

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*

This material is reproduced from ISO documents under International Organization for Standardization (ISO) Copyright License number IHS/ICC/1996. Not for resale. No part of these ISO documents may be reproduced in any form, electronic retrieval system or otherwise, except as allowed in the copyright law of the country of use, or with the prior written consent of ISO (Case postale 56, 1211 Geneva 20, Switzerland, Fax +41 22 734 10 79), IHS or the ISO Licensor's members.



Reference number
ISO/TR 7871:1997(E)

ISO/TR 7871:1997(E)

Contents

0 Introduction iv

1 Scope and general principles 1

2 Preparations for cusum charting 2

3 Presentation..... 7

4 Decision rules for monitoring and control 8

5 Decision rules for retrospective analysis 22

6 Computational CUSUM techniques.....25

7 Examples of applications.....30

Appendix A Measures of variation.....37

Appendix B Scaling the chart.....42

Appendix C Calculation of local averages 45

Appendix D Replotting the CUSUM chart.....47

Appendix E The full V-mask 48

Appendix F Local decision lines 51

Appendix G Distribution-free CUSUM tests 53

Appendix H Alternative standard schemes.....56

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.

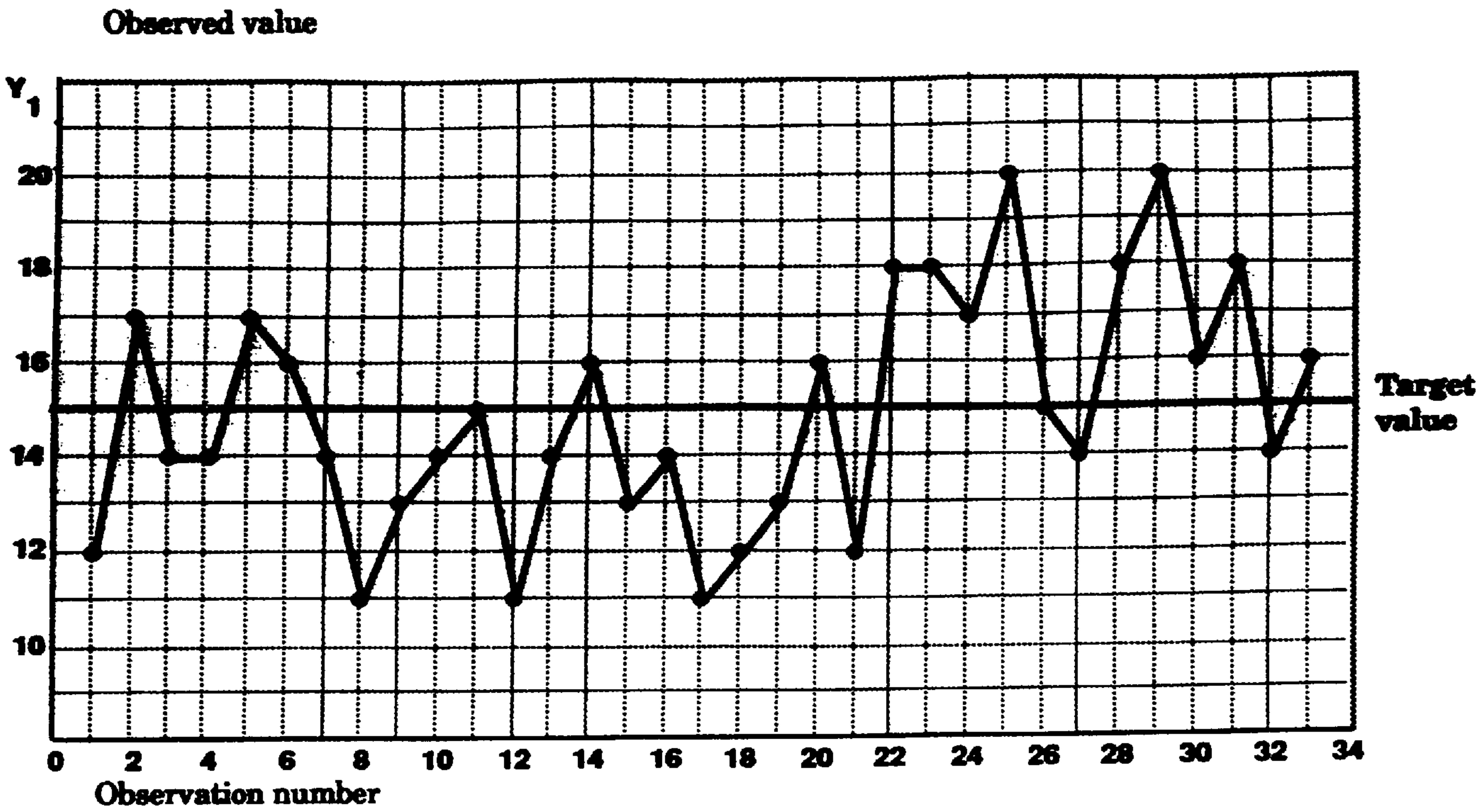


Figure 1 : Conventional chart of data from table 1

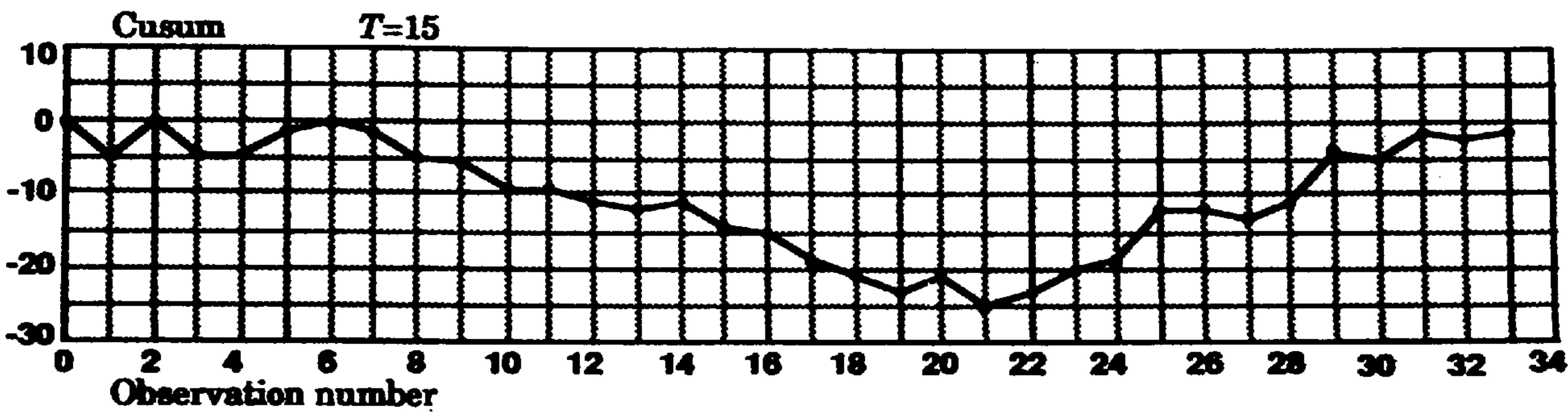


Figure 2 : Cusum chart of data from table 1

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 an inappropriate target value may result in the chart running off the upper or 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.

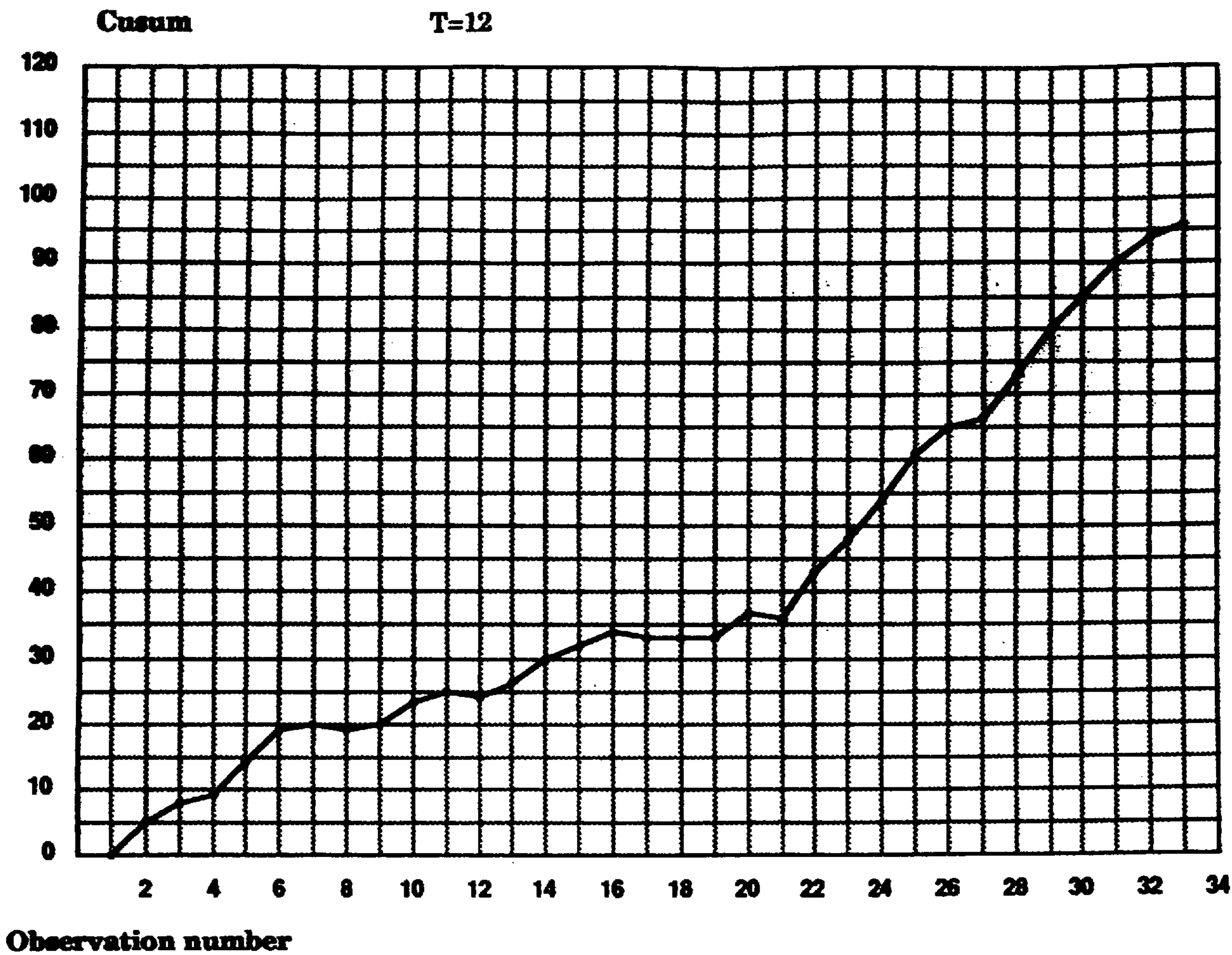


Figure 3 : Cusum chart of data from table 1, with reference value 12

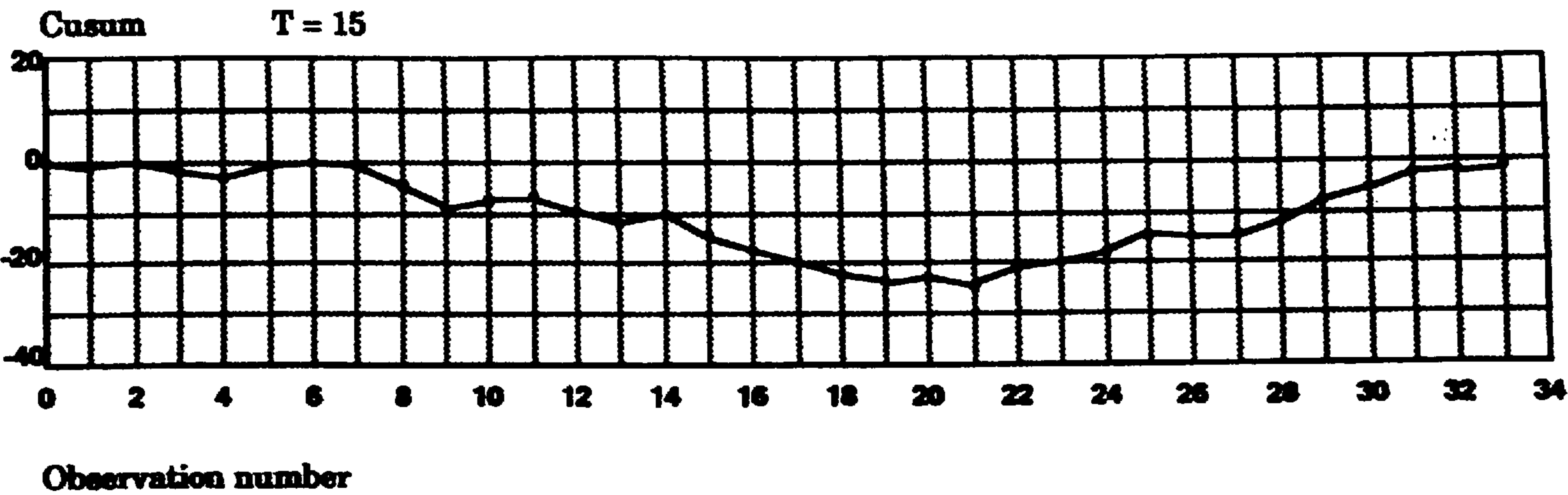


Figure 4 : Cusum chart of data from table 1, with reference value 15 but compressed cusum scale

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 :

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

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 multiples of the standard error of the plotted values (σ_e), and this standardized scale factor will carry the α (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 :

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

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 from 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 :

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

where :

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

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 $\bar{y}$, $\sigma_{\bar{y}} = \sigma / \sqrt{n}$, 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 :

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

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

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

$$\sigma = \sqrt{np_0 q_0} \quad (4)$$

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 ($\delta_j = y_j - y_{j-1}$) 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.

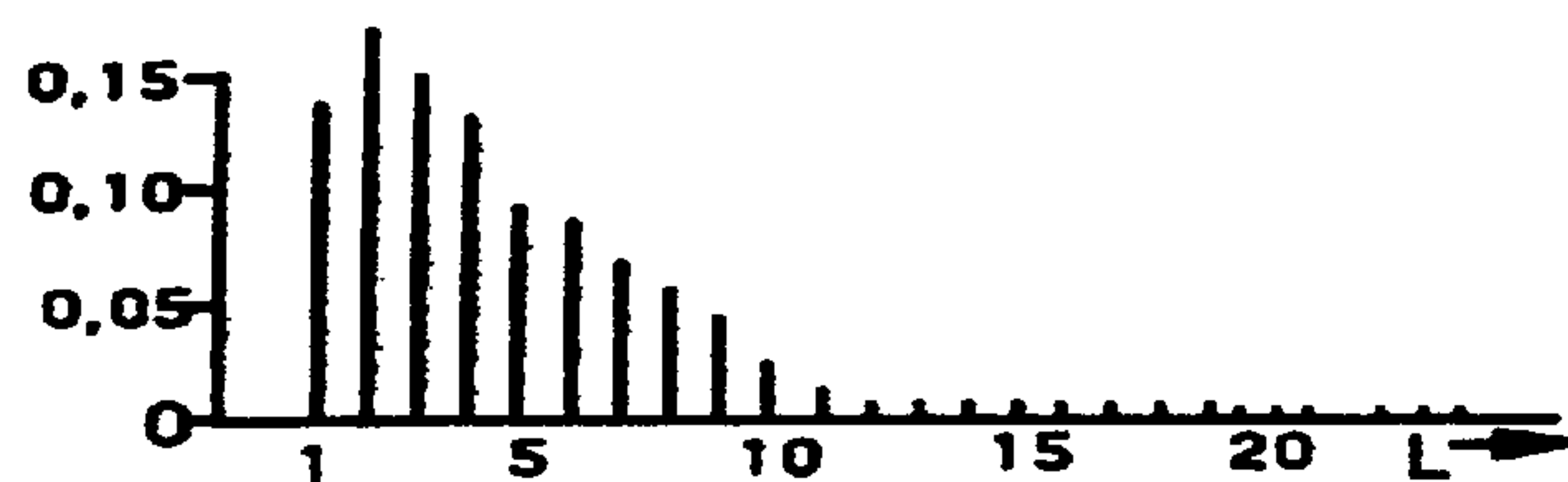
It should be noted that the run length itself is subject to statistical variation. Sometimes one may be "lucky" in obtaining no false alarms over a long run, or detecting a change very quickly. Sometimes an "unfortunate" run of samples may generate false alarms, or mask a real change so that it does 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

$\sigma_L = 4$

Median = 3,5

Mode = 2

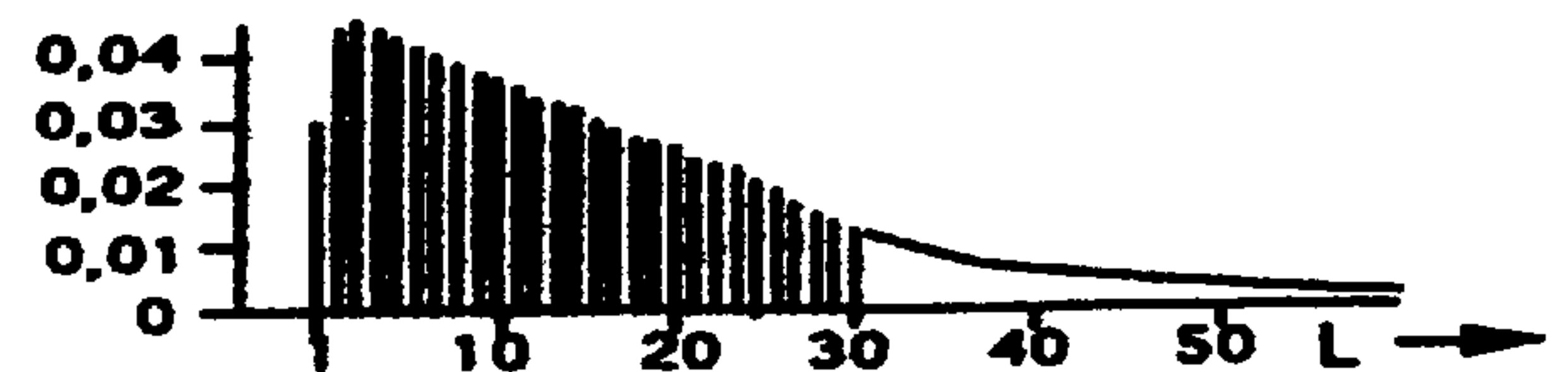
2,5 % of runs ≥ 16 1 % of runs ≥ 19 , etc.

b) ARL = 20

$\sigma_L = 18,9$

Median = 14

Mode = 3

12,5 % of runs ≥ 41 7,5 % of runs ≥ 50 , etc.

c) ARL = 96,5

$\sigma_L = 95,6$

Median = 67

Mode = 4

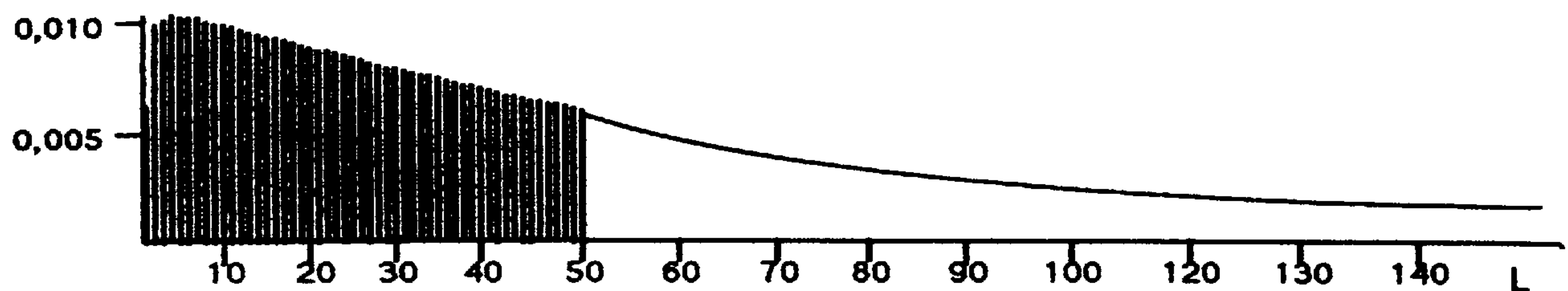
12,5 % of runs ≥ 200 7,5 % of runs ≥ 250 , etc.

Figure 5 : Distributions of characteristic run lengths

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.

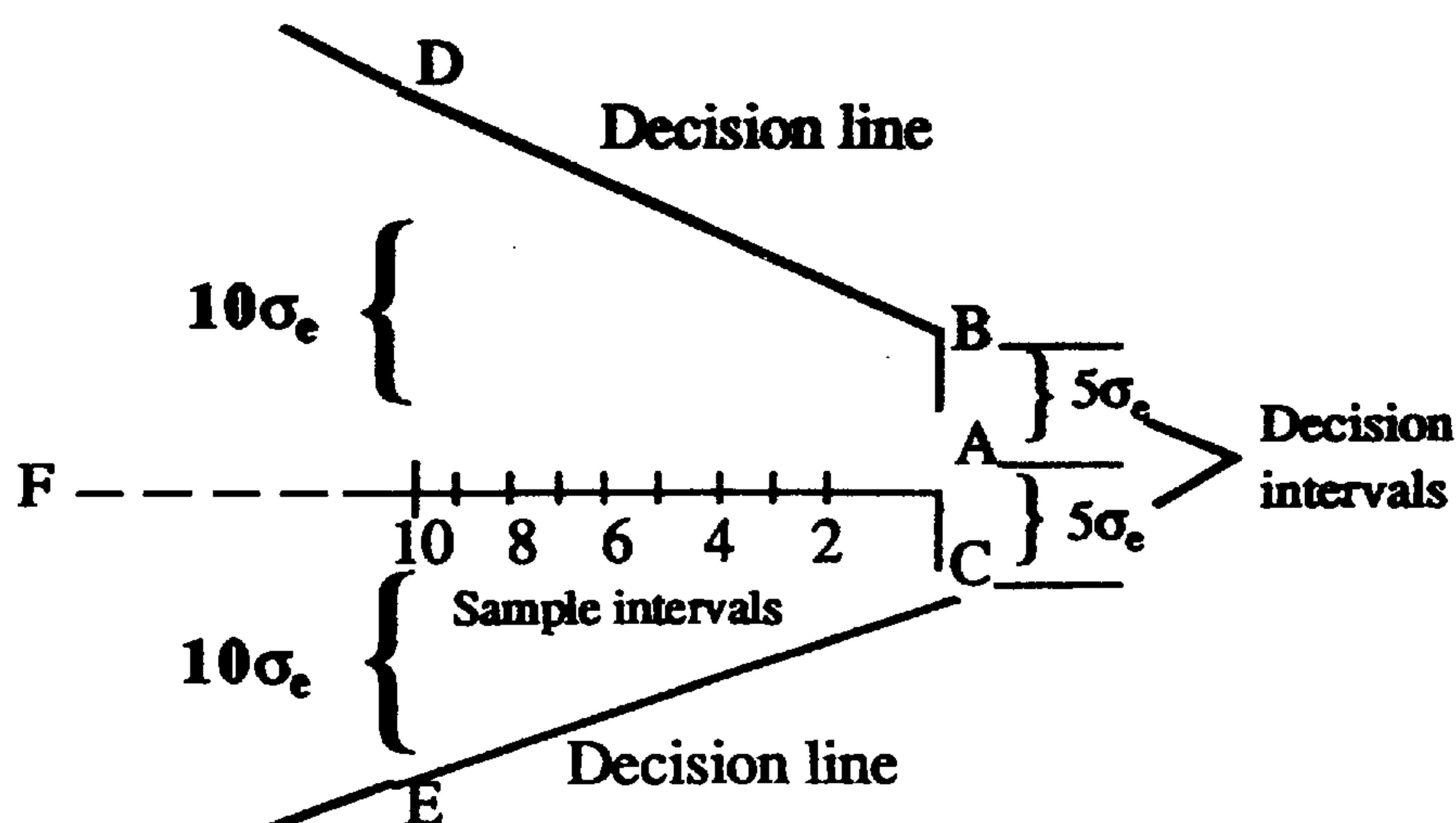


Figure 6 - General-purpose truncated V-mask

Several points are worth noting here :

- 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 ;
- 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 ;
- 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.

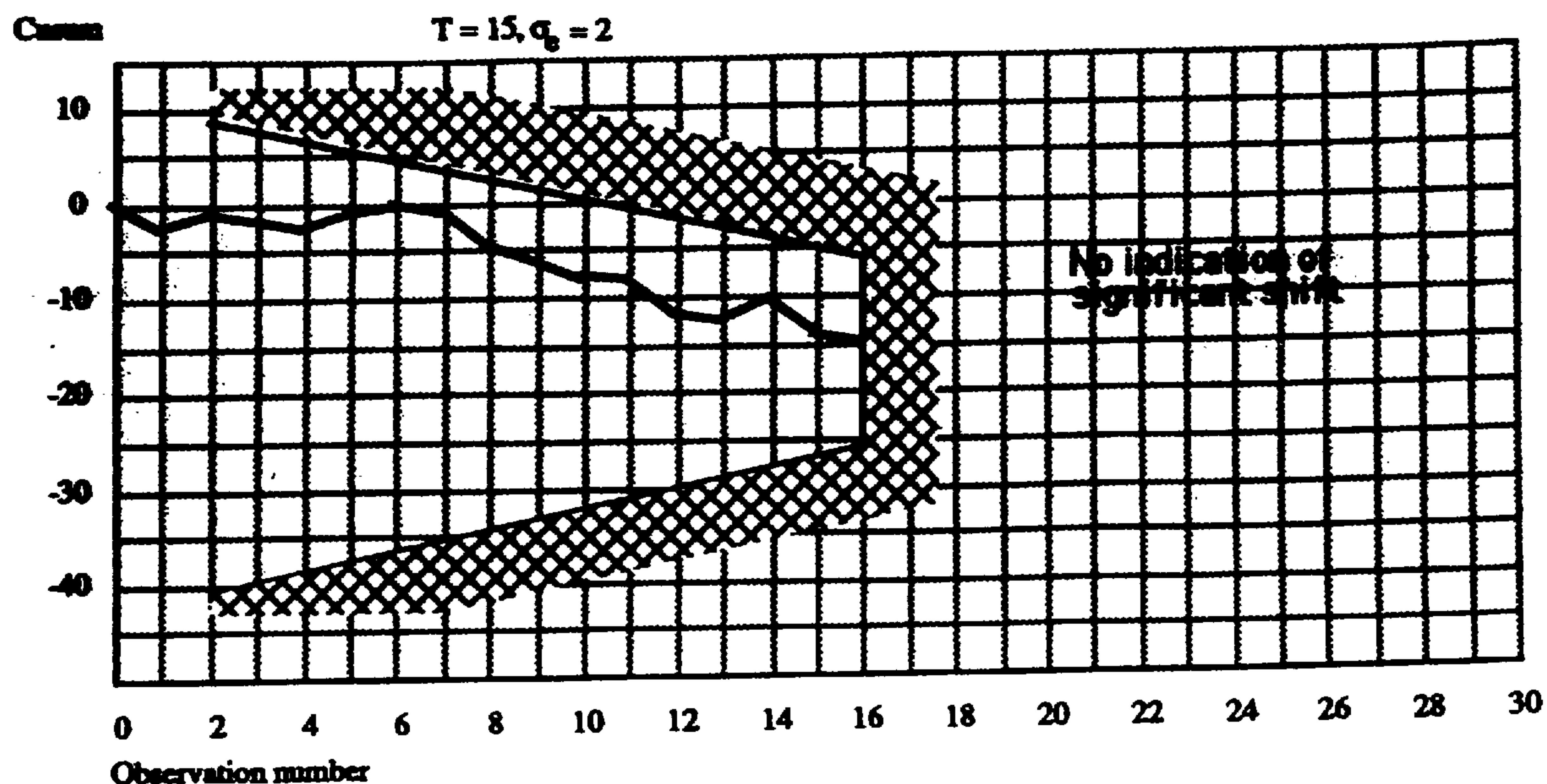


Figure 7 : Truncated V-mask applied to cusum chart : no indication of shift

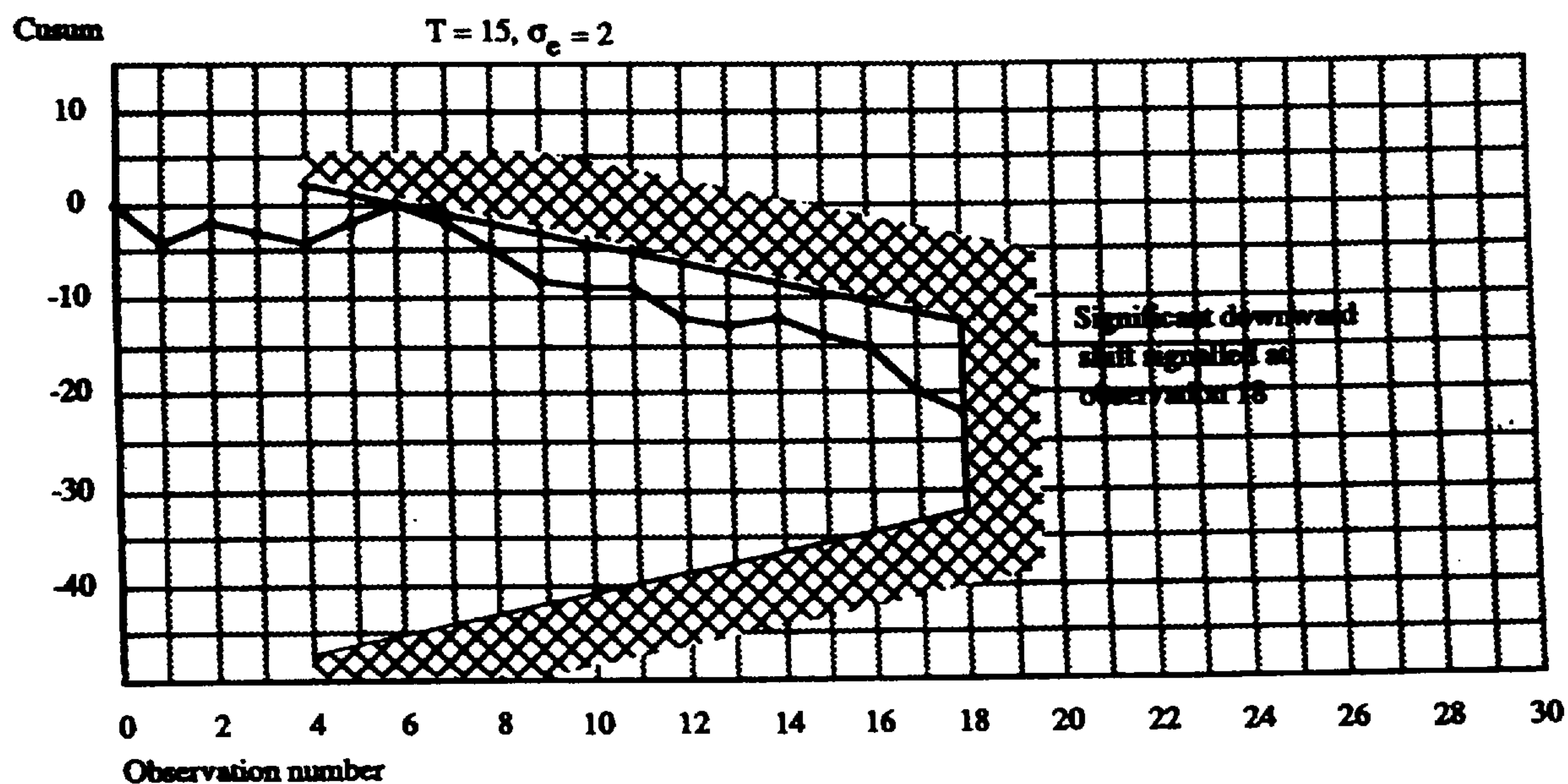


Figure 8 : Truncated V-mask applied to cusum chart : indication of shift

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.