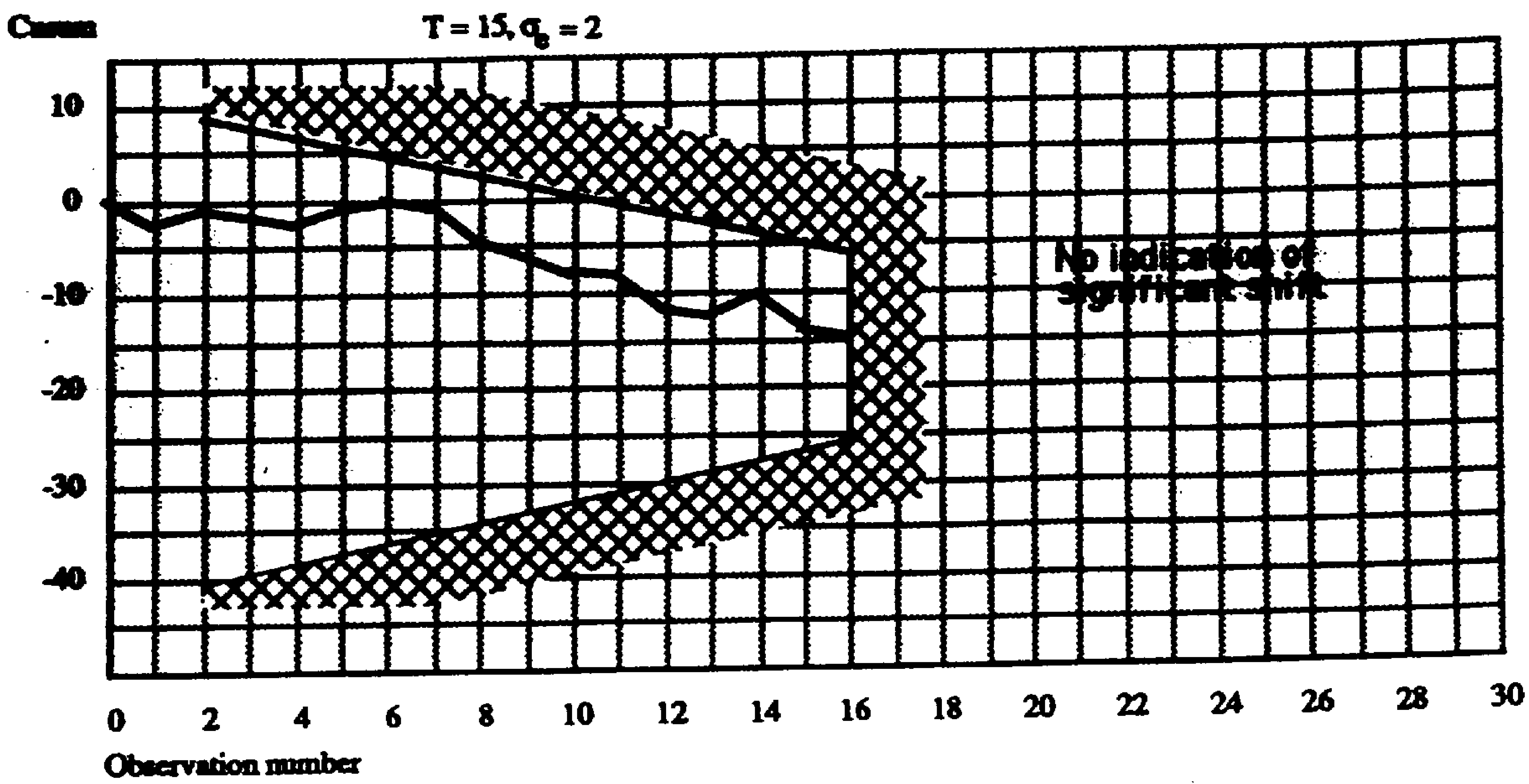


Figure 7 : Truncated V-mask applied to cusum chart : no indication of shift

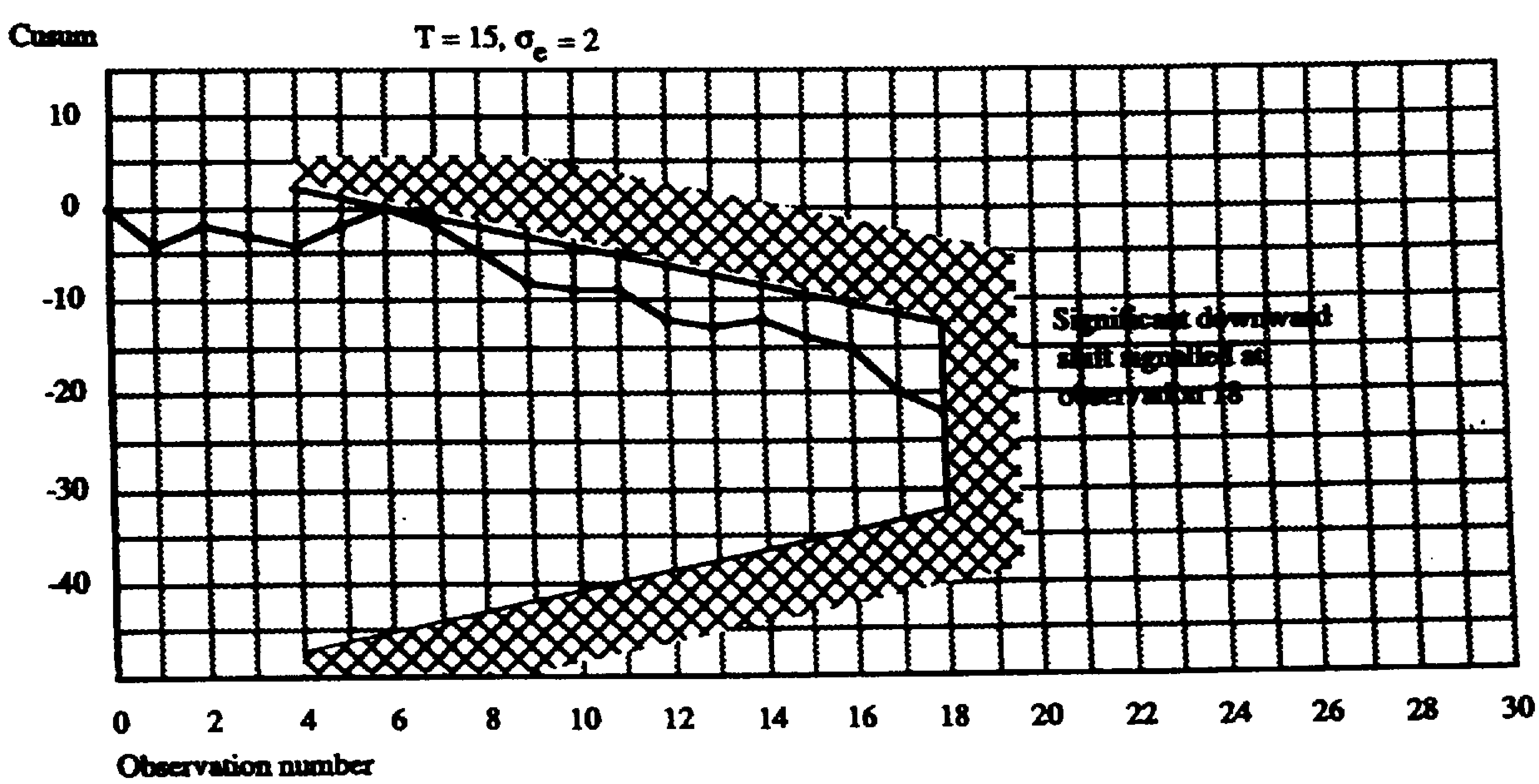


Figure 8 : Truncated V-mask applied to cusum chart : indication of shift

f) In a control situation, appropriate action would be indicated after plotting sample value 18 and obtaining a signal from application of the V-mask. An estimate of the new mean level is useful in determining the correct adjustment, and use of the method described in appendix C of this standard yields :

$C_{18} = -22,\ C_6 = 0$

$$\bar{y} = 15 + \frac{-22-0}{18-6} = 13,16$$

If action is delayed or ineffective, this estimate may be up-dated at a later stage if required.

4.3 The full V-mask

Appendix E details the construction of full V-masks.

4.4 Local decision lines

Appendix F details the construction of local decision lines.

4.5 Average run length characteristics of basic decision rule

4.5.1 This clause presents the average run length properties of the basic decision rule formulated in 4.2 and incorporated in the various graphical methods of cusum application by masks and decision lines. Readers familiar with Shewhart-type control charts may find it useful to compare this ARL curve with those for two well-established procedures :

a) Shewhart-type control chart with action limits only ;

b) Shewhart-type control chart with action limits and warning limits.

The ARL data refer to observations having a normal distribution with known standard error (σ_e) which is unaffected by changes in mean level of the observations. For non-normal distributions, or where variation may be affected by changes in mean, the ARL data may be considerably affected, both for Shewhart-type control charts and cusums. The average run lengths refer to two-sided control of the average.

The ARL data are listed in table 2 and presented graphically in figure 9.

Table 2 : Average run lengths for cusum and Shewhart-type control charts (h = 5 and f = 0,5)

Shift in process average from target value (units of σ_e)	CUSUM	Action limit 3,09 σ_e	Action and warning 3,09 σ_e, 1,96 σ_e	Action limit 3 σ_e	Action and warning 3 σ_e, 2 σ_e
0	465	500	323	370	278
0,25	139	374	236	281	199
0,50	38	201	109	155	101
0,75	17	103	50	81	49
1,00	10,4	55	26	44	26
1,50	5,8	18	8,9	15	8,8
2,00	4,0	7,3	4,2	6,3	4,1
2,50	3,1	3,6	2,5	3,2	2,0
3,00	2,6	2,2	1,8	2,0	1,7
4,00	2,0	1,2	1,2	1,2	1,2

Readers familiar with Shewhart-type control charts will recognize that in many countries the multipliers 3,09σ_e and 1,96σ_e are conventional, while in many other countries the multipliers 3σ_e and 2σ_e are conventional. As table 2 makes clear, the CUSUM ARL curves are significantly better than Shewhart-type charts using either convention ; better in the sense of long on-target ARLs combined with short ARLs in the important off-target range of about 1,0σ_e.

Decision rules :

CUSUM has H = 5σ_e, f = 0,5 (slope of decision line F = 0,5σ_e per sample interval).

Shewhart-type control charts :

Action limits at ± 3,09σ_e correspond to "1 in 500" points (99,8 %) of distribution.

Action limits ± 3,0σ_e correspond to "1 in 370" points (99,13 %) of distribution.

Warning limits at ± 1,96σ_e correspond to "1 in 20" points (5,0 %) of distribution.

Warning limits at ± 2,0σ_e correspond to "1 in 22" points (4,6 %) of distribution.

Signal generated by single action value only, or

Signal generated by single action value or two successive warning values on the same side of the target value.

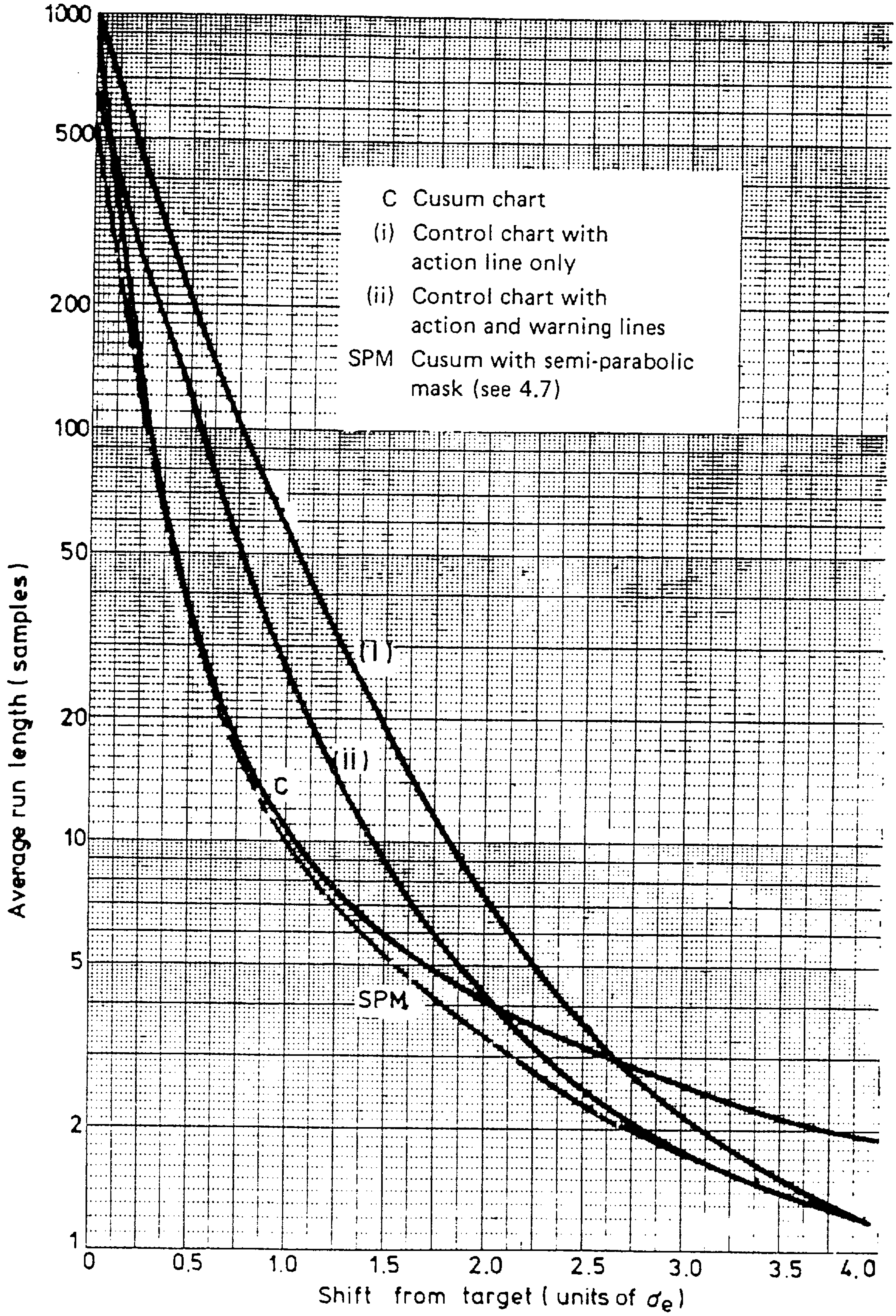


Figure 9 : Average run length data for cusum and Shewhart-type control charts (see table 2)

4.5.2 In general, when the design parameters of a CUSUM scheme are chosen to give an on-target ARL roughly equal to that of a conventional Shewhart-type scheme, the CUSUM scheme can be designed to give much shorter ARLs than the Shewhart scheme for shifts from target in the important range 0,5σ_e to 1,5σ_e. This is a decisive advantage for CUSUM, and is illustrated in table 2 and figure 9. Note the logarithmic ARL scale in figure 9. The Shewhart-type charts are faster in responding to very large process shifts, as can be seen from the crossing of the ARL curves in figure 9. However, both types of charts respond very rapidly to very large shifts and the practical difference is minimal.

4.5.3 The user of the basic cusum decision rule may find it useful to have some indication of the possible effect of non-normality on the ARL characteristics. In general, skewness with the longer tail in the same sense as the direction of potential shift in one-sided control will result in shortening the ARL at target, but it will have little effect on the ARL for large shifts in mean. Conversely, if the shorter tail is in the direction of potential shift, the ARL at target level will be considerably lengthened (again with little effect on ARL for large shifts).

If the distribution is discrete rather than continuous (as may occur with coarse measurements as well as with counting of nonconformities of defects), the ARL near target is again shortened, with little effect on ARL for large shifts.

The effects of skewness and discreteness may be illustrated by two examples :

a) for a "folded" normal distribution, as may occur in observing absolute deviations from a target quantity, the ARL at target conditions falls from 930 samples (for the normal distribution) to 370 samples ;

b) for a Poisson distribution with mean = 4 (so that the standard deviation is 2) a cusum scheme with H = 10 and f = 0,5 (slope of decision line = 1) has ARL 420 samples as compared with 930 samples for a normal distribution with the same mean and standard deviation.

In both cases, shifts which would yield ARL less than about 20 for a normal distribution yield very similar ARLs for the non-normal cases.

Table 3 : Effect of incorrect value of σ_e on ARL for one-sided control

Process average (as function of true σ_e)	Over-estimate σ_e by 10 %	Correct estimate of σ_e	Under-estimate σ_e by 10 %
T	3 000	930	410
T ± 0,5 σ_e	45	38	35
T ± σ_e	10	10,5	10
T ± 1,5 σ_e	6	5,8	6
T ± 2 σ_e	4,4	4,1	4,5

4.5.4 The importance of having a reliable estimate of σ_e has been noted in 2.3 of this standard. Under or over-estimation of σ_e will distort the ARL characteristics, again having little effect for large shifts but gross effects on ARL near target conditions. Thus for 10 % errors in estimation of σ_e, the distortions shown in table 3 could occur for one-sided control.

4.5.5 Finally, it should be noted that the effects of non-normality, error in estimating parameters and other possible departures from explicit or implicit assumptions apply in similar degree to the use of conventional control charts, and are not a peculiarity of cusum methods.

4.6 Modifications to the standard cusum procedure

Among the various modifications to the standard CUSUM that have been developed, three are described briefly in this standard as being especially useful.

The CUSUM design can be made more sensitive to shifts of a defined magnitude by adjusting the design parameters h and f. For the value f = 0,5 discussed in this standard, greatest sensitivity is focused on shifts of magnitude equal to σ_e.

A combined Shewhart-CUSUM scheme can provide enhanced sensitivity to very large deviations with little degradation of the other regions of the ARL curve.

A Fast Initial Response (FIR) CUSUM can provide greater sensitivity to being off-target during startup or after a change has been made in the process being monitored or controlled.

4.6.1 Combined Shewhart-CUSUM

The traditional action limits on Shewhart-type charts are set at either 3 (Shewhart version) or 3,09σ (99,9 percentile of normal distribution), with warning limits at 2 (Shewhart) or 1,96σ (97,5 percentile of normal distribution). To combine a Shewhart-control limit (SCL) with a CUSUM scheme to test for upward movement, we take the current observatory y_i (which may be an average for a routine sample) and test whether $y_i > m + SCL\sigma$, where m is the desired mean value and σ is the know standard deviation of the y observations. if y_i passes this test, it is added to the CUSUM in the usual way, if y_i fails this test then an out of control state is signalled. The modification if we test against downward movement is clearly to test whether $y_i < m - SCL\sigma$. The average run length of this modified procedure is given in table 5 for both SCL = 3,0 and SCL = 3,1. The increased sensitivity to large deviations is clear since 3σ displacements are detected (on average) more quickly by the revised scheme than 4σ displacement by the original scheme and four displacements have ARL's to signal approximately 1,2 in place of 2 sampling intervals. The sensitivity may itself be an embarrassment if there is an appreciable probability of obtaining faulty y observations unrepresentative of the process at the time of sampling. There is more likely to happen when y is in fact a single observation of the process (not an average of a routine sample). This sensitivity to spurious values may be compensated for by treating an observation as an outlier if it lies outside specified limits and replacing it by a second sample if practical or by the next sample if not. Two successive outliers require immediate attention even if they lie on opposite sides of the desired mean value. This form of presenting the procedure is to be preferred to one which ignores the outliers completely and to a lesser extent it is to be preferred to a procedure which replaces the outlier by the nearest value which would not be classified as outlying.

4.6.2 Combined snub-nosed scheme

Greater sensitivity to large departures from the in-control state can also be achieved by running parallel CUSUM procedures with different values for h and f. For example if the values h = 5, f = 0,5 are used to control the smaller displacements of the mean then using h = 1,55 and f = 1,55 will reproduce the effect of SCL = 3,1 and also produce an approximation to the use of warning limits. The ARL will be approximately the same as the SCL = 3,1 scheme for low ARL's. In general the true average run length will not be greater than the minimum of the average run lengths under the parallel CUSUM schemes. For many purposes this knowledge will be sufficient without a detailed calculation of the average run length for the combined schemes. The control procedures requires only two running totals, it is referred to as a snub-nosed CUSUM mask.

Table 4 : Data for construction of a typical snub-nosed CUSUM mask

Distance from datum (sample intervals) i	Half-width of mask at 1 (units of σ_e) y	
0	1,55	$y = 1,55\ (i + 1)$
1	3,1	
2	4,65	
3	6,2	
4	7,0	
5	7,5	$y = 5 + 0,5\ i$
6	8,0	
7	8,5	
8	9,0	
9	9,5	
10	10,0	
15	12,5	
20	15,0	

The equations may be used for construction if required.

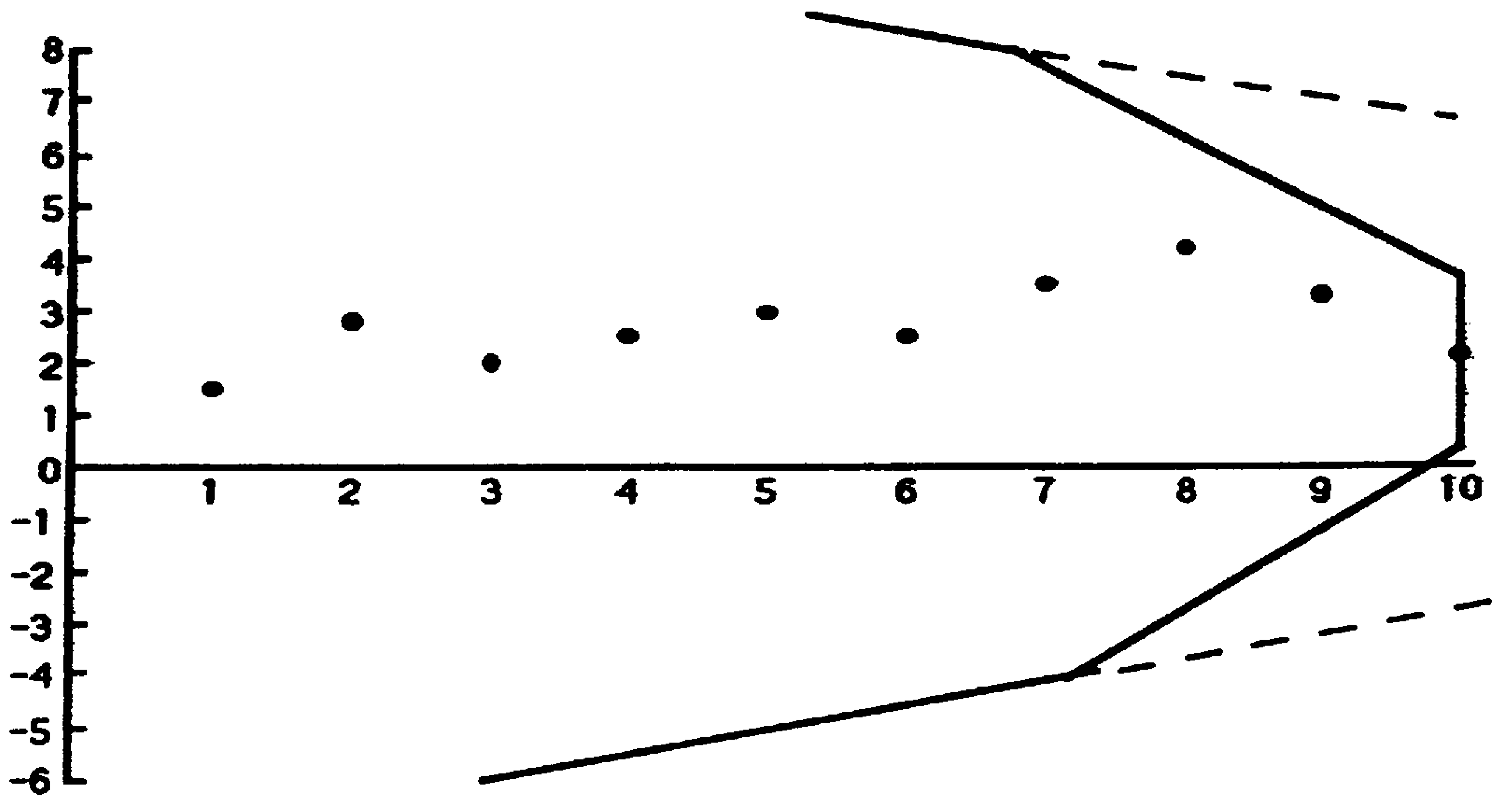


Figure 10 : Snub-nosed CUSUM mask

Table 5 : Comparison of average run lengths for various modified procedures

Displacement of process mean (units of σ_e)	Standard CUSUM	Single sided schemes		Snub-nosed scheme	FIR scheme
		Combined SCL = 3,0	Shewhart-CUSUM SCL = 3,1	h = 5,0	f = 0,5 S = 2,5
0	930	448	529	472	896
0,25	140	112	119	114	125
0,50	38	34,8	35,6	35,5	28,8
0,75	17	16,1	16,4	16,1	11,2
1,0	10,5	9,8	10,0	9,9	6,3
1,5	5,8	5,3	5,4	5,2	3,4
2,0	4,1	3,5	3,6	3,2	2,4
2,5	3,2	2,4	2,5	2,3	1,9
3,0	2,6	1,8	1,9	1,7	1,5
4,0	1,9	1,2	1,2	1,2	1,2

Table 6 : ARL's for one sided CUSUM scheme S(0) = H/2

Parameters		Displacement of current mean (multiple of true value of sigma)										
H	K	S(0)	0,000	0,250	0,500	0,750	1,00	1,50	2,00	2,50	3,00	4,00
2,50	0,25	1,25	22,87	10,32	5,668	3,663	2,667	1,760	1,368	1,169	1,068	1,005
4,00	0,25	2,00	66,57	20,25	8,991	5,291	3,700	2,354	1,770	1,441	1,233	1,040
6,00	0,25	3,00	255,4	38,76	13,48	7,382	5,054	3,169	2,372	1,932	1,638	1,227
8,00	0,25	4,00	684,3	63,25	17,86	9,416	6,390	3,967	2,940	2,383	2,046	1,605
10,00	0,25	5,00	1972,0	93,75	22,13	11,43	7,724	4,767	3,509	2,815	2,381	1,933
2,00	0,50	1,00	34,40	15,19	7,785	4,626	3,126	1,873	1,395	1,174	1,069	1,006
3,00	0,50	1,50	108,0	33,39	13,25	6,755	4,208	2,353	1,680	1,348	1,165	1,023
4,00	0,50	2,00	316,4	66,57	20,25	8,991	5,291	2,862	2,014	1,586	1,325	1,067
5,00	0,50	2,50	895,9	124,9	28,76	11,24	6,348	3,372	2,362	1,856	1,540	1,159
6,00	0,50	3,00	2492,0	225,4	38,76	13,48	13,48	7,382	3,875	2,703	1,774	1,311
1,50	0,75	0,75	39,44	18,70	9,747	5,658	3,659	2,015	1,431	1,182	1,070	1,006
2,25	0,75	1,12	131,9	46,22	18,72	9,004	5,140	2,499	1,670	1,314	1,138	1,017
3,00	0,75	1,50	426,6	108,0	33,39	13,25	6,755	3,010	1,950	1,489	1,243	1,040
3,75	0,75	1,87	1345,0	242,8	56,40	18,36	8,428	3,528	2,251	1,696	1,386	1,085
4,50	0,75	2,25	4193,0	534,0	91,74	24,32	10,12	4,043	25,60	1,922	1,559	1,160
1,00	1,00	0,50	33,54	17,78	10,04	6,084	3,982	2,135	1,466	1,190	1,071	1,006
1,50	1,00	0,75	89,74	39,44	18,70	9,747	5,658	2,608	1,657	1,282	1,115	1,012
2,00	1,00	1,00	250,2	88,45	34,40	15,19	7,785	3,126	1,873	1,395	1,174	1,023
2,50	1,00	1,25	699,9	195,9	61,68	22,87	10,32	3,664	2,107	1,529	1,252	1,040
3,00	1,00	1,50	1934,0	426,6	108,0	33,39	13,25	4,208	2,353	1,680	1,348	1,067
3,50	1,00	1,75	5292,0	918,6	185,9	47,59	16,56	4,752	2,606	1,843	1,460	1,107
0,70	1,50	0,35	66,43	34,92	19,31	11,27	6,966	3,204	1,887	1,367	1,146	1,016
1,10	1,50	0,55	181,2	83,78	40,61	20,79	11,36	4,279	2,225	1,502	1,207	1,026
1,50	1,50	0,75	542,8	216,1	89,74	39,44	18,70	5,658	2,608	1,657	1,282	1,04
1,90	1,50	0,95	1748,0	585,2	203,5	75,30	30,51	7,325	3,020	1,828	1,371	1,062
2,30	1,50	1,15	5871,0	1619,0	464,3	142,8	48,99	9,260	3,447	2,012	1,473	1,09

Parameters :

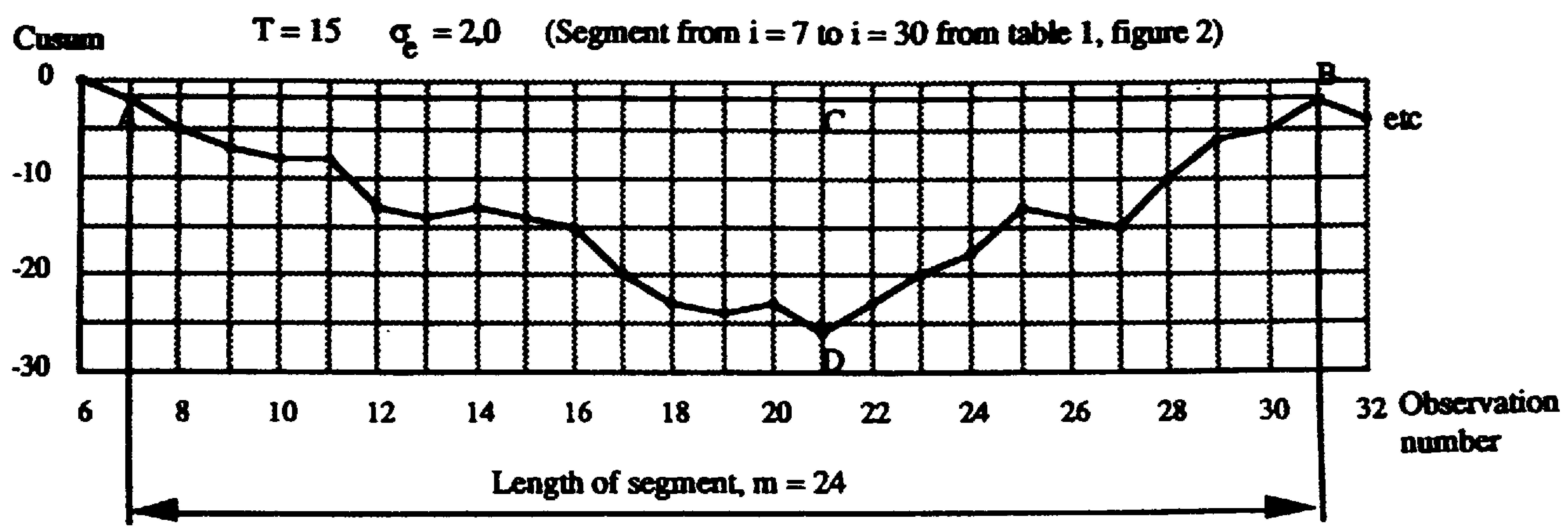
- H = Decision interval for the CUSUM scheme ;
- K = Reference value ;
- S(0) = Headstart value.

4.7 Non-standard average run lengths

The average run lengths described for the various schemes are appropriate when the CUSUM is set at zero when the change from in-control conditions occurs. Although this condition is not realistic in practice, the ranking it produces for alternative CUSUM schemes is satisfactory since the expected value of the cumulative sum is zero in the steady state in-control condition and the distribution of starting points is concentrated about zero.

There are circumstances, however, in which the average run lengths from a starting point other than zero are important. When starting up a process, or readjusting the process after an out of control situation has been signalled, there may be some doubt as to whether the process is in control to begin with. There are advantages here in having an early warning of unsatisfactory adjustment to the process parameters. This can be achieved by giving the decision boundary (either upper or lower) the standard starting value of zero. With a "head start", s, those values of the process mean that are at some distance from the reference value have an effective decision interval close to (h - s) if the mean is on the unfavorable side and an effective decision interval h if the mean is on the favorable side of the reference value. Intermediate values have intermediate effective decision intervals. Table 6 gives some values of ARL's for selected CUSUM schemes with "head starts" equal to h/2 are of interest. A comparison with other schemes is given for h = 5,0, f = 0,5 and head start equal to 2,5 in table 5.

The use of a non zero starting value for the cumulative sum is known as the fast initial response procedure. The effect of the initial setting lasts until either the CUSUM attains the value zero before signalling an out of control state or signals an out of control state without passing through the value zero. In the first case the CUSUM behavior reverts to the standard mode starting from zero, in the second case the out of control signal is given earlier than would have been the case with the standard mode of operation for most sequences and never later.



CD = 24 = V_{max} V_{max}/σ_e = 12,0 for m = 24
p (from figure 12) < 0,001, change is highly significant

Figure 11 : Application of span test to CUSUM chart

5 Decision rules for retrospective analysis

5.1 Introduction

5.1.1 The fundamental difference between retrospective analysis and monitoring is that for the latter (as detailed in the previous clause) the mean levels in all suspected segments are compared with a target value, specified in advance. In the case of retrospective analysis, the interest centres on differences between adjacent segments ; sometimes they may be prior grounds for the definition of the segments (e.g. changes in operating or experimental conditions occurred at known points in the sequence), but often the basis for segmentation and estimation of points of change will be the appearance of the cusum chart itself.

5.1.2 Where a decision rule is applied to test whether the apparent difference between segments may be merely a manifestation of the underlying variation in the sequence, the problem is analogous to the use of tests of significance of the difference in mean between two groups of observations. Indeed, where the segmentation is based on external factors, conventional tests of the "no difference in mean between two groups of observations" hypothesis may be applied, and the CUSUM regarded merely as a useful device for data presentation (and perhaps further diagnosis). However, where the segmentation is based on inspection of the chart, the initial hypotheses are formulated after inspection of the data, so that conventional tests are no longer applicable.

5.1.3 In the course of retrospective analysis of a long series of observations, numerous tests of the differences between successive segments may be applied. Each of these tests carries a risk of type 1 errors : an incorrect decision that a difference exists when no real change has occurred. When many tests are carried out, the overall significance of apparent differences between segments should therefore be interpreted with caution. It is suggested that, when testing a difference between two segments of total length m in a series of N observations in all, the required significance level, α, for individual tests should be modified to $m\alpha/N$ to avoid excessive type 1 errors in examining the complete series. Alternatively, the observed probability of the test statistics for individual tests may be multiplied by N/m.

5.2 "Span" tests

5.2.1 When inspecting a cusum chart for possible points of change in average level of the plotted variable, the basis for segmentation is the occurrence of local maxima or minima in the CUSUM. A useful criterion for assessing the significance of such change is the maximum extent to which the CUSUM deviates from a straight line joining the ends of the sequence within which a change is suspected.

5.2.2 For a CUSUM chart, the maximum vertical height, V_{max} standardized by dividing the observed height by the standard error of the observations, is measured by drawing a chord (AB in figure 11) between the points at the ends of the sequence, and measuring the maximum vertical distance (CD) from the CUSUM to the cord. This maximum will occur at an apparent turning point in the CUSUM : either a local maximum or minimum, or a point where the slope becomes steeper or flatter ; such points are termed "corners".

The value of V_{max}/σ_e is then referred to the nomogram of figure 12 for the appropriate span, i.e. the horizontal length of the segment AB, in sample intervals. The P-scale of figure 12 then gives the probability of exceeding V_{max}/σ_e in a sequence of length m from a series of independent standard normal observations.

5.2.3 Alternatively, where the original CUSUM calculations are available, the maximum vertical height may be calculated from :

$$V_{max} = C_r - C_i - \frac{r-i}{j-1}\left(C_j - C_i\right), m = j - i$$

where r is the sample number corresponding to a suspected change point in the sequence i + 1 to j (i.e. i, r and j are successive "corners" on the CUSUM charts).

Thus for the sequence i = 7, r = 21, j = 31 on figure 2, from table 1 :

C_i = - 1, C_r = - 26, C_j = - 2

hence :

$$V_{max} = -26 - (-1) - \frac{21-7}{31-7}\left(2 - (-1)\right) = -25 - \frac{14}{24} x - 1 = -24,42$$

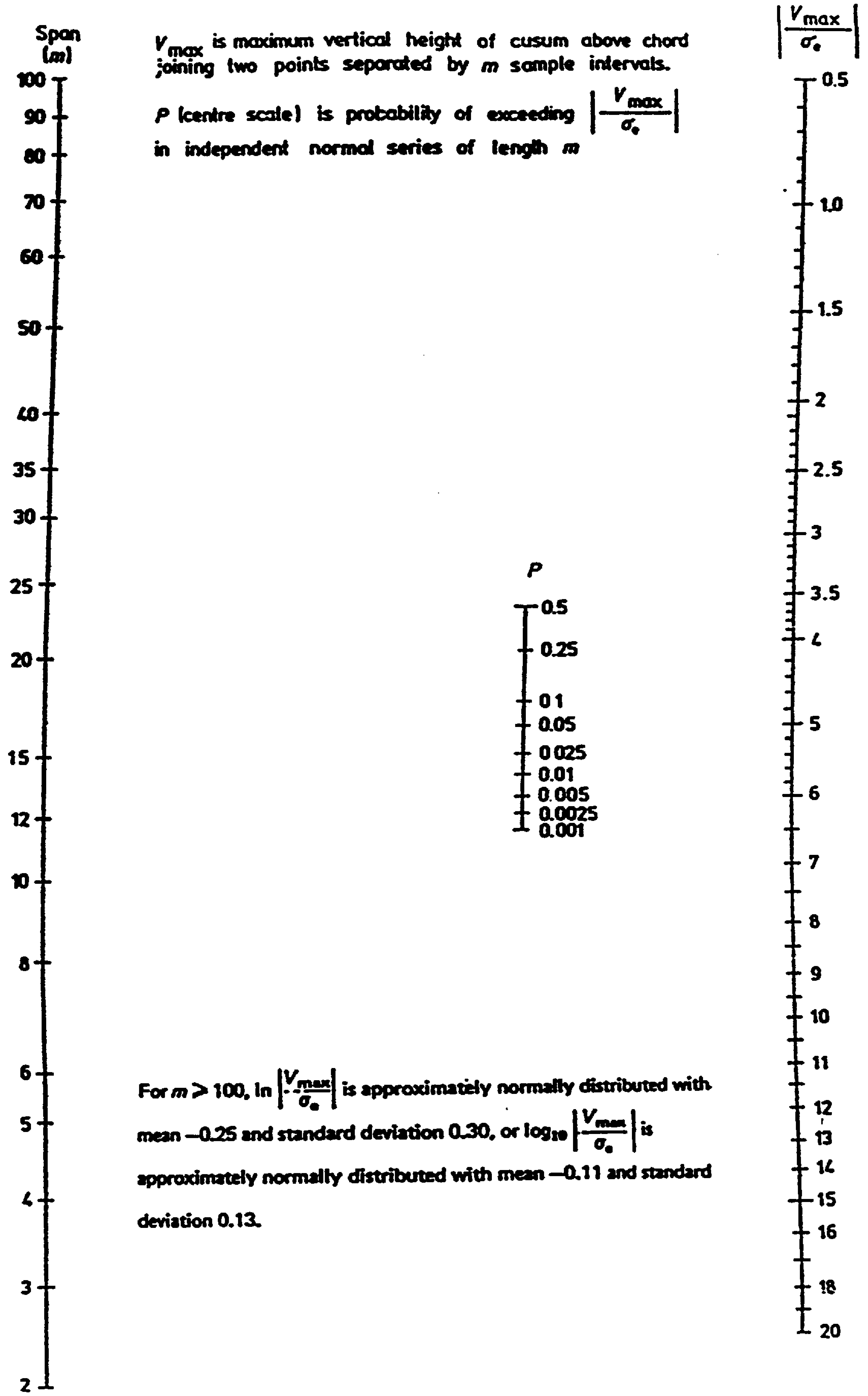


Figure 12 : Nomogram for evaluating maximum vertical height (Vmax) of CUSUM over span of m sample intervals

With σ_e = 2,0 for this series :

$$|V_{max} / \sigma_e| = \frac{24,42}{2} = 12,21$$

From figure 12, this value lies well below the 0,001 point for m = 24, and the change in mean level is a highly significant indication of a real shift in the average level of the series under observation. The analysis by direct measurement from the chart gives the same conclusion, subject only to limitations of graphic precision.

5.3 Distribution-free CUSUM tests

Appendix G details distribution-free CUSUM tests.

6 Computational CUSUM techniques

6.1 CUSUMS without charts

6.1.1 Where the purpose of a CUSUM procedure is mainly to detect off-standard conditions, rather than to present a graphical summary of sequential data, the information available at any point in the sequence (usually a time sequence in monitoring applications) can be recorded in a CUSUM tabulation instead of a chart. A numerical decision rule then replaces the mask used with a chart. To introduce the tabulation method, it will be useful to present the use of a decision line on a CUSUM chart a slightly different format.

6.1.2 Instead of erecting the vertical decision interval through a suspected change point, and then drawing the sloping decision line from the end of the vertical line (as in appendix F and figures 22 and 23), begin by drawing a line of the same slope as the decision line, but passing through the suspected change point (AB in figure 13). This "reference line" will slope upward if the mean level is suspected to have increased, and downward if a decrease in mean is apparent. A signal will then be generated if the CUSUM path rises more than H units above an upward reference line or falls more than H units below a downward reference line. This alternative construction will yield identical signals to the usual decision line or mask (given equal slopes and H-values).

6.1.3 The decision line CD is thus parallel to, and H vertical units above or below, the sloping reference line AB. Now slope on a CUSUM chart represents average level in the observations, so that the reference lines with slopes ± F (i.e. $\pm f\sigma_e = \pm 0{,}5\sigma_e$ in this case) correspond to hypothetical mean levels to T ± F (here $T \pm f\sigma_e$, i.e. $T \pm 0{,}5\sigma_e$). If we construct a CUSUM chart by calculating $d_{1i} = y_i - (T + 0{,}5\sigma_e)$ or $d_{2i} = y_i - (T - 0{,}5\sigma_e)$ and then plotting :

$$S_{1i} = \max. [S_{1,i-1} + d_{1i,}0]$$

or

$$S_{2i} = \min. [S_{2,i-1} + d_{2i,}0]$$

where :

$$S_{1,0} = S_{2,0} = 0$$

we have a partial CUSUM plot of sequences where the mean either exceeds the upper reference value $K_1 = T + 0{,}5\sigma_e$ or falls short of the lower reference value $K_2 = T - 0{,}5\sigma_e$. A horizontal decision line may now be drawn at +H or -H units from the origin of the CUSUM scale ; if the CUSUM path crosses either of these lines ; a signal occurs. The sequence of figure 13 is plotted in this format in figure 14.

6.1.4 A computational CUSUM has the following advantages :

a) the decision rules are condensed into two parameters, the decision interval H and the reference shift given by $F = K_1 - T = T - K_2$. Each may be referred to parameters expressed in terms of standard error. $H = h\sigma_e$, $F = f\sigma_e$. It may also be noted that, where the exact $2\sigma_e$ scale convention is adopted, f/2 has an interpretation as the tangent of the angle between the decision line and the sample (horizontal) axis of the chart. For f = 0,5, we thus have $f = \tan^{-1} 0{,}25 = 14°$. This interpretation is invalidated by any departure from a scale factor of $2\sigma_e$ per sample interval.

b) any decisions become independent of precision if plotting values on the chart, as the table of calculations can be used directly. Any upper CUSUM exceeding H or lower CUSUM falling below - H produces a signal, and a chart is unnecessary ;

c) dispensing with a chart, whist sacrificing direct presentation of the data, permits the use of routine tabulation sheets or automatic data storage methods, including computers. This can be especially useful when many processes are controlled simultaneously, or where automatic instrumentation is linked to an on-line computer system ;

d) in many situations, satisfactory quality may be achieved by maintaining the process average within an acceptable quality band rather than at a single target value. Readers familiar with control charts will recognize this as the situation where modified control limits are applicable for a "high precision" process. In these cases, two target levels may be chosen to correspond to the outer edges of the acceptable region, T_1 et T_2. The reference values for tabulation are then set at $K_1 = T_1 + f\sigma_e$ et $K_2 = T_2 - f\sigma_e$. The CUSUM is then dormant until any value exceeds K_1 or falls below K_2, and attention is drawn to the process when a possible out-of-control situation occurs. Under these circumstances, simultaneous activity in the upper and lower CUSUMs is unlikely. Furthermore, some difficulty may be encountered plotting a satisfactory CUSUM chart where T_1 and T_2 are widely separated, and two CUSUM schemes (either charts based on T_1, T_2 or tabulations using K_1, K_2) are preferable.

6.1.5 The calculations on which this modified CUSUM is based are shown in table 7. It is apparent that when there are some disadvantages in using this procedure :

a) the chart no longer gives a complete presentation of the series of observations ; plotting occurs only when a possible change is being evaluated ;

b) two CUSUM schemes need to be operated simultaneously, and they may both be active for short periods (as samples 22 to 25) ;

c) the two reference values (K_1, K_2) have less intuitive appeal than the target value ;

d) in consequence of a), local averages can be calculated only over segments whose means differ from the target by $0{,}5\sigma_e$ or more.

Table 7 : CUSUM calculations for modified procedure

i	x_i	Upper CUSUM		Lower CUSUM	
		x_i - 16	$S_1 = \sum(x_i - 16)$	x_i - 14	$S_2 = \sum(x_i - 14)$
1	12	- 4	0	- 2	- 2
2	17	1	1	3	0
3	14	- 2	0	0	0
4	14	- 2	0	0	0
5	17	1	1	3	0
6	16	0	1	2	0
7	14	- 2	0	0	0
8	11	- 5	0	- 3	- 3
9	13	- 3	0	- 1	- 4
10	14	- 2	0	0	- 4
11	15	- 1	0	1	- 3
12	11	- 5	0	- 3	- 6
13	14	- 2	0	0	- 6
14	16	0	0	2	- 4
15	13	- 3	0	- 1	- 5
16	14	- 2	0	0	- 5
17	11	- 5	0	- 3	- 8
18	12	- 4	0	- 2	-10 1)
19	13	- 3	0	- 1	- 11
20	16	0	0	2	- 9
21	12	- 4	0	- 2	- 11
22	18	2	2	4	- 7
23	18	2	4	4	- 3
24	17	1	5	3	0
25	20	4	9	6	0
26	15	- 1	8	1	0
27	14	- 2	6	0	0
28	18	2	8	4	0
29	20	4	12 1)	6	0
30	16	0	12	2	0
31	18	2	14	4	0
32	14	- 2	12	0	0
33	16	0	12	2	0
1) Decision interval reached or exceeded = Decision line touched or crossed in figure 14.					

$T = 15$, $\sigma_e = 2{,}0$ Upper reference value $K_1 = T + (0{,}5 \times 2{,}0) = 16$

Lower reference value $K_2 = T - (0{,}5 \times 2{,}0) = 14$

Decision interval $H = 5{,}0 \times 2{,}0 = 10$

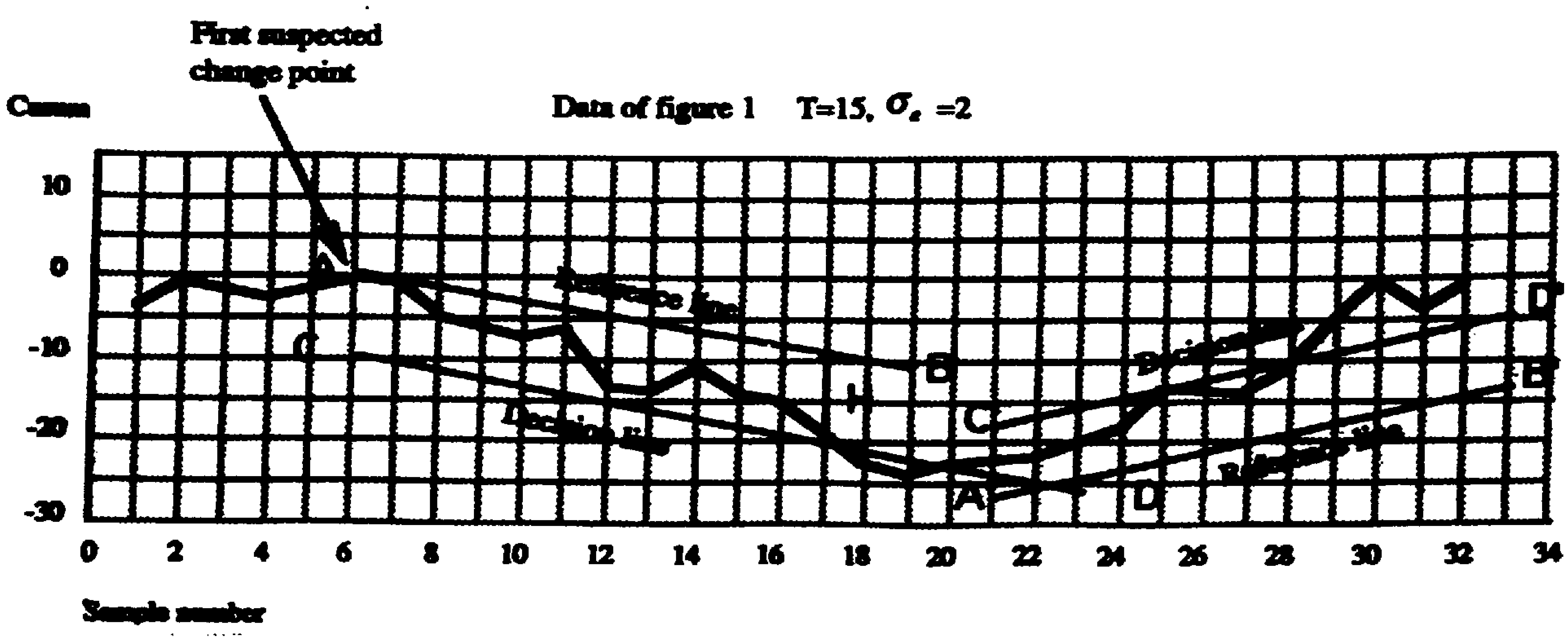


Figure 13 : Alternative presentation of decision lines

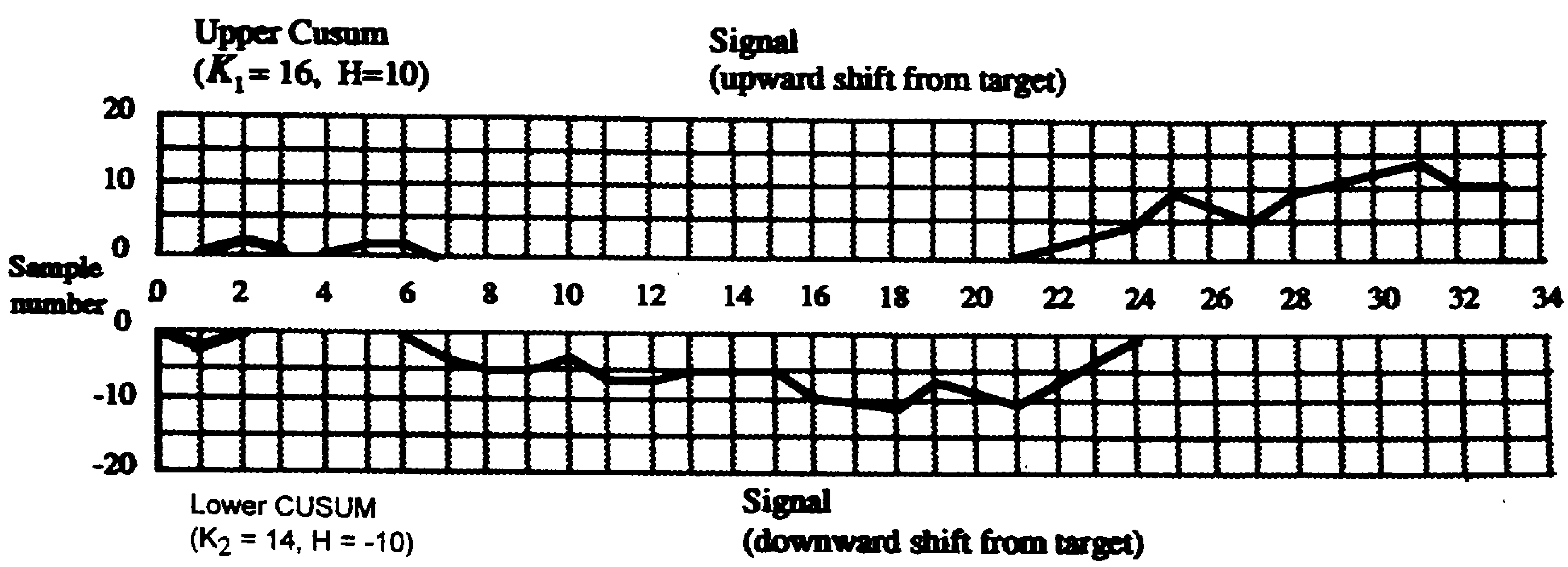


Figure 14 : Modified CUSUMs S_1 and S_2

6.1.6 For particular purposes, parameters other than h = 5 and f = 0,5 may be preferred (see appendix H).These will give different ARL characteristics suitable to some circumstances. This subject will be developed in other standards, but it may be noted that any pair of (h, f) parameters may be readily converted to a CUSUM mask, using $h\sigma_e$ as the vertical half-height at the root of the mask, and $f\sigma_e$ per sample interval as the slope of the mask. Figure 15 shows a simple means of setting up such a mask. In some cases (where only upward or downward shifts, but not both are of interest) only a half-mask will be required.

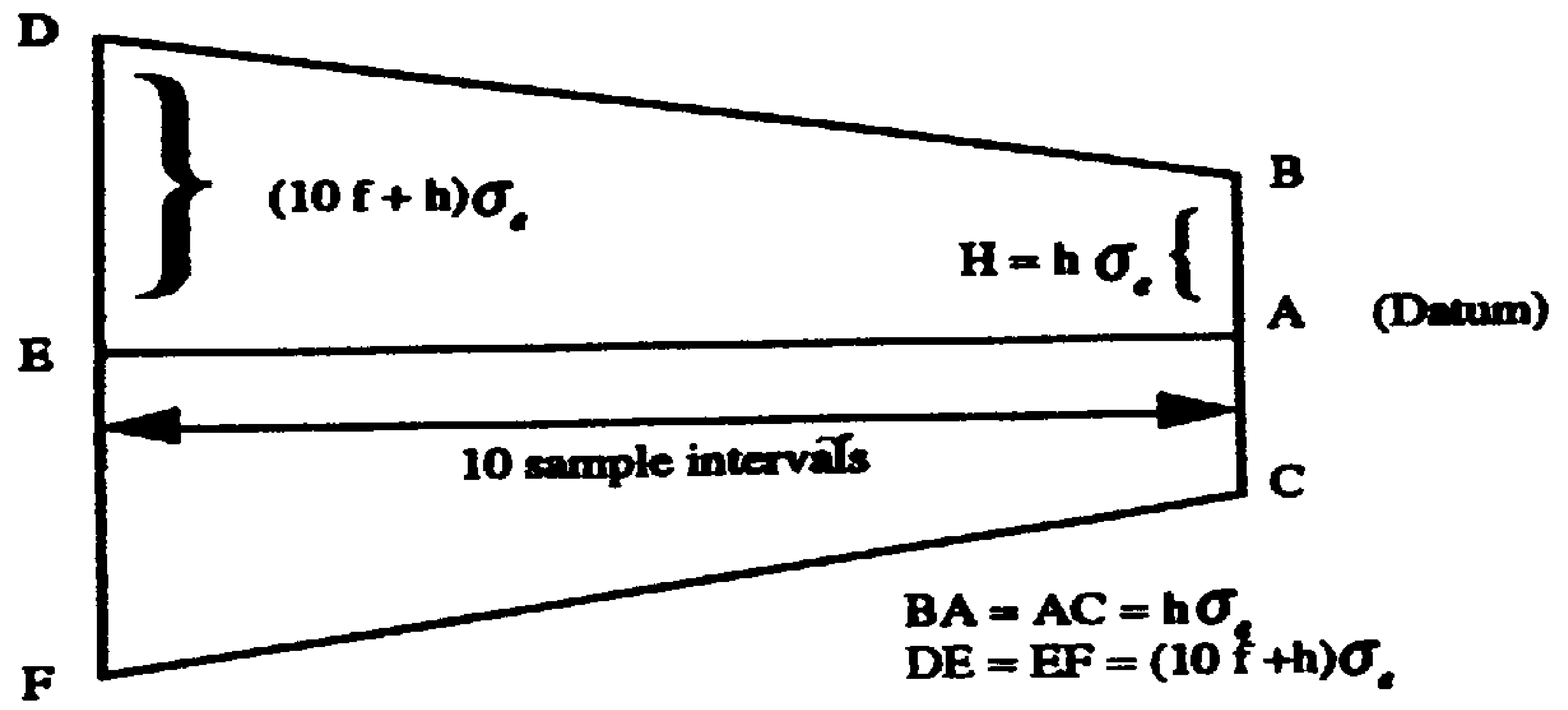


Figure 15 : Generalized CUSUM mask construction

7 Examples of applications

7.1 Example 1

7.1.1 A common quality control problem is the need to keep a product characteristic within a tolerance interval which is narrow relative to the process capability so that very little latitude is permitted to the mean process level. Consider the case of fibre boards. The specification for individual boards requires the width to be measured 100 mm from each end of the board and these widths, to the nearest mm, must lie in the interval 1 197 mm to 1 203 mm, if the board is to conform to specification. The acceptability of batches of boards is determined by a sampling clause with an AQL of 10 % for nonconforming boards. The standard deviation for width measurements made within one batch is reliably estimated to be 1,6 mm, and this implies that approximately 3 % of measurements would be recorded as outside limits if the distribution were normal centered at 1 200 mm. Since there are two measurements on a board effectively independent of each other, this would give 6 % of boards outside the specification limits. Even a 1 mm deviation of the process mean would give rise to greater than 10 % of nonconforming boards.

7.1.2 These circumstances dictate the choice of 1 200 mm as the reference value and other parts of the specification require that 5 boards be selected from each batch, hence the statistic used for the CUSUM chart is the average of the 10 widths measured on the 5 boards. The average of the 10 widths has a standard error σ_e of $1{,}6 / \sqrt{10}$ mm or 0,506 mm. For simplicity we use σ_e = 0,5 mm to give H = 2,5 mm and k = 0,25 mm. The first 19 points of a CUSUM chart prepared with these values and with the sample averages given in table 8 are illustrated in figure 16. The CUSUM wanders without crossing the V-mask until the 19th observation. The immediately preceding points indicate a change around samples 16 or 17. The chart also records that between samples 15 and 16 the machine was overhauled and since the overhaul the mean width is estimated to be :

$$1200 + \left[\frac{(4{,}5) - (0{,}2)}{4}\right] \text{mm} = 1\,201{,}1 \text{ mm}$$

7.1.3 After the process has been adjusted the cumulative sum can be restarted from zero without affecting subsequent performance of the CUSUM chart. If, however, it is desired to use the fast initial response feature then although the CUSUM will start from zero there will be two dummy points corresponding to the "head-start" in both the upward and downward directions and indicated by crosses on figure 16.

Table 8 : Average width of 5 fibre board panels selected from each sampling period

Sample n°	Average width (mm)	Average width -1 200 mm	Cumulative sum
1	1199,5	- 0,5	- 0,5
2	1201,9	+ 1,9	+ 1,4
3	1199,8	- 0,2	+ 1,2
4	1199,5	- 0,5	+ 0,7
5	1199,0	- 1,0	- 0,3
6	1198,5	- 1,5	- 1,8
7	1202,1	+ 2,1	+ 0,3
8	1199,8	- 0,2	+ 0,1
9	1201,2	+ 1,2	+ 1,3
10	1199,7	- 0,3	+ 1,0
11	1200,2	+ 0,2	+ 1,2
12	1200,7	+ 0,7	+ 1,9
13	1199,2	- 0,8	+ 1,1
14	1199,1	- 0,9	+ 0,2
15	1200,0	0,0	+ 0,2
16	1199,5	- 0,5	- 0,3
17	1201,3	+ 1,3	+ 1,0
18	1201,3	+ 1,3	+ 2,3
19	1202,2	+ 2,2	+ 4,5
Break in cumulative sum			
20	1199,7	- 0,3	- 0,3

7.2 Example 2

7.2.1 It sometimes happens that a pilot quality monitoring procedure has to be set up without a preparatory study of process capability. Obviously, such procedures should be revised as soon as some operating experience is gained but a specification may provide the only basis for the initial or tentative scheme.

As an example, consider the hypothetical case of a packer required to meet a specification that controls both average weight and individual weights for pre-packaged goods.

7.2.2 For packages of nominal weight 250 g, the requirements are that the mean weight should be not less than nominal, and that few packages should have weights below 241 g. In this context, it appears reasonable to interpret the "specification" in terms of a process mean over 1 h production, and "few packages" as implying not more than 2,5 % of packages in the long run. It may also be reasonable to assume that, under controlled conditions, a filling machine will produce normally distributed package weights. Thus if the process mean happens to be at just the nominal weight, a standard deviation of 4,59 g will yield just 2,5 % of packages below 241 g. The packer might thus regard a standard deviation of 4 g as giving a satisfactory product, provided the mean weight is carefully controlled.

Let it be further supposed that the packer wishes to control the mean level slightly above the nominal, so that, if it should fall to 250 g this condition will be detected within 1 h. With a standard deviation for individual packages of 4 g, a number of options are open.

7.2.3 If the packer chooses a sample size of 4 packages, the standard error of the sample mean (given random samples of the packages produced between sampling times) will be $s / \sqrt{4} = 2$ g. If a sampling interval of 10 min is chosen, the question to be answered is "at what target level must the process be operated so that a drop of 250 g will be detected with ARL 6 samples, given $\sigma_e = 2$ g ?"

Reference to figure 9 indicates that a shift of about 1,5 standard errors is detectable with this ARL. The target level should thus be set a $1{,}5\sigma_e = 3$ g above the nominal, i.e. 253 g.

Alternatively, the packer may prefer, say, samples of six items every 15 min so that an ARL of four samples is implied. The target then has to be set about $2\sigma_e$, above nominal, and with $s / \sqrt{6} = 1{,}63$, a target level of 253,3 (slightly more than the previous value) is required. In fact the total rate of sampling (24 per hour) in the two cases is identical, and in the absence of complications from within sample and between-sample components of variation, small frequent samples are preferable to large infrequent ones. There is also (with frequent sampling) the advantage that a gross change may be detected in the first sample taken, soon after the actual point of change.

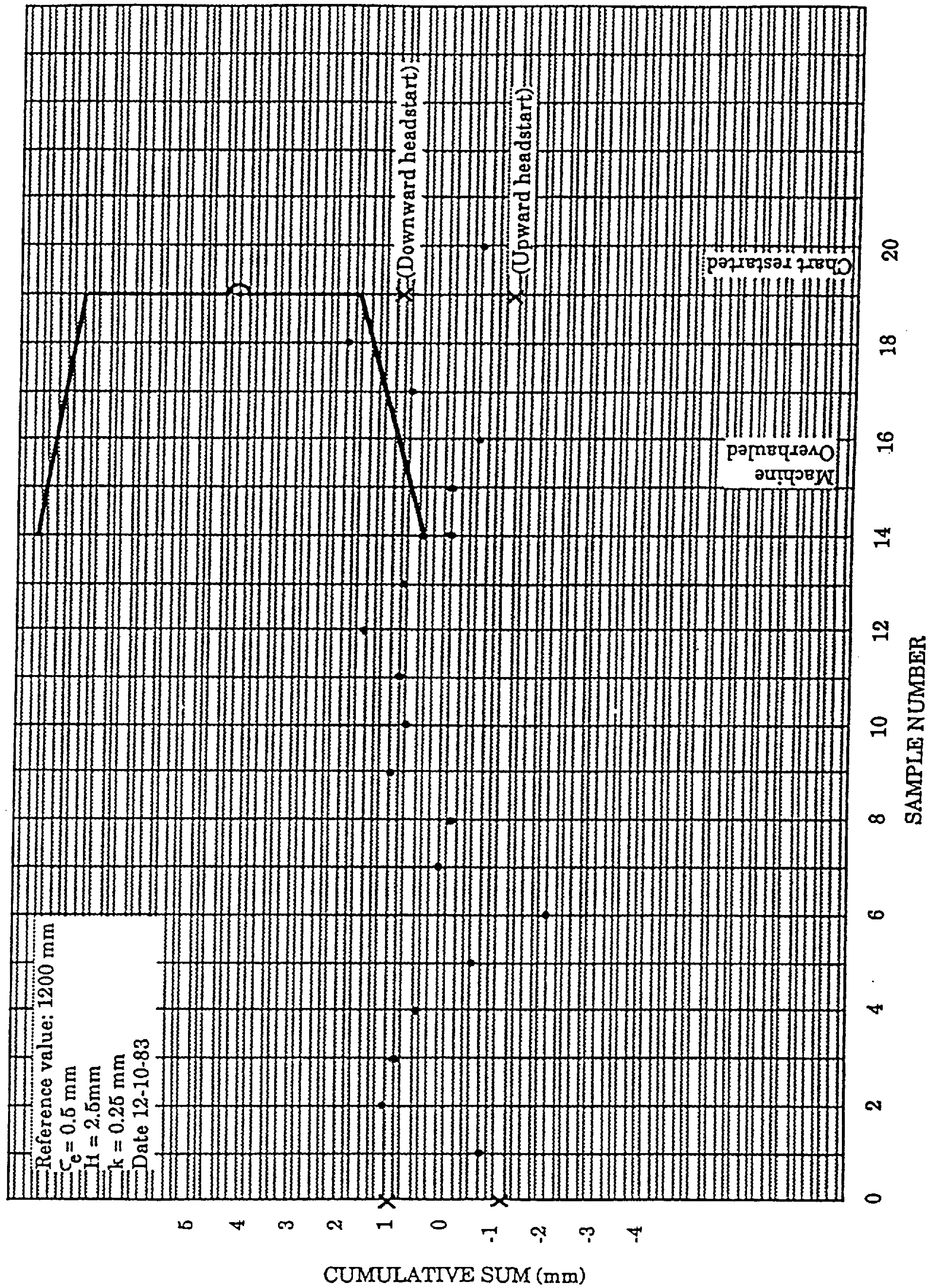


Figure 16 : Average width for samples of 5 fibre boards

7.2.4 However, the packer may regard 253 g as representing an excessive "overfill" : on average, the packages exceed the nominal weight by 1,2 %. If it is wished to reduce the target fill to a value closer to the nominal weight, the prospect of a higher level of sampling must be faced. For example, if a target is chosen for a fill of 252 g, with a sample size of 4, the difference between target and nominal weights is now one standard error. A shift of this magnitude requires about 10 samples, on average, for its detection (from figure 9, ARL = 10 for $1\sigma_e$ shift), implying sampling every 6 min.

7.2.5 The example should not be taken too far. It is purely an indication of how a pilot scheme might be set up from a specification in the absence of process data. In a real situation, the assumed standard deviation of 4 g would need to be monitored, and the scheme redesigned on the basis of available data as soon as practicable. In particular, the above procedure should not be regarded as an effective means of satisfying legal requirements for package fill control, which should be based on real performance data. Such data might not involve true random sampling, so that $\frac{\sigma}{\sqrt{n}}$ may not be a suitable estimate of standard error.

7.3 Example 3

7.3.1 In the manufacture of a synthetic resin for use in glues and plastics, it was considered important that the proportions of volatile matter should not exceed an average level of 1 %. Observations made on the process over a series of batches, produced under what were regarded as stable conditions, yielded the following data :

Number of batches : 42

Overall mean : 0,8893 (% volatile matter)

Standard error of means :

$$\left.\begin{aligned} &\text{a) } \sqrt{\frac{1}{41}\sum\left(\bar{y}-\bar{\bar{y}}\right)^2} = 0{,}0669 \\ &\text{b) } \sqrt{\frac{1}{2{,}41}\sum\left(\bar{y}_i-\bar{y}_{i-1}\right)^2} = 0{,}0652 \end{aligned}\right\} \text{See appendix A of this standard}$$

With so little difference between the two estimates of standard error, its appears unlikely that the process suffers cyclic fluctuations in average, and a rounded value σ_e = 0,065 may be assumed to be appropriate for setting up a CUSUM procedure.

7.3.2 In this case, there were no specified requirements for the process, the 1 % upper limit being regarded as a level at which fairly rapid detection should be achieved. The level of 0,89 % was regarded as a satisfactory process target, and considerations of sampling frenquency did not arise : every batch is adequately tested. The objective is to detect any sustained upward shift from the target level, and to evaluate within how many batches a shift of any given magnitude will be detected.

A CUSUM computer scheme was set up with target value T = 0,89 %, σ_e = 0,065 %, reference value K = T + 0,5σ_e = 0,923 and decision interval H = 5σ_e = 0,325. A typical series of test results are listed in table 9. The decision interval is exceeded at sample 31. A calculation of the average level since the minimum of the CUSUM shows the mean is 0,947 %. Thus although the critical value of 1 % has not been reached, some fundamental shift appears to have occurred, requiring investigation and corrective action.

7.3.3 We can check the average run length to a signal of upward movement by using figure 9. This shows that when the mean is at the target level 0,89 % the average run length is 1 000 batches, at 0,95 % it is 15 batches and for a process mean of 1,0 % the average run length drops to 5 batches. Thus, on average, a shift of mean to 1,0 % or more volatile matter would be detected within five batches. This was regarded as adequate in the light of the time for production of batches, and the possibility of blending batches of high volatile matter with others, provided they did not occur in larger numbers of excessive concentrations.

Table 9 : Batch % volatile matter

Batch n°	Batch % volatile (y)	y - T - K	Modified CUSUM S_i	Batch n°	Batch % volatile (y)	y - T - K	Modified CUSUM S_i
1	0,87	- 0,053	0,000	17	0,84	- 0,083	0,000
2	0,93	0,007	0,007	18	0,91	- 0,013	0,000
3	0,89	- 0,033	0,000	19	0,93	0,007	0,007
4	0,94	0,017	0,017	20	0,96	0,037	0,044
5	0,84	- 0,083	0,000	21	0,99	0,067	0,111
6	0,89	- 0,033	0,000	22	0,92	- 0,003	0,108
7	0,86	- 0,063	0,000	23	1,01	0,087	0,195
8	0,80	- 0,123	0,000	24	0,98	0,057	0,252
9	0,83	- 0,093	0,000	25	0,97	0,047	0,299
10	0,87	- 0,053	0,000	26	0,87	0,053	0,246
11	1,05	+ 0,127	0,127	27	0,93	0,007	0,253
12	0,87	- 0,053	0,074	28	0,94	0,017	0,270
13	0,94	0,017	0,091	29	0,96	0,037	0,307
14	0,89	- 0,033	0,058	30	0,91	- 0,013	0,294
15	0,92	- 0,003	0,055	31	0,98	0,057	0,351
16	0,77	- 0,153	0,000	32	0,95	0,027	0,378

CUSUM calculations use T + K = 0,923.

7.4 Example 4

7.4.1 Example 4 concerns the retrospective analysis of data. For this example, 100 values of % ash-content in deliveries of coal are presented. The recipient of the first 40 deliveries, who has concluded that the quality over this trial period is satisfactory, and using the data obtained, wishes to monitor subsequent deliveries against sustained shifts in mean level.

7.4.2 The data for 100 deliveries are shown in table 10. The mean of the first 40 is 7,03 % and the standard error (estimated from mean square differences) is 1,60. The recipient might reasonably adopt a target value of 7 % for a CUSUM chart, and it may be useful to plot the CUSUM for the trial period as well as for subsequent deliveries. The CUSUM chart for the whole series is shown in figure 17 but note that, for deliveries 1 to 40, the CUSUM forms part of a retrospective analysis of accumulated data, whereas the data for deliveries 41 to 100 would be plotted sequentially as they become available.

Table 10 : Data for ash-content in coal

(Read the following across rows)

6,63	5,58	3,36	4,79	7,39	13,46	8,13	7,49	8,49	7,03
8,97	8,48	9,99	5,54	7,29	5,72	4,64	8,52	7,94	6,55
5,31	5,75	6,05	6,07	5,40	7,79	8,69	7,46	4,93	8,95
5,53	4,55	5,50	7,66	9,56	7,71	6,78	6,40	7,73	5,32
6,78	9,24	8,05	7,88	6,79	7,93	4,44	9,13	7,92	4,78
11,96	7,27	11,29	6,72	11,76	4,99	9,47	7,65	7,63	6,99
8,84	9,09	3,20	6,07	7,61	7,44	12,65	6,65	5,39	10,58
8,13	5,42	7,48	8,71	8,88	5,83	3,91	12,07	5,18	5,89
9,43	5,54	4,64	8,59	6,35	10,47	7,21	12,28	6,90	8,41
11,08	5,28	15,71	7,65	7,61	7,57	3,80	6,11	6,74	9,90

7.4.3 The CUSUM chart reveals some striking features in the data, both in the "trial" sequence 1 to 40 and the "production" sequence 41 to 100. After a short run (deliveries 1 to 4) below 7 % there is a run above 7 % (deliveries 7 to 13, average 7,9 %) followed by a stable sequence until the end of the trial run, and continuing to delivery 50. There is then a general tendency for a rising CUSUM i.e. deliveries averaging more than 7 %. With short sequences of steep slope, longer runs of lesser slope. The sequence 86 to 93, averaging 10 % ash, is especially noteworthy. One also gains the impression of increased variability in the latter part of the series, and in fact the standard error of the last 60 observations is around 2,5 compared with 1,6 for the first 40 observations. Applying the span test to the sequence from 6 to 33, the maximum height above the chord joining 5 and 33 occurs at delivery 13 and is 15,8. The length of the sequence, m, is 28 and σ_e is 1,60. Using the nomogram in figure 12 the probability of exceeding observed value of $|V_{max.} / \sigma_e|$ is less than 0,001.

Appendix A

Measures of variation

A.1 Standard deviation and related measures

The standard deviation may be relevant either when used explicitly in connection with individual observations (in a weights and measures context, often termed "original values"), or when used implicitly to provide an appropriate measure of variation in sample statistics. The most obvious example is that of the standard error of the sample mean, but other cases arise in monitoring, for example, sample medians, ranges or standard deviations.

In either context, the individual observations are assumed to be random samples from an homogeneous segment of the process. Sometimes this assumption is obviously ill-founded and other appropriate measures of variation have then to be considered. In most cases it is at least advisable to check on the assumption by calculating alternative measures and noting any discrepancies, or by testing for serial dependences.

If a measure of the process standard deviation is required, there are several means of estimating it. Note the three most useful methods as follows :

a) For the complete set of N values, calculate the conventional estimate, s :

$$s = \sqrt{\frac{1}{N-1}\sum (y_i - \bar{y})^2} \qquad \text{(A.1)}$$

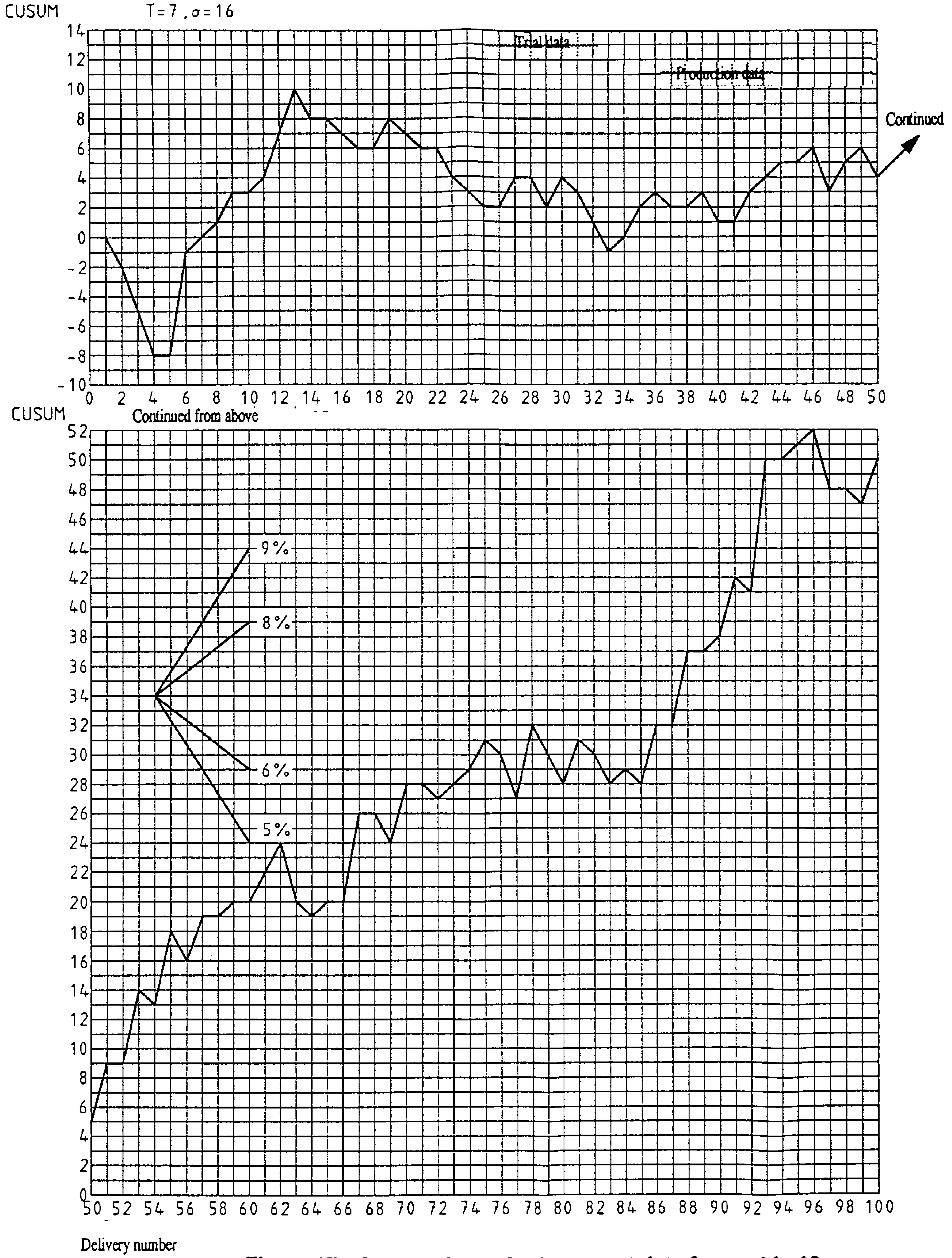


Figure 17 - Cusum chart of ash-content data from table 10

then $\hat{\sigma}$, the estimate of σ, = s. If the data to be plotted are individual observations, s now provides the basis for scaling the chart and drawing inferences about changes in average level.

b) Alternatively, the within-sample estimates of variance may be averaged over the m samples. This procedure is especially usefull if the samples are not of the same size (as may occur if items or test results are lost). Denoting the size of the jth sample by n_j and the variance estimate by s_j^2 the combined estimate is :

$$s^2 = \frac{\sum_{j=1}^{m}(n_j - 1)s_j^2}{\sum_{j=1}^{m}(n_j - 1)}$$

then $\hat{\sigma} = \sqrt{s^2}$ (A.2)

Where all sample sizes are identical, this simply becomes the average of the variances :

$$\hat{\sigma}^2 = \frac{1}{m}\sum s_j^2 \qquad (A.3)$$

c) Where the data are already formed into m samples of n observations, a rapid estimate of s (and hence of σ_e) may be obtained from the average of the ranges in these groups.

For average range $\overline{R}$, the estimate is given by :

$$\hat{\sigma} = \overline{R} / d_n, \text{ where } \overline{R} = \frac{1}{m}\sum_{j=1}^{m} R_j \qquad (A.4)$$

where d_n is the conversion factor for the appropriate sample size, from table 11.

Table 11 : Factors for converting mean range to an estimate of standard deviation $\hat{\sigma} = \overline{R} / d_n$

Sample size n	2	3	4	5	6	8	10
Factor, d_n	1,13	1,69	2,06	2,33	2,53	2,85	3,08

If the sample means are to be plotted, for sample size n use :

$$\hat{\sigma}_e = \hat{\sigma} / \sqrt{n} \qquad (A.5)$$

A.2 As noted in 2.6, the preceding estimates are often unsuitable because of the nature of the process or sampling procedure. In such cases, a direct estimate of the standard error of the sample statistic is preferable. Such sample statistics may be averages, proportions, medians, etc. Denoting them by y_j, then for m sample statistics the following is an estimate of the standard error of the sample statistic :

$$\hat{\sigma}^2 = \frac{1}{m-1}\sum_{j=1}^{m}\left(y_j - \bar{y}\right)^2 \text{ where } \bar{y} = \frac{1}{m}\sum_{j=1}^{m} y_j \qquad \text{(A.6)}$$

NOTE : The alternative measures based on between sample variability need to cover a number of samples. It is recommended that at least 20 samples should be included in any measure of between sample variability.

A.3 Even the above estimates may sometimes fail to yield satisfactory measures of the true variation. In cases where shifts in mean level occur in the sample data from which the standard error is to be estimated, these shifts will inflate the estimate, and a procedure based on successive differences (δ_j) between the sample values :

$$\delta_j = y_j - y_{j+1}$$

Then either :

a) square and sum these differences, and estimate :

$$\hat{\sigma} = \sqrt{\frac{1}{2(m-1)}\sum_{j=1}^{m-1}\delta_j^2} \qquad \text{(A.7)}$$

or

b) sum the absolute values of differences, calculate their average and use :

$$\sigma = \frac{8}{9}\bar{\delta} = \frac{8}{9}\left[\frac{\sum_{j=1}^{m-1}\left|\delta_j\right|}{m-1}\right] \qquad \text{(A.8)}$$

(The factor 8/9 is approximately 1/1,128, treating the successive differences as ranges in overlapping pairs of values and using the conversion factor d_2 = 1,128 to estimate the standard error from these ranges).

A.4 When setting up a CUSUM chart for data presentation purposes, the choice of procedure for measuring the variability may be made largely on grounds of the nature of the data, method of sampling and possible convenience of calculation. However, where statistical tests for change points or shift in level are to be preformed, care should be exercised in the selection, and the possibility of serial dependencies between successive values, or of cyclic phenomena, should be considered.

A useful general test for anomalies in the series of observations from which the standard error is estimated is as follows :

a) estimate σ_e using the formula (A.7) or (A.8) and denote this estimate by the symbol A ;

b) also estimate σ_e using the formulae (A.5) or (A.6) and denote this value by B ;

c) form the ratio $(A/B)^2$.

For a series of 20 or more observations from a normal population, this ratio is itself approximately normally distributed with mean 1 and standard error $1/\sqrt{(k+2)}$. Thus a simple test at the 5 % significance level is to check whether the ratio lies within the range :

$$1 \pm \sqrt{\frac{2}{k+2}}$$

Where the observed ratio (A/B) significantly exceeds 1, the implication is of negative serial correlation (alternation or over-control). A ratio significantly below 1 may occur with cycling, other forms of positive serial correlation (e.g. lag effects) or changes in mean level within the sequence of observations, whether regular or irregular step changes, drifting or trends.

Appendix B

Scaling the chart

B.1 The relationship between the CUSUM and sample number scales will determine the appearance of the chart. Too large a CUSUM scale will produce a chart with saw-tooth appearance on which small changes will be difficult to detect, while large shifts in average will quickly send the CUSUM path off the top or bottom of the graph sheet. Too small a scale will effectively damp out minor variations but may also suppress real changes in the mean level of the data.

A widely accepted convention is that the distance on the vertical (CUSUM) scale equal to one sample interval should represent approximately $2\sigma_e$ units of the plotted variable. This convention may be applied in one of three ways.

B.2 Decide on a convenient interval for the i-scale (sample number axis). This may be, for example, 5 mm (or 0,2 in) for a chart intended for "desk" use, or 1 cm, 2 cm (or 1/2 in, 1in) for a chart designed for wall display.

Mark off this ample interval on the CUSUM scale, and let it represent approximately $2\sigma_e$ units of the CUSUM values to be plotted. For graph grids divided on a decimal system, it is advantageous to round the value of $2\sigma_e$ to a convenient number such as 1,0, 2,0, 2,5, 4,0, 5,0 or some multiple thereof. If $2\sigma_e$ happens to lie midway between two convenient values, it is preferable to round down. Thus if $2\sigma_e$ then with a scale of 1 cm per sample on the i-scale, use 1 cm = 5 units on the CUSUM scale (rather than 1 cm = 10 units).

B.3 Conversely, decide on a convenient CUSUM scale, and then adopt a distance equivalent to $2\sigma_e$ as the i-scale length for one sample interval.

This has the minor advantage of preserving the exact $2\sigma_e$ relationship (so that, for example, a change from a horizontal CUSUM path to one at 45° represents a shift of $2\sigma_e$ in mean level). Some decision procedures for CUSUMS may be interpreted in terms of angles (or their tangents) when this scale convention is adopted.

B.4 Where identical scale factors and equality of sample-interval scale are desirable for comparison between charts of two or more variables on the same sequence of items, a standardized chart may be useful. In this case, instead of the CUSUM $C = \sum(y_r - T)$, the deviations from target are divided by the standard error, i.e. :

$$C = \sum_{r=1}^{i} \frac{(y_r - T)}{\sigma_e} \qquad \text{(B.1)}$$

This procedure sets the standard error of the plotted variable to unity, and one unit on the i-scale corresponds to two units on the CUSUM scale. Different charts may then be compared by alignment or superimposing and decision rules can be implemented directly from the standardized parameters detailed in the appendix, tables and charts, without the necessity of adapting them to any special convention.

Against this must be set the disadvantage of losing direct contact with the actual variable involved. The additional standardization operation may be a further objection in some circumstances.

B.5 Other considerations may affect the choice of scale factor, and the special case of counted data (where prominence must obviously be given to integer values) is considered in detail in B.6. The mode of rounding may affect the choice of both scale factor and graph grid, e.g. measurements in hours, months, inches, etc. may be more readily plotted on a grid divided in twelfths than one based on decimal divisions.

B.6 Binary data, or counts of events containing frequent zeros

Whilst CUSUM methods are appropriate to counted (as well as measured) data, problems of scaling may arise when the counts are confined to 0's and 1's (binary events, such as failure = 0, success = 1) or comprise frequent zero counts (as in inspecting samples for nonconforming units when quality is good). These problems are readily overcome by suitable choice of the interval between sample points on the horizontal axis of the CUSUM chart.

An approximate indication is required of the expected proportion, p, of successes or failures (which-ever are fewer) scoring 1. In the case of event counts, the average number of events (e.g. nonconformities or faults) per sample should be estimated ; if this exceeds one, no special scaling is necessary, and the usual $2\sigma_e$ convention may be adopted, where σ_e now becomes the standard deviation of the number of events per sample.

For binary counts, or event counts averaging less than one per sample, proceed as follows :

a) calculate the average number of samples required to yield one binary score (1) or one event ;

b) round this value up to a convenient integer for plotting, and adopt it as the horizontal interval for the CUSUM chart. Note that for small p or low frequency of events, the interval between individual sample points will be very small : a reminder that individual samples contain little information in these circumstances ;

c) mark off the vertical (CUSUM) scale in intervals of the same length as the horizontal scale, and label as consecutive even integers upward and downward from zero ;

d) the CUSUM is now plotted as :

$C = \sum(y - p)$ for binary data where p is the frequency of 1-scores ;

or

$C = \sum(y - \mu)$ for small event counts, where μ is the mean number of events per sample.

The formulation of decision criteria requires a more advanced level of statistical knowledge than is assumed here and is therefore deferred to other standards. The use of conventional tests of significance (based on appropriate distributional assumptions) remains valid, however, subject to the modifications noted in 5.1.3 and appendix G, and, even without decision procedures, the CUSUM chart is an effective means of presenting binary or "rare event" data.

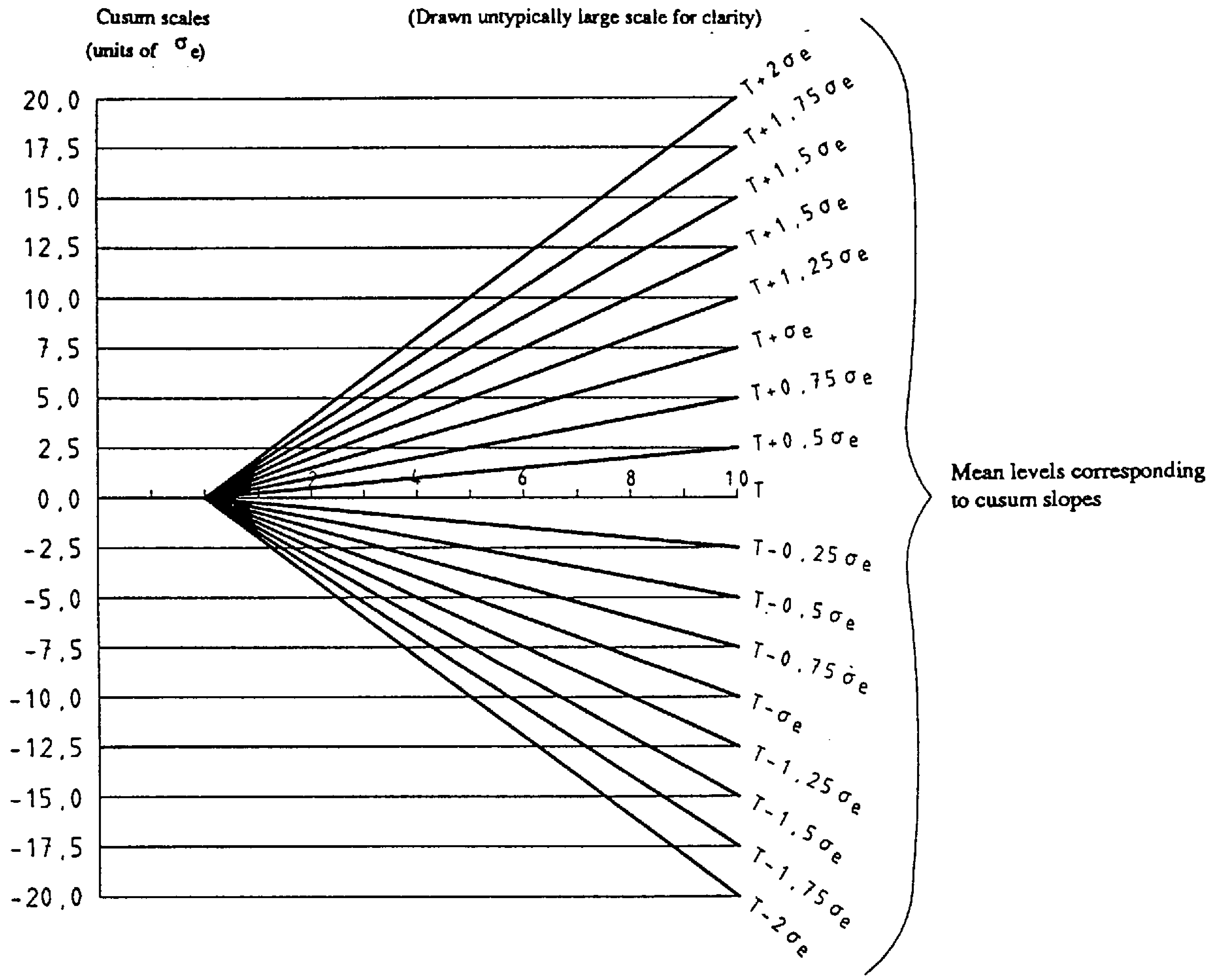


Figure 18 : Construction of CUSUM protractor (slope guide)

Appendix C

Calculation of local averages

C.1 One of the principle virtues of the CUSUM presentation is the clarity with which changes in average level are seen as changes in the CUSUM slope, and the facility the CUSUM technique offers for calculation of the local averages on either side of a suspected change point. This is achieved without recourse to recalculation from the original data.

C.2 A simple method of assessing approximate local averages is to incorporate a slope guide or "protractor" on the chart. Any inclination of the CUSUM path may then be compared with the protractor, and a useful indication of the average level thus obtained.

The protractor (see figure 18) comprises a pencil of lines, each drawn at an angle corresponding to the CUSUM inclination for a hypothetical average level. To construct the lines, it is only necessary to recognize that, for any mean level $\bar{y}$, the CUSUM path will have an average slope of $\bar{y}$ - T CUSUM units per sample interval. The choice of useful values of $\bar{y}$ will, of course, depend on the application, but frequently a range covering T ± $2\sigma_e$ in steps between $0{,}25\sigma_e$ and $0{,}5\sigma_e$ will be found suitable.

To construct the protractor, a horizontal line (i.e. parallel to the i-axis) is drawn to represent an average level T. The line should be of length corresponding to ten sample intervals.

Now for each chosen $\bar{y}$, mark off a pair of points 10 ($\bar{y}$ - T) CUSUM units above and below the right-hand end of the baseline. Join each point to the left-hand end of the baseline to complete the pencil of lines.

In using the protractor to estimate local averages, a parallel rule, roller rule or a piece of transparent material (e.g. clear plastics or acetate sheet as used for overhead projector transparencies) ruled with a set of parallel lines will be useful. Alternatively, the protractor itself may be drawn on transparent material, and used by superimposing on the CUSUM chart over the segment of interest.

C.3 A more precise estimate of a local average can be obtained from the numerical values of the CUSUM at two points in the sequence. These values may be read from the chart or taken directly from the CUSUM calculations if these are available.

The first step is to define the segment whose average is required. Sometimes external considerations will dictate this, e.g. the average for a week or month may be needed for reporting purposes, or some change in conditions in an experiment may indicate the appropriate segment. In other cases, inspection of the CUSUM chart may reveal suspected change points as changes in the CUSUM slope (the point at which such a change occurs on a CUSUM chart is termed a "corner").

The first and last sample numbers in the required segment are defined as i+1 and j. Thus sample number i is the last point in the previous segment (or point 0 if the first segment on the chart is being dealt with). Note, however, that any change in CUSUM from C_i and C_{i+1} is due to the contribution of y_{i+1}, the first point in the segment under consideration. The two CUSUM values required are thus C_i and C_j.

C.4 The average departure from target over the segment is then the average contribution of the sample values to the CUSUM and is simply : change in CUSUM from C_i to C_j divided by the number of values in segment.

The actual average level for the segment is then obtained by adding the target value to the average CUSUM contribution, giving :

$$\bar{y}_{i+1,y} = T + \frac{C_j - C_i}{j - i}$$

It is apparent that the CUSUM contains all the information required for this estimate, and there is no necessity to refer back to the original data unless any check on anomalous values is indicated by the appearance of the CUSUM chart.

Appendix D

Replotting the CUSUM chart

D.1 Circumstances occur when it may be desirable to replot a CUSUM chart, in part or in whole. Replotting the whole chart (for example, because a change of scale is desirable) presents no new problems. However, sometimes the CUSUM may, because of a shift in the average level (or a wrong choice of target), tend to run off the graph at the top or bottom. In this case, only part of the chart may need to be replotted. Two situations occur.

D.2 A change in mean occurs, taking the CUSUM off the graph. No change in target value is required, perhaps because some specification value has been adopted. In this case, a new CUSUM scale, displaced by an appropriate number of units, can be drawn at any convenient point and the CUSUM replotted. An alternative (and equivalent) procedure is merely to reduce the CUSUM by an appropriate amount to bring it back to a convenient location on the chart.

Provided that the chart is labelled to indicated this adjustment, no problems of interpretation arise, due to the fact that the magnitude of the CUSUM has no absolute importance, the inferences being drawn from changes in CUSUM and slopes on the chart.

D.3 Where a change in average has occurred, and is expected to persist, it may be appropriate to adjust the target value to match (approximately) the new mean level. This procedure will provide a maximum sensitivity for detection of any further change. Obviously the chart should be clearly labelled to indicate the change of target, as interpretation of slopes will be affected by the adjustment.

Appendix E

The full V-mask

E.1 Decision rules may also be applied by means of a mask based on a complete V. This has the same basic properties as the truncated mask, but the arms are brought to a vertex (0 in figure 19). This means that there is no longer a datum point, and the vertex should be positioned on the chart so that it is a distance OA ahead (i.e. to the right) of the latest point of interest. For obvious reasons OA is known as the lead distance.

E.2 This form of mask was originally described in connection with charts scaled strictly on the $2\sigma_e$ rule, and the half angle ϕ (or its tangent, f) is then a useful characteristic. The slope of the decision line (as for the truncated mask) is $0{,}5\sigma_e$ per sample interval, and with scale factor $2\sigma_e$, the tangent of ϕ becomes 0,5/2 = 0,25, corresponding to an angle of 14°. Then the lead distance d, tan ϕ and decision interval H are related by :

$$\frac{H}{\sigma_e} = 2d\ \tan\phi \text{ or (where } h = H/\sigma_e\text{) and a = 2} \qquad (E.1)$$

giving in this case d = 5/(2 x 0,25) = 10 sample intervals. For any other scale convention, a, where a σ_e units on the CUSUM scale correspond to one sample interval :

$$h = \frac{H}{\sigma_e} = ad\ \tan\phi,\ \text{or } d = \frac{h}{f} \qquad (E.2)$$

The simplest way of constructing the mask is to draw the truncated form as figure 19 and extend the arms to meet at 0. Alternatively, the lead distance and decision interval (d, H) may be calculated from (E.1) or (E.2) and the construction of figure 20 adopted.

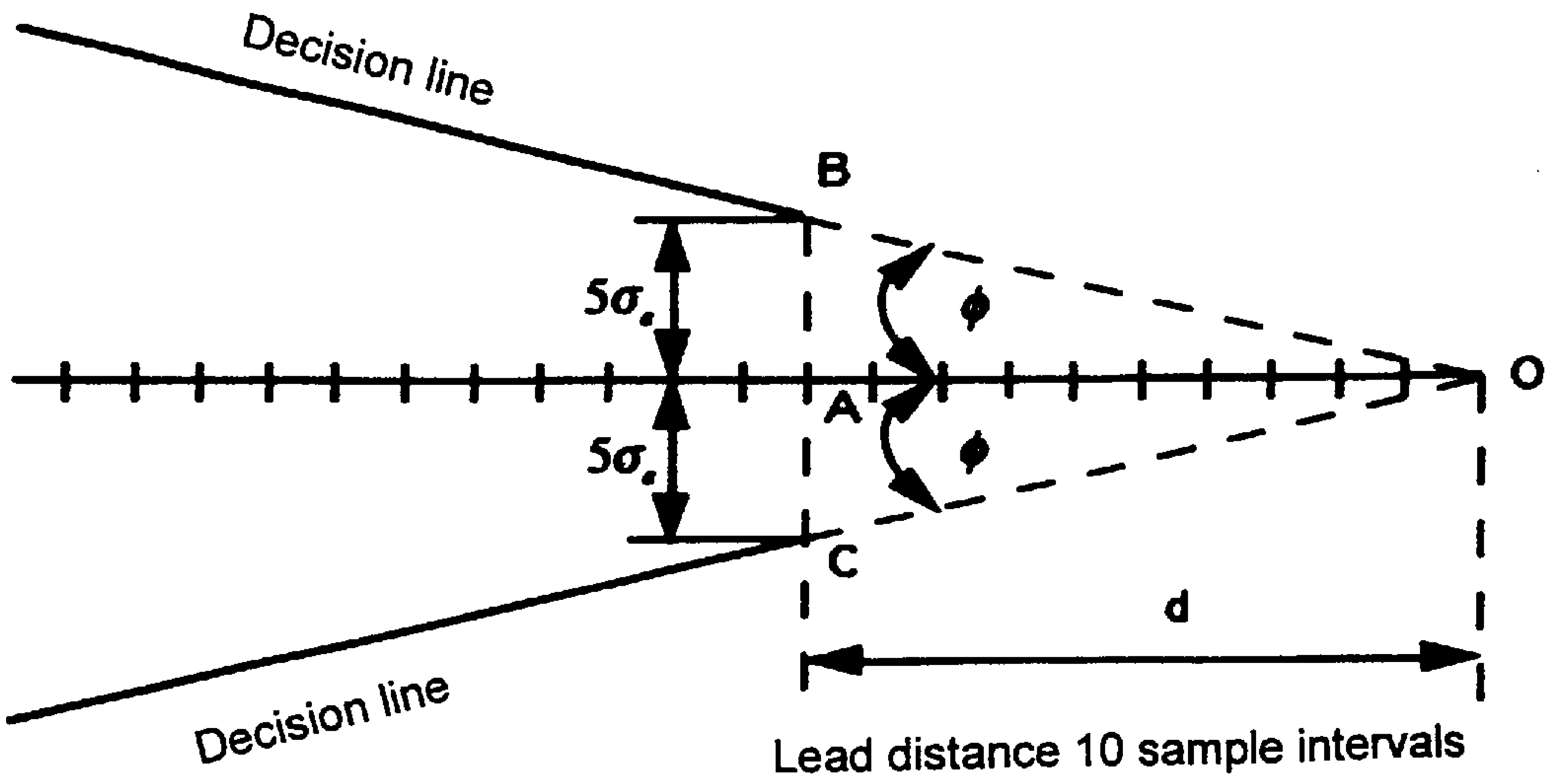
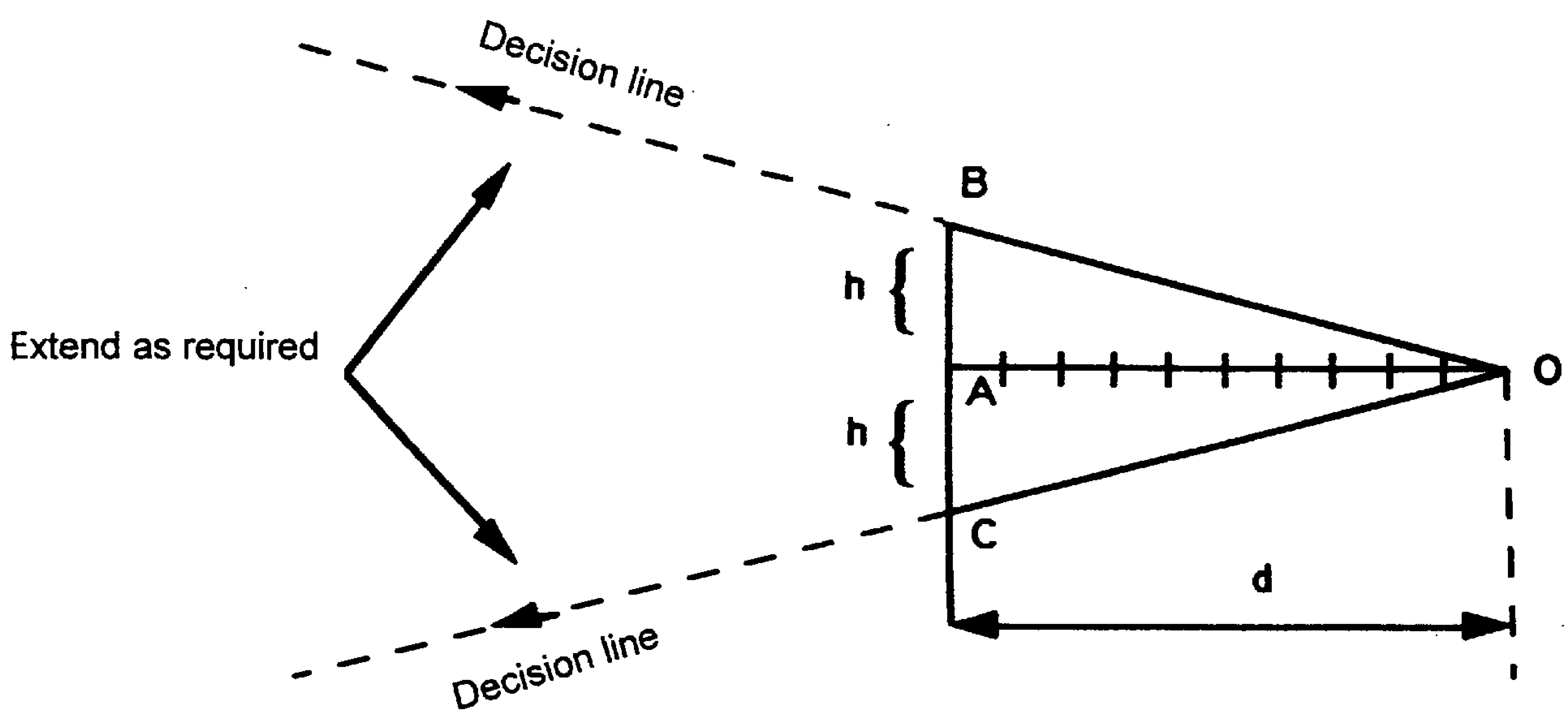


Figure 19 : Full CUSUM V-mask



For decision rule of 4,22, $h = 5\sigma_e$, d = 10 samples

Figure 20 : Alternative method of constructing full V-mask

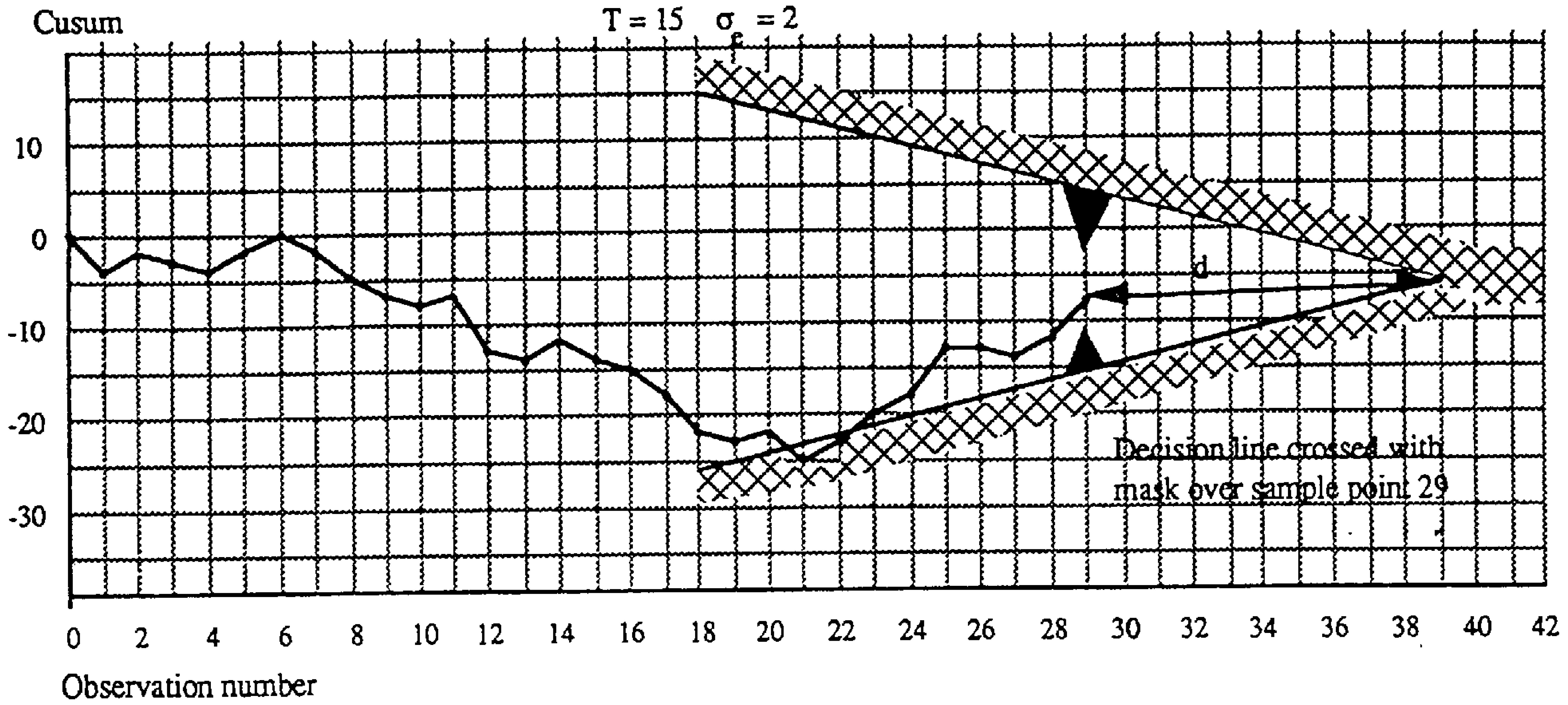


Figure 21 : Full V-mask applied to CUSUM chart (T = 15, σ_e = 2)

E.3 In practice, the full v-mask will yield identical decisions to those signalled by the truncated form. Therefore only one example of the application of the full mask is shown. This is at sample 29 on the chart for the data of table 1, and is shown in figure 21.

E.4 As for the truncated mask, some features may be incorporated to aid in using the full-V form. If drawn on transparent material, the datum point may be indicated to avoid counting out sample intervals for the lead distance. Where a cut-out mask of opaque material is used, indicator marks may be included to assist in aligning the mask correctly (as shown in figure 21). If the solid mask is preferred, a small hole may be located at the datum point, and the mask placed so that the sample point of interest appears in the aperture.

Appendix F

Local decision lines

F.1 Two further methods of implementing the decision rule are available. One does not involve the use of charts, being applied directly to a CUSUM tabulation. This procedure is described in 6.1. The other method is to construct local decision lines on the chart whenever a possible change is suspected. This has two advantages : firstly, there is no need to prepare a permanent mask, and secondly the decision line may be drawn when a CUSUM slope suggests a possible shift in average level. The decision is not, of course, actually signalled until the decision line is touched or crossed, but as the CUSUM nears the decision line appropriate preparations for action may be initiated. On the other hand, there is the risk that over-enthusiastic construction of decision lines may add unnecessary clutter to the chart.

F.2 The method uses the same constructions as for the full and truncated masks, but it is necessary to inspect the CUSUM, and locate a suspected point of change. This will generally be apparent, often corresponding to a local maximum or minimum on the chart (as in figure 2, where points 6 and 7 are a local maximum and point 21 is a local minimum).

Where (as for points 6 and 7) it is not clear which of two points to use, the general rule is that the line should be constructed with its datum at point i-1, where C_i differs from C_{i-1} by at least $2f\sigma_e$. For the present decision rule f = 0,25, so that y_i should differ by at least half a standard error from y_{i-1}, in the direction of the suspected change. In the case of samples 6 and 7 of table 1 and figure 2, we have (with $\sigma_e = 2$) :

$$C_7 = -1,\ C_6 = 0,\ (C_7 - C_6)/\sigma_e = -1/2$$

Thus the slight downward slope from C_7 to C_6 is just sufficient to justify basing the decision line on sample point 6.

F.3 Having located a likely point of change, 0, a vertical line OA of length $5\sigma_e$ is drawn in the same sense (upward or downward) as the suspected shift in average. A horizontal line OD is also drawn from the suspected change-point towards the right. At C, 10 samples to the right of O, a second vertical line BC is drawn from the horizontal base, of length $10\sigma_e$. The ends of the two vertical lines are joined to produce the decision line AB, extended to the right as required. The construction is illustrated in figure 22.

In more general terms, we have :

$OA = H = h\,\sigma_e$

$OC = d$

$BC = 2H = 2h\,\sigma_e$

Slope of $BC = \tan\phi = f = \dfrac{h}{ad} = \dfrac{H}{2ad\sigma_e}$

The construction shown is for testing an apparent upward change in mean. For a suspected downward change, *OA*, *CB* are drawn downward from the horizontal base line.

For the rule in 4.2, $H = 5\,\sigma_e$, $d = 10$.

F.4 The application of a local decision line is illustrated in figure 23, again using the data of table 1. A line erected on sample point 6 yields a signal at sample 18, the same point at which the V-masks signal the downward shift from target.

Appendix G

Distribution-free CUSUM tests

G.1 In order to apply a distribution-free CUSUM test, it is necessary to locate the median value of the segment under observation. Scores of +1 or -1 are then given to observations respectively greater of less than the median (any values exactly equal to the median scoring 0). A CUSUM chart, or merely a cumulative total, of the ordered scores is then formed, and the maximum absolute value noted. For values of $m \geq 10$, critical levels for the score (under the null hypothesis of no change in level within the segment) may be obtained from table 12.

G.2 Again using the same data :

a) $i = 7$, $r = 21$, $j = 24$, $m = 17$

Median of sequence 8 to 24 is 14. The scores and cumulatives are then as shown in table 13.

Table 12 : Percentage points for CUSUM score (+1, -1) about median (distribution-free test)

Significance level 1)	Critical level is next integer above $f\sqrt{m}$	$m \geq 10$
0,10	$f = 1,22$	
0,05	$f = 1,36$	
0,025	$f = 1,48$	
0,01	$f = 1,63$	
0,005	$f = 1,73$	
0,001	$f = 1,95$	

1) For other values of α, use $f = 0,697 + 0,403\ \mu_\alpha$, where μ_α is the standard normal deviate with one tail probability α.

Table 13 : Scores and cumulatives of sequence 8 to 24

Sample	Score	Cumulative
8	-1	-1
9	-1	-2
10	0	-2
11	1	-1
12	-1	-2
13	0	-2
14	1	-1
15	-1	-2
16	0	-2
17	-1	-3
18	-1	-4
19	-1	-5
20	1	-4
21	1	-3
22	1	-2
23	1	-1
24	1	0

For m = 17, the critical (absolute) score for α = 0,05 is 1,36 $\sqrt{17}$ = 5,6, and the observed score does not reach the next higher integer, 6. Using the distribution-free procedure, the change cannot yet be judged significant .

b) i = 7, r = 21, j = 31, m = 24

The median now becomes 14,5 (the average of the 12th and 13th ordered values in the segment). The new table of scores is as shown in table 14.

Table 14 : Scores and cumulatives of sequence 8 to 31

Sample	Score	Cumulative	Sample	Score	Cumulative
8	-1	-1	20	+1	-7
9	-1	-2	21	-1	-8
10	-1	-3	22	+1	-7
11	+1	-2	23	+1	-6
12	-1	-3	24	+1	-5
13	-1	-4	25	+1	-4
14	+1	-3	26	+1	-3
15	-1	-4	27	-1	-4
16	-1	-5	28	+1	-3
17	-1	-6	29	+1	-2
18	-1	-7	30	+1	-1
19	-1	-8	31	+1	0

The critical value is now the next integer above 1,36 $\sqrt{24}$ = 6,66, i.e. 7. This is reached by the (absolute) scores at samples 18 to 23, indicating a significant change within the segment.

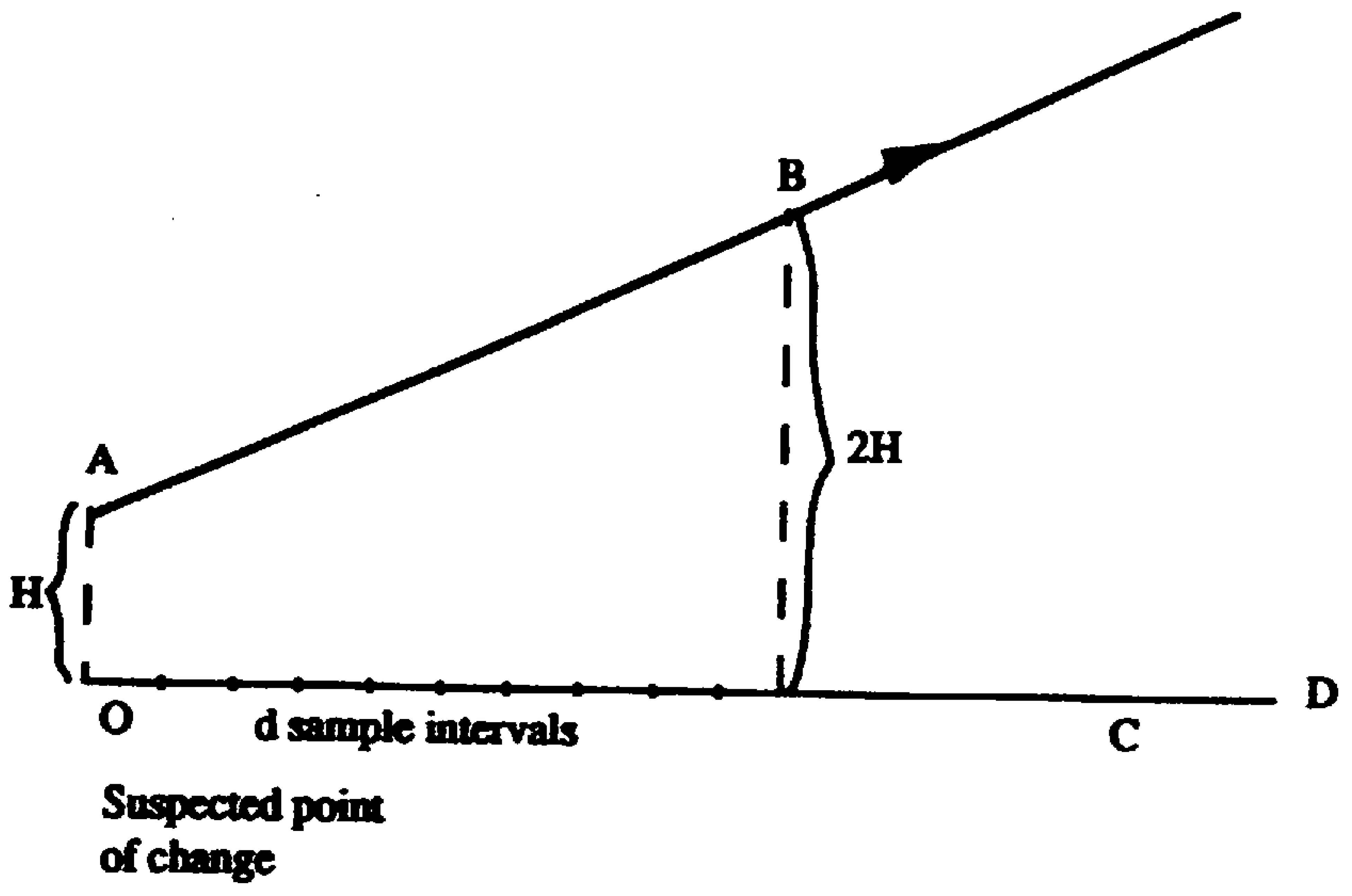


Figure 22 : Construction of local decision line

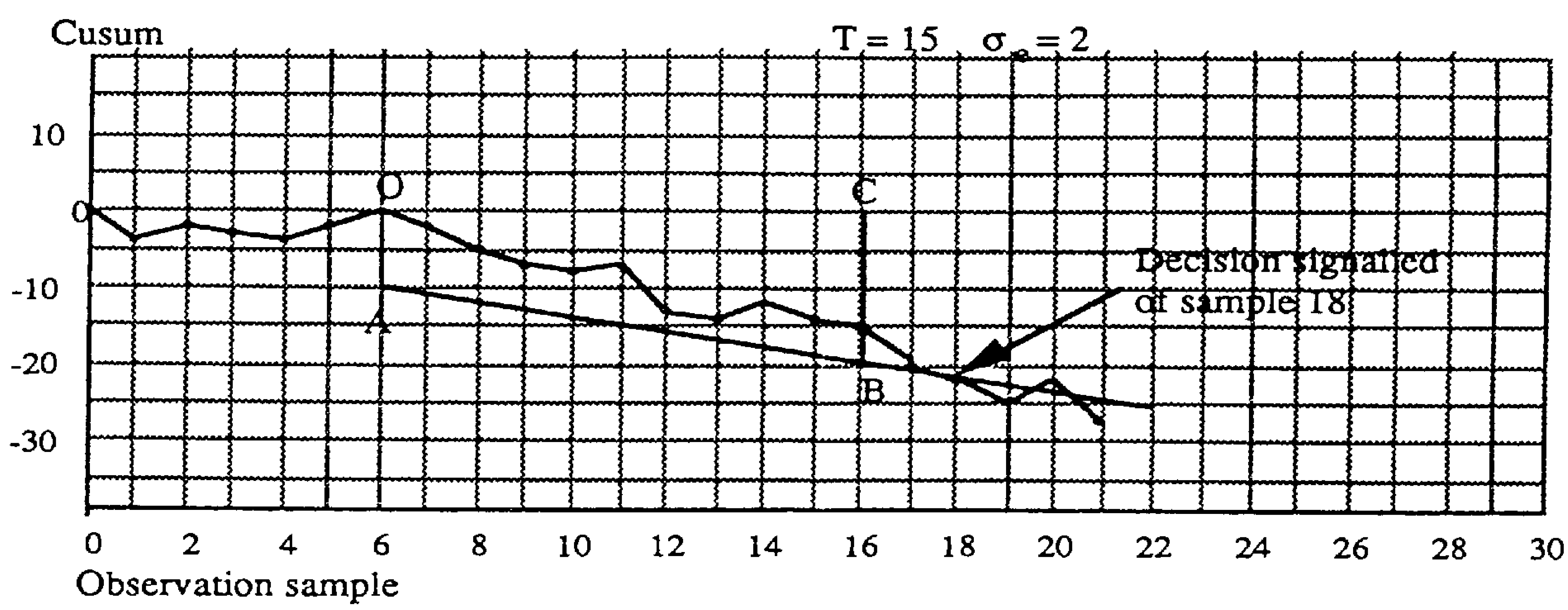


Figure 23 : Decision line constructed on CUSUM chart (T = 15, σ_e = 2)

Appendix H

Alternative standard schemes

H.1 The schemes presented in this subclause extend the range of decision rules in two ways :

a) an option for shorter ARL at Target level is provided, and permits the more rapid detection of undesirable conditions at the expense of a higher rate of false alarms ;

b) where shifts of a particular magnitude are of especial interest, the most suitable of three alternatives may be adopted to provide near-optimum efficiency in detection.

H.2 Schemes having ARL's at AQL of around 700 to 1 000 samples are here referred to as C1 schemes ; those with shorter ARL's at AQL, around 140 to 200, are denoted C2 schemes.

Within each of these groups, three subdivisions are made on the basis of the magnitude of shift which is of greatest interest or importance. These subdivisions are :

a) critical shift less than 0,75 σ_e ;

b) critical shift 0,75 σ_e to 1,5 σ_e ;

c) critical shift greater than 1,5 σ_e.

It will be noted from the ARL data provided in Table 16, that there is little difference in ARL behavior of neighboring schemes at or near the boundaries of the subdivisions. Thus if the critical shift is around 0,7 to 0,8 standard errors, wither scheme a) or scheme b) may be adopted.

H.3 The CUSUM decision rule parameters (h, f) are listed in table 15. Average run length data are provided in table 16. Data for two Shewhart procedures are also listed :

- S1 : Conventional Shewhart chart with Action Line at T + 3,09 σ_e (or T - 3,09 σ_e) and warning line at T +1,96 σ_e (or T - 1,96 σ_e). These are the usual 1 in 1 000 and 1 in 40 limits, and for one-way control yield an ARL of 640 samples at the Target value. This is roughly comparable with the 700 to 1 000 ARL for the C1 procedures.

- S2 : Limits at displacements of 2,65 σ_e and 1,65 σ_e providing 1 in 250 and 1 in 20 probabilities at Target level, and a corresponding ARL of 167. This is compatible with the ARL's at AQL of the C2 procedures.

Table 15 : CUSUM parameters for alternative decision rules

Critical from target,	C1 schemes		C2 Schemes	
in units of σ_e	h	f	h	f
< 0,75	8	0,25	5	0,25
0,75 - 1,5	5	0,5	3,5	0,5
>1,5	2,5	1,0	1,8	1,0

NOTE : Boxed parameters are those for the basic rule, C1(b). Corresponding Shewhart type chart parameters are :

	Action line	Warning line
S1	T ± 3,09 σ_e	T ± 1,96 σ_e
S2	T ± 2,65 σ_e	T ± 1,65 σ_e

Table 16 : Average run length data and alternative decision rules

Shift from Target in units of σ_e	C1 Schemes			S1	C2 Schemes			S2
	a	b	c		a	b	c	
0	730	930	715	640	140	200	170	167
0,25	85	140	205	265	38	55	68	76
0,5	29	38	68	11	17	22	30	38
0,75	16,4	17	27	52	10,5	11,5	15	20
1,0	11,4	10,5	13,4	26	7,4	7,4	8,8	11,7
1,5	7,1	5,8	5,4	8,9	4,7	4,3	4,0	5,0
2,0	5,2	4,1	3,25	4,15	3,5	3,0	2,5	2,8
2,5	4,2	3,2	2,3	2,46	2,8	2,4	1,9	1,86
3,0	3,5	2,6	1,85	1,75	2,4	2,0	1,5	1,44
4,0	2,6	1,9	1,32	1,19	1,9	1,5	1,12	1,09

NOTE : Boxed and arrows indicate range of shifts for which each scheme is most effective.

ICS 03.120.30

Descriptors: statistical analysis, quality control, statistical quality control, data representation, graphic methods, charts, general conditions.

Price based on 40 pages