INTERNATIONAL STANDARD

ISO 7870-4

First edition
2011-07-01

Control charts —

Part 4: **Cumulative sum charts**

Cartes de contrôle —

Partie 4: Cartes de contrôle de l'ajustement de processus

Reference number
ISO 7870-4:2011(E)

Contents

Page

Foreword

ISO (the International Organization 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 and 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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 7870-4 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 4, *Applications of statistical methods in process management*.

This first edition of ISO 7870-4 cancels and replaces ISO/TR 7871:1997.

ISO 7870 consists of the following parts, under the general title *Control charts*:

— *Part 1: General guidelines*

— *Part 3: Acceptance control charts*

— *Part 4: Cumulative sum charts*

The following part is under preparation:

— *Part 2: Shewhart control charts*

Additional parts on specialized control charts and on the application of statistical process control (SPC) charts are planned.

Introduction

This part of ISO 7870 demonstrates the versatility and usefulness of a very simple, yet powerful, pictorial method of interpreting data arranged in any meaningful sequence. These data can range from overall business figures such as turnover, profit or overheads to detailed operational data such as stock outs and absenteeism to the control of individual process parameters and product characteristics. The data can either be expressed sequentially as individual values on a continuous scale (e.g. 24,60, 31,21, 18,97...), in “yes”/“no”, “good”/“bad”, “success”/“failure” format, or as summary measures (e.g. mean, range, counts of events).

The method has a rather unusual name, cumulative sum, or, in short, “cusum”. This name relates to the process of subtracting a predetermined value, e.g. a target, preferred or reference value from each observation in a sequence and progressively cumulating (i.e. adding) the differences. The graph of the series of cumulative differences is known as a cusum chart. Such a simple arithmetical process has a remarkable effect on the visual interpretation of the data as will be illustrated.

The cusum method is already used unwittingly by golfers throughout the world. By scoring a round as “plus” 4, or perhaps even “minus” 2, golfers are using the cusum method in a numerical sense. They subtract the “par” value from their actual score and add (cumulate) the resulting differences. This is the cusum method in action. However, it remains largely unknown and hence is a grossly underused tool throughout business, industry, commerce and public service. This is probably due to cusum methods generally being presented in statistical language rather than in the language of the workplace.

This part of ISO 7870 is a revision of ISO/TR 7871:1997. The intention of this part is, thus, to be readily comprehensible to the extensive range of prospective users and so facilitate widespread communication and understanding of the method. The method offers advantages over the more commonly found Shewhart charts in as much as the cusum method will detect a change of an important amount up to three times faster. Further, as in golf, when the target changes per hole, a cusum plot is unaffected, unlike a standard Shewhart chart where the control lines would require a constant adjustment.

In addition to Shewhart charts, an EWMA (exponentially weighted moving average) chart, can be used. Each plotted point on an EWMA chart incorporates information from all of the previous subgroups or observations, but gives less weight to process data as they get “older” according to an exponentially decaying weight. In a similar manner to a cusum chart, an EWMA chart can be sensitized to detect any size of shift in a process. This subject is discussed further in another part of this International Standard.

Control charts —

Part 4:
Cumulative sum charts

1 Scope

This part of ISO 7870 provides statistical procedures for setting up cumulative sum (cusum) schemes for process and quality control using variables (measured) and attribute data. It describes general-purpose methods of decision-making using cumulative sum (cusum) techniques for monitoring, control and retrospective analysis.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*

ISO 3534-2, *Statistics — Vocabulary and symbols — Part 2: Applied statistics*

3 Terms and definitions, abbreviated terms and symbols

For the purposes of this document, the terms and definitions given in ISO 3534-1 and ISO 3534-2 and the following apply.

3.1 Terms and definitions

3.1.1
target value
T
value for which a departure from an average level is required to be detected

NOTE 1 With a charted cusum, the deviations from the target value are cumulated.

NOTE 2 Using a “V” mask, the target value is often referred to as the reference value or the nominal control value. If so, it should be acknowledged that it is not necessarily the most desirable or preferred value, as may appear in other standards. It is simply a convenient target value for constructing a cusum chart.

3.1.2
datum value
⟨tabulated cusum⟩ value from which differences are calculated

NOTE The upper datum value is $T + f\sigma_e$, for monitoring an upward shift. The lower datum value is $T - f\sigma_e$, for monitoring a downward shift.

3.1.3
reference shift
F, *f*
⟨tabulated cusum⟩ difference between the **target value** (3.1.1) and **datum value** (3.1.2)

NOTE It is necessary to distinguish between *f* that relates to a standardized reference shift, and *F* to an observed reference shift, $F = f\sigma_e$.

3.1.4
reference shift
F, *f*
⟨truncated V-mask⟩ slope of the arm of the mask (tangent of the mask angle)

NOTE It is necessary to distinguish between *f* that relates to a standardized reference shift, and *F* to an observed reference shift, $F = f\sigma_e$.

3.1.5
decision interval
H, *h*
⟨tabulated cusum⟩ cumulative sum of deviations from a **datum value** (3.1.2) required to yield a signal

NOTE It is necessary to distinguish between *h* that relates to a standardized decision interval, and *H* to an observed decision interval, $H = h\sigma_e$.

3.1.6
decision interval
H, h
⟨truncated V-mask⟩ half-height at the datum of the mask

NOTE It is necessary to distinguish between *h* that relates to a standardized decision interval, and *H* to an observed decision interval, $H = h\sigma_e$.

3.1.7
average run length
L
average number of samples taken up to the point at which a signal occurs

NOTE Average run length (*L*) is usually related to a particular process level in which case it carries an appropriate subscript, as, for example, L_0, meaning the average run length when the process is at target level, i.e. zero shift.

3.2 Abbreviated terms

ARL average run length

CS1 cusum scheme with a long ARL at zero shift

CS2 cusum scheme with a shorter ARL at zero shift

DI decision interval

EWMA exponentially weighted moving average

FIR fast initial response

LCL lower control limit

RV reference value

UCL upper control limit

3.3 Symbols

a scale coefficient

C cusum value

C_r difference in the cusum value between the lead point and the out-of-control point

c_4 factor for estimating the within-subgroup standard deviation

δ amount of change to be detected

Δ standardized amount of change to be detected

d lead distance

d_2 factor for estimating the within-subgroup standard deviation from within-subgroup range

F observed reference shift

f standardized reference shift

H observed decision interval

h standardized decision interval

J index number

φ size of process adjustment

K cusum datum value for discrete data

k number of subgroups

L_0 average run length at zero shift

L_δ average run length at δ shift

μ population mean value

m mean count number

n subgroup size

p probability of “success”

$\bar{R}$ mean subgroup range

r number of plotted points between the lead point and the out-of-control point

σ process standard deviation

σ_0 within-subgroup standard deviation

$\hat{\sigma}_0$ estimated within-subgroup standard deviation

σ_e standard error

s observed within-subgroup standard deviation

$\overline{s}$ average subgroup standard deviation

$s_{\overline{x}}$ realized standard error of the mean from k subgroups

T target value

T_m reference or target rate of occurrence

T_p reference or target proportion

τ true change point

t observed change point

V_{avg} average voltage

$\hat{V}_{avg}$ estimated average voltage

w difference between successive subgroup mean values

x individual result

$\overline{x}$ arithmetic mean value (of a subgroup)

$\overline{\overline{x}}$ mean of subgroup means

4 Principal features of cumulative sum (cusum) charts

A cusum chart is essentially a running total of deviations from some preselected reference value. The mean of any group of consecutive values is represented visually by the current slope of the graph. The principal features of a cusum chart are the following.

a) It is sensitive in detecting changes in the mean.

b) Any change in the mean, and the extent of the change, is indicated visually by a change in the slope of the graph:

   1) a horizontal graph indicates an “on-target” or reference value;

   2) a downward slope indicates a mean less than the reference or target value: the steeper the slope, the bigger the difference;

   3) an upward slope indicates a mean more than the reference or target value: the steeper the slope, the bigger the difference.

c) It can be used retrospectively for investigative purposes, on a running basis for control, and for prediction of performance in the immediate future.

Referring to point b) above, a cusum chart has the capacity to clearly indicate points of change; they will be clearly indicated by the change in gradient of the cusum plot. This has enormous benefit for process management: to be able to quickly and accurately pinpoint the moment when a process altered so that the appropriate corrective action can be taken.

A further very useful feature of a cusum system is that it can be handled without plotting, i.e. in tabular form. This is very helpful if the system is to be used to monitor a highly technical process, e.g. plastic film manufacture, where the number of process parameters and product characteristics is large. Data from such a process might be captured automatically, downloaded into cusum software to produce an automated cusum analysis. A process manager can then be alerted to changes on many characteristics on a simultaneous basis. Annex B contains an example of the method.

5 Basic steps in the construction of cusum charts — Graphical representation

The following steps are used to set up a cusum chart for individual values.

Step 1: Choose a reference, target, control or preferred value. The average of past results will generally provide good discrimination.

Step 2: Tabulate the results in a meaningful (e.g. chronological) sequence. Subtract the reference value from each result.

Step 3: Progressively sum the values obtained in Step 2. These sums are then plotted as a cusum chart.

Step 4: To obtain the best visual effect set up a horizontal scale no wider than about 2,5 mm between plotting points.

Step 5: For reasonable discrimination, without undue sensitivity, the following options are recommended:

a) choose a convenient plotting interval for the horizontal axis and make the same interval on the vertical axis equal to 2σ (or $2\sigma_e$ if a cusum of means is to be charted), rounding off as appropriate; or

b) where it is required to detect a known change, say δ, choose a vertical scale such that the ratio of the scale unit on the vertical scale divided by the scale unit on the horizontal scale is between δ and 2δ, rounding off as appropriate.

NOTE The scale selection is visually very important since an inappropriate scale will give either the impression of impending disaster due to the volatile nature of the plot, or a view that nothing is changing. The schemes described in a) and b) above should give a scale that shows changes in a reasonable manner, neither too sensitive nor too suppressed.

6 Example of a cusum plot — Motor voltages

6.1 The process

Suppose a set of 40 values in chronological sequence is obtained of a particular characteristic. These happen to be voltages, taken in order of production, on fractional horsepower motors at an early stage of production. But they could be any individual values taken in a meaningful sequence and expressed on a continuous scale. These are now shown:

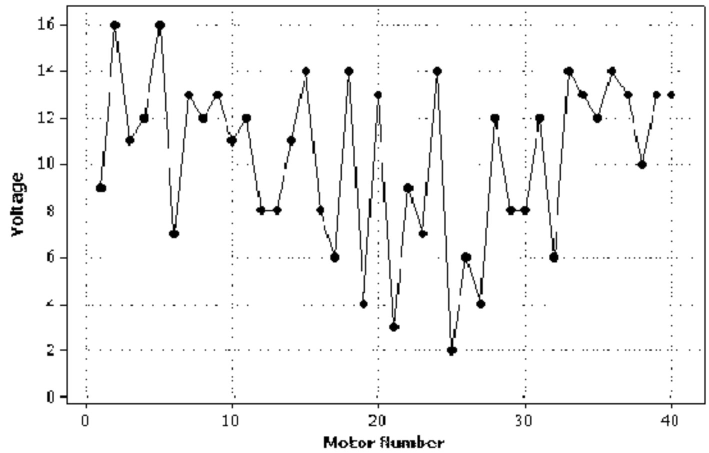
9, 16, 11, 12, 16, 7, 13, 12, 13, 11, 12, 8, 8, 11, 14, 8, 6, 14, 4, 13, 3, 9, 7, 14, 2, 6, 4, 12, 8, 8, 12, 6, 14, 13, 12, 14, 13, 10, 13, 13.

The reference or target voltage value is 10 V.

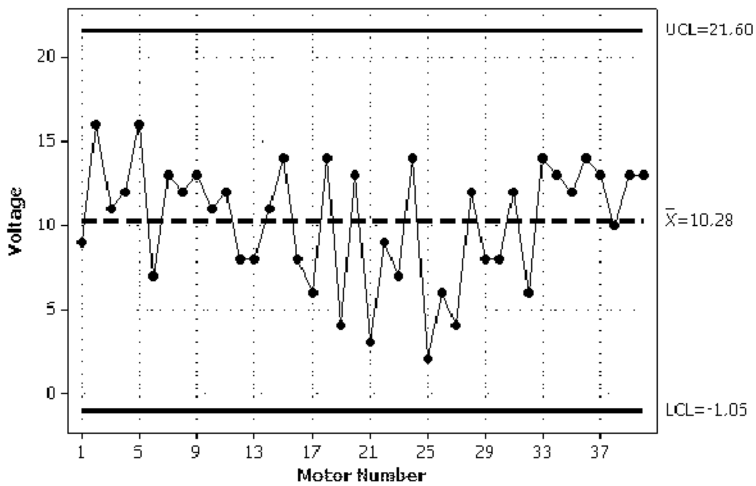
6.2 Simple plot of results

In order to gain a better understanding of the underlying behaviour of the process, by determining patterns and trends, a standard approach would be simply to plot these values in their natural order as shown in Figure 1 a).

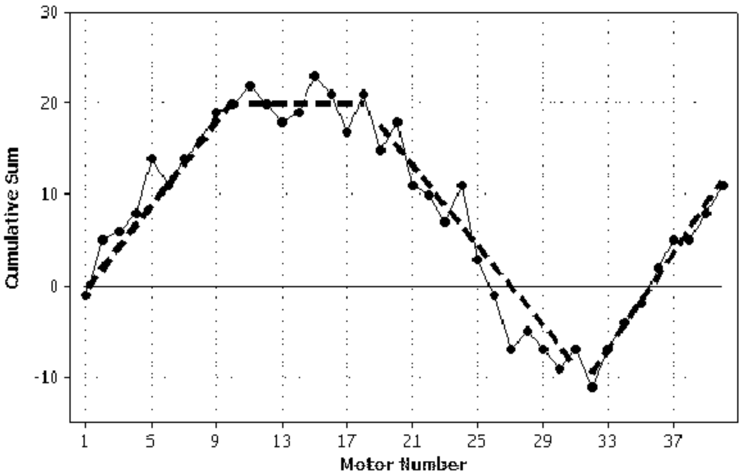
Apart from indicating a general drop away in the middle portion from a high start and with an equally high finish, Figure 1 a) is not very revealing because of the extremely noisy, or spiky, data throughout.



a) Simple plot of motor voltages



b) Standard control chart for individuals



c) Cusum chart

Figure 1 — Motor voltage example

6.3 Standard control chart for individual results

The next level of sophistication would be to establish a standard control chart for individuals as in Figure 1 b).

Figure 1 b) is even less revealing than the previous figure. It is, in fact, quite misleading. The standard statistical process control criteria to test for process stability and control are

a) no points lying above the upper control limit (UCL) or below the lower control limit (LCL),

b) no runs of seven or more intervals upwards or downwards,

c) no runs of seven points above or below the centreline.

The answer to all these criteria is "no". Hence, one would be led to the conclusion that this is a stable process, one that is "in control" around its overall average value of about 10 V, which is the target value. Further standard analysis would reveal that although the process is stable, it is not capable of meeting specification requirements. However, this analysis would not in itself provide any further clues as to why it is incapable of meeting the requirements.

The reason for the inability of the standard control chart for individuals to be of value here is that the control limits are based on actual process performance and not on desired or specified requirements. Consequently, if the process naturally exhibits a large variation the control limits are correspondingly wide. What is required is a method that is better at indicating patterns and trends, or even pinpointing points of change, in order to help determine and remove primary sources of variation.

NOTE By using additional tools, such as an individual and moving range chart, the practitioner can study other process variation issues.

6.4 Cusum chart — Overall perspective

Another option here, the one recommended, would be to plot a cusum chart. Figure 1 c) illustrates the cusum plot of the same data.

It was not immediately apparent from the previous charts where, or whether, any significant changes in process level occurred, whereas the cusum chart indicates a well-defined pattern. The best fitting (by eye) indicates four changes in process level, changing after the 10th, 18th and 31st motors.

It is noted, from Clause 4, that an upward/downward slope indicates a value higher/lower than the preferred value and a horizontal line is indicative of a process at the preferred value. Hence, it is seen that this process appears to be on target only for a short period between around motor 11 and 18. Motors 1 to 10 were running higher similarly to motors 33 onwards, whereas the process between about motors 19 and 32 was delivering motors with low voltages.

These changes and their significance are further discussed and interpreted in detail in 6.6.

In a real life situation, the next step would be to seek out what happened operationally at these points of production to cause such changes in voltage performance. This poses certain questions directed specifically at improving the consistency of performance at the 10 V level. For instance, how did the build characteristics of motor 32 differ from those of 33? Or, what happened to the test gear calibration at this point? Did this correspond with a shift, manning or batch change? And so on. Used in this way, whatever the situation, the cusum chart can be a superb diagnostic tool. It pinpoints opportunities for improvement.

6.5 Cusum chart construction

The construction of a cusum chart using individual values, as in this example, is based on the very simple steps given in Clause 5.

Step 1: Choose a reference value, RV. Here the preferred or reference value is given as 10 V.

Step 2: Tabulate the results (voltages) in production sequence against motor number as in Table 1, column 2 (and 6). Subtract the reference value of 10 from each result as in Table 1, column 3 (and 7).

Step 3: Progressively sum the values of Table 1, column 3 (and 7) in column 4 (and 8). Plot column 4 (and 8) against the observation (motor) number as in Figure 1 c), taking note of the scale comments in Steps 4 and 5.

Table 1 — Tabular arrangement for calculating cusum values from a sequence of individual values

(1) Motor no.	(2) Voltage	(3) Voltage –10	(4) Cusum	(5) Motor no.	(6) Voltage	(7) Voltage –10	(8) Cusum
1	9	–1	–1	21	3	–7	+11
2	16	+6	+5	22	9	–1	+10
3	11	+1	+6	23	7	–3	+7
4	12	+2	+8	24	14	+4	+11
5	16	+6	+14	25	2	–8	+3
6	7	–3	+11	26	6	–4	–1
7	13	+3	+14	27	4	–6	–7
8	12	+2	+16	28	12	+2	–5
9	13	+3	+19	29	8	–2	–7
10	11	+1	+20	30	8	–2	–9
11	12	+2	+22	31	12	+2	–7
12	8	–2	+20	32	6	–4	–11
13	8	–2	+18	33	14	+4	–7
14	11	+1	+19	34	13	+3	–11
15	14	+4	+23	35	12	+2	–7
16	8	–2	+21	36	14	+4	–4
17	6	–4	+17	37	13	+3	–2
18	14	+4	+21	38	10	0	+2
19	4	–6	+15	39	13	+3	+5
20	13	+3	+18	40	13	+3	+5

6.6 Cusum chart interpretation

6.6.1 Introduction

When a cusum chart is used in retrospective diagnostic mode, as in this example, it is usually better not to focus on individual plotting points but to draw the minimum number of straight lines that are representative of lines of best fit by eye, through the data as in Figure 1 c).

One has to be very careful then not to interpret either the slope of these lines or their relative position related to the vertical axis, as with conventional data plots. It should be noted, too, that the vertical axis no longer represents actual voltages.

A straight line with an upward/downward slope does not indicate that the process level is increasing/decreasing, as is customary, but rather that it is constant at a value more/less than the reference value. The steeper the slope, the greater the difference. A horizontal line indicates that the process level is constant at the reference value. The interpretation of the cusum chart for the motor is now discussed in more detail.

6.6.2 The basics of interpretation of a cusum chart using "imaginary noiseless" data

Suppose that the sequence of the first 18 motor voltages had been 10, 10, 10, 13, 13, 13, 10, 10, 10, 9, 9, 9, 10, 10, 10, 8, 8, 8, as shown in Table 2, column 2, and that the reference value is still 10 V.

Table 2 — Imaginary motor data to illustrate the basic interpretation of a cusum chart

(1) Motor no.	(2) Voltage	(3) Voltage – 10	(4) Cusum
1	10	0	0
2	10	0	0
3	10	0	0
4	13	+3	+3
5	13	+3	+6
6	13	+3	+9
7	10	0	+9
8	10	0	+9
9	10	0	+9
10	9	–1	+8
11	9	–1	+7
12	9	–1	+6
13	10	0	+6
14	10	0	+6
15	10	0	+6
16	8	–2	+4
17	8	–2	+2
18	8	–2	0

The resulting cusum chart will now look as in Figure 2.

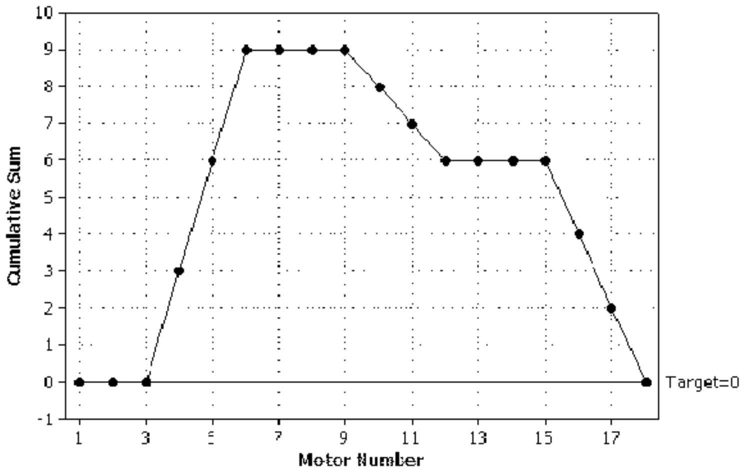


Figure 2 — Cusum chart of imaginary motor voltage data to illustrate its interpretation

In comparing the actual voltages of Table 2, column 2 with the cusum chart of Figure 2, it is seen that:

a) motors 1 to 3, 7 to 9 and 13 to 15 were all at the reference value of 10 V and that these are all represented by horizontal lines in the cusum chart. It will also be noted that the positions of the horizontal lines with respect to the vertical scale are not related to these actual motors but rather to previous performances;

b) motors 4 to 6 were at a value higher than the reference value, namely 13 V, and that these motors are represented by an upward slope on the cusum chart. This is obvious here as there is no variability in voltage between the motors to confuse the issue. If there were noise then the equation to calculate the average value over the period from the particular slope is:

$$\text{Average voltage} = \text{Reference value} + \left(\frac{\text{Cusum value at the end of line} - \text{Cusum value at the start of line}}{\text{Number of observation intervals}} \right)$$

$$V_{\text{avg}} = 10 + \left(\frac{9-0}{3} \right) = 13$$

c) similarly for motors 9 to 12:

$$V_{\text{avg}} = 10 + \left(\frac{6-9}{3} \right) = 9$$

d) and for motors 16 to 18:

$$V_{\text{avg}} = 10 + \left(\frac{0-6}{3} \right) = 8$$

Summarizing, the different slopes on the cusum chart indicate that from motors:

— 1 to 3, 7 to 9 and 13 to 15, the voltage remained constant at a value of 10;

— 4 to 6, the voltage also remained constant but at a value of 13;

— 10 to 12, the voltage remained constant at a value of 9; and

— 16 to 18, the voltage remained constant at a value of 8.

This was obvious by referring back to the “noiseless” data here. But it is not immediately apparent when referring to the actual “noisy” data in Table 1, columns 2 and 6.

6.6.3 Interpretation using “actual” data

The cusum chart of Figure 1 c) shows:

a) the average voltage level from motor number 1 to 10 is at a higher value than the reference voltage. The calculated value is given by the slope thus:

$$\text{Average voltage} = \text{Reference value} + \left(\frac{\text{Cusum value at the end of line} - \text{Cusum value at the start of line}}{\text{Number of observation intervals}} \right)$$

$$V_{\text{avg}} = 10 + \left(\frac{20 - 0}{10} \right) = 12 \text{ V}$$

b) similarly for motors 11 to 18, the average voltage = 10 as the line is horizontal;

c) for motors 19 to 31:

$$\hat{V}_{\text{avg}} = 10 + \left(\frac{-12 - 20}{13} \right) \approx 7{,}5 \text{ V}$$

d) for motors 32 to 40:

$$V_{\text{avg}} = 10 + \left[\frac{11 - (-12)}{9} \right] \approx 12{,}6 \text{ V}$$

Summarizing, the cusum chart enables us to calculate variable period moving averages matched to actual process performance. This represents a considerable advance on the standard predetermined and inflexible moving average approach more commonly used. The summary estimate of results is given in Table 3.

Table 3 — Average voltages for motors in terms of variable moving average periods

Motors	Average motor voltage
1 to 10	12,0
11 to 18	10,0
19 to 31	7,5
32 to 40	12,6

As an alternative to this method of calculating the relationship between cusum slope and average voltage, one can simply calculate local moving averages for each constant level portion of the cusum chart.

For example, for motors 1 to 10, by calculation:

$$V_{\text{avg}} = \frac{(9 + 16 + 11 + 12 + 16 + 7 + 13 + 12 + 13 + 11)}{10} = 12{,}0$$

This use of individual voltages will sometimes give slightly different results to the slope method. This results from the smoothing out of local variation in the data by putting a straight line through individual points.

6.7 Manhattan diagram

Having established estimates of points of change in voltage level and their values, it is often found convenient, to further simplify and enhance the presentation, to go to an extra stage of presenting the data in “noiseless” form, in terms of the original vertical axis indicating actual voltages. This presentation is inspired by the Manhattan rectilinear skyline and is consequently known as a Manhattan diagram.

It is simply an expression of the results, shown in 6.6.3 a), b), c) and d), as a conventional plot of voltage against motor production sequence. This is shown in Figure 3 for comparison with the cusum data in Figure 1 c) and the original noisy data in Figure 1 a).

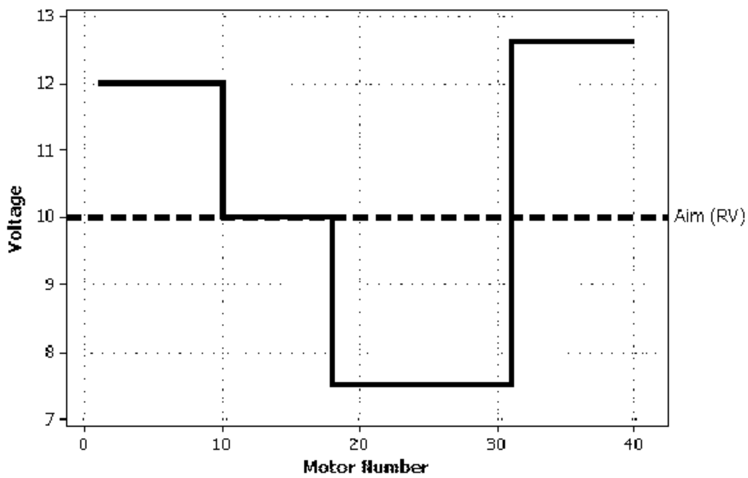


Figure 3 — Manhattan plot of motor data

Figures 3, 1 c) and 1 a) summarize the role and value of the cusum method in investigative mode through retrospective analysis of process performance. They show what can be achieved using readily understandable language, and simple visual enhancement methods without the intrusion of, or recourse to, mathematical symbolism or formal statistical expressions.

Because of the simplicity and unambiguous nature of the Manhattan diagram, it is sometimes useful to look on the cusum diagram as the intermediate technical stage and simply present the data in Manhattan format to facilitate wider non-technical communication, understanding and application.

7 Fundamentals of making cusum-based decisions

7.1 The need for decision rules

Decision rules might be needed to rationalize the interpretation of a cusum chart. When an appropriate decision rule so indicates, some action is taken, depending on the nature of the process. Typical actions are:

a) *for in-process control*: adjustment of process conditions;

b) *in an improvement context:* investigation of the underlying cause of the change; and

c) *in a forecasting mode:* analysis of and, if necessary, adjustment to the forecasting model or its parameters.

7.2 The basis for making decisions

Establishing the base criteria against which decisions are to be made is obviously an essential prerequisite.

To provide an effective basis for detecting a signal, a suitable quantitative measure of "noise" in the system is required. What represents noise, and what represents a signal, is determined by the monitoring strategy adopted, such as how many observations to take, and how frequently, and how to constitute a sample or a subgroup. Also, the measure used to quantify variation can affect the issue.

It is usual to measure inherent variation by means of a statistical measure termed either of the following.

a) *Standard deviation*: where individual observations are the basis for plotting cusums

   The individual observations for calculation of the standard deviation are often taken from a homogeneous segment of the process data. This performance then becomes the more onerous criterion from which to judge. Any variation greater than this inherent variation is taken to arise from special causes indicating a shift in the mean of the series or a change in the natural magnitude of the variability, or both.

b) *Standard error*: where some function of a subgroup of observations, such as mean, median or range, forms the basis for cusum plotting

   The concept of subgrouping is that variation within a subgroup is made up of common causes with all special causes of variation occurring between subgroups. The primary role of the cusum chart is then to distinguish between common and special cause variation. Hence, the choice of subgroup is of vital importance. For example, making up each subgroup of four consecutively from a high-speed production process each hour, as opposed to one taken every quarter of an hour to make up a subgroup of four every hour, would give very different variabilities on which to base a decision. The standard error would be minuscule in the first instance compared with the second. One cusum chart would be set up with consecutive part variation as the basis for decision-making as opposed to 15 min to 15 min variation for the other chart. The appropriate measure of underlying variability will depend on which changes it is required to signal.

However, the prerequisite that stability should exist over a sufficient period to establish reliable quantitative measures, such as standard deviation or standard error, is too restrictive for some potential areas of application of the cusum method.

For instance, observations of a continuous process can exhibit small unimportant variations in the average level. It is required that it is against these variations that systematic or sustained changes should be judged. Illustrations are:

a) an industrial process is controlled by a thermostat or other automatic control device;

b) the quality of raw material input can be subject to minor variations without violating specification; and

c) in monitoring a patient's response to treatment, there might 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.

On the other hand, samples can comprise output or observations from several sources (administrative regions, plants, machines and operators). As such, there might be too much local variation to provide a realistic basis for assessing whether or not the overall average shifts. Because of this factor, data arising from a combination of sources should be treated with caution, as any local peculiarities within each contributing source might be overlooked. Moreover, variation between the sources might mask any changes occurring over the whole system as time progresses.

One of the important assumptions in cusum procedures is that the process standard deviation σ is stable. Therefore, before constructing the cusum procedure, any process should be assessed to see if it is in a state of statistical control (by using the R-chart, s-chart or moving range chart) so that a reliable estimate of σ can be obtained.

Serial correlation between observations can also manifest itself — namely, one observation might have some influence over the next. An illustration of negative serial correlation is the use of successive gauge readings to estimate the use of a bulk material, where an overestimate on one occasion will tend to produce an underestimate on the next reading. Another example is where overordering in one month is compensated by underordering in the subsequent month. Positive serial correlation is likely in some industrial processes where one batch of material might partially mix with preceding and succeeding batches.

Budgetary and accounting interval ends, project milestones and contract deadlines can affect the allocation of successive business figures, such as costs and sales on a period-to-period basis, and so on.

In view of these aspects, it is necessary to consider other quantitative measures of variation in the series or sequences of data and the circumstances in which they are appropriate.

Such measures of variation on which to base decision-making using cusums are developed, in a quantitative sense, in Annex A. Recommendations are also made as to which to choose depending on the circumstances.

7.3 Measuring the effectiveness of a decision rule

7.3.1 Basic concepts

The ideal performance of a decision rule would be for real changes of at least a prespecified magnitude to be detected immediately and for a process with no real changes to be allowed to continue indefinitely without giving rise to false alarms. In real life this is not attainable. A simple and convenient measure of actual effectiveness of a decision rule is the average run length (ARL).

The ARL is the expected value of the number of samples taken up to that which gives rise to a decision that a real change is present.

If no real change is present, the ideal value of the ARL is infinity. A practical objective in such a situation is to make the ARL large. Conversely, when a real change is present, the ideal value of the ARL is 1, in which case the change is detected when the next sample is taken. The choice of the ARL is a compromise between these two conflicting requirements. Making an incorrect decision to act when the process has not changed gives rise to “overcontrol”. This will, in effect, increase variability. Not taking appropriate action when the process has changed gives rise to “undercontrol”. This will also, in effect, increase variability and also results in increasing cost of production.

Of course the ARL itself is subject to statistical variation. Sometimes one can be fortunate in obtaining no false alarms over a long run, or in detecting a change very quickly. Occasionally, an unfortunate run of samples can generate false alarms or mask a real change so that it does not yield a signal. The actual pattern of such variation deserves attention once in a while. Generally, however, the ARL is looked upon as a reasonable measure of effectiveness of a decision rule. Summarizing, the aim is:

True process condition	Required cusum response	Ideal response
At or near target	Long ARL (few false alarms)	ARL = infinity
Significant departure from target	Short ARL (rapid detection)	ARL = 1

7.3.2 Example of the calculation of ARL

The ARL concept is not particular to cusums. Take a standard Shewhart control chart with control limits set at ±3 standard deviations from the centreline. This is illustrated in Figure 4 for a normal distribution.

The distribution shown is termed “standardized” in that it has a zero mean and unit standard deviation.

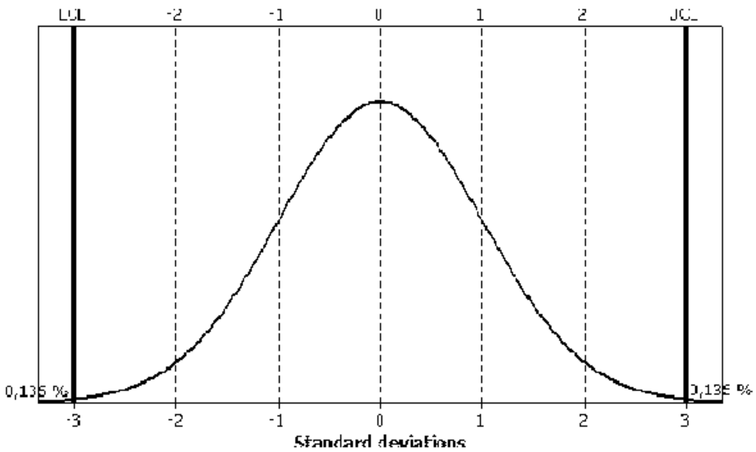


Figure 4 — Plot of standardized normal distribution

It is seen from Figure 4 that some 0,135 % of the observations are expected, on average, to fall beyond each of these limits when the process average is on the centreline or target value. This can readily be translated into an average run length, ARL, by calculating 1/0,001 35 = 741. In other words, we would expect, on average, to see a value beyond the upper control limit only once in every 741 observation intervals. Such a value would trigger an erroneous signal of a change in level when, in fact, such a change has not occurred.

Hence the need, in practice, to design a control system that ensures a high ARL when the process is running at the target value.

When two-sided limits are considered, with the process mean still on target, the ARL is halved, it now being 1/(0,001 35 + 0,001 35) = 370.

Suppose that the process mean shifts one standard deviation towards the upper control limit. The expectation is then that some 2,28 % will lie above the upper control limit. The ARL in respect of the UCL then becomes 1/0,022 8 = 44 for this single-sided limit. In other words, on average, it will take some 44 observation intervals to signal a shift in the mean of one standard deviation.

When two-sided limits are considered here only 0,003 2 % is expected below the LCL as the process mean is four standard deviations away from the LCL. As 1/(0,000 032 + 0,022 8) does not materially affect the ARL calculated for a single limit, for a one standard deviation shift in the mean, the ARL for a double-sided limit is approximately the same as for a single-sided one, namely 44.

Summarizing:

With the mean at the target value	ARL for a two-sided limit is half that of a single-sided limit
As the shift in the mean increases	ARL for a two-sided limit approaches that of a single-sided limit

Of course, in practice, other signalling rules such as the addition of warning limits, runs above and below the mean, and so on, will secure more rapid detection of shift but at the expense of an increase in spurious signals when the process is on target. The Shewhart chart is very attractive and popular because of its extreme simplicity and its effectiveness in detecting isolated special causes which give rise to large shifts.

However, it is recognized that it has an inherent limitation in signalling other than large shifts even if they persist without seriously prejudicing the extent of false alarms. This indicates a role for quite a different method in order to achieve a more rapid detection of shift while retaining long ARLs when on target. The cusum method is well suited to this.

8 Types of cusum decision schemes

8.1 V-mask types

The simplest decision rules for use in conjunction with cusum charts are embodied in V-type masks. There are four slightly different forms of mask, but all are identical in principle and effect. Their purpose is explained in the subclauses that follow. The types are:

a) truncated V-mask;

b) semi-parabolic mask;

c) snub-nosed V-mask; and

d) full V-mask.

8.2 Truncated V-mask

8.2.1 Configuration and dimensions

A "general-purpose" truncated V-mask is illustrated in Figure 5. It comprises a datum point indicated by O in the figure. Two vertical lines are set off from the datum, OB and OC, each of $5\sigma_e$ (i.e. $H = 5\sigma_e$) units in length. These two lines are known as decision intervals. Two sloping arms, BA and CD, termed decision lines, may be extended as required to encompass the plotted cusum points. Dimensionally, EO is equal to 10 observation intervals and the vertical distances EA and ED are both $10\sigma_e$ (i.e. to give a gradient of $F = 0,5\sigma_e$) units in length.

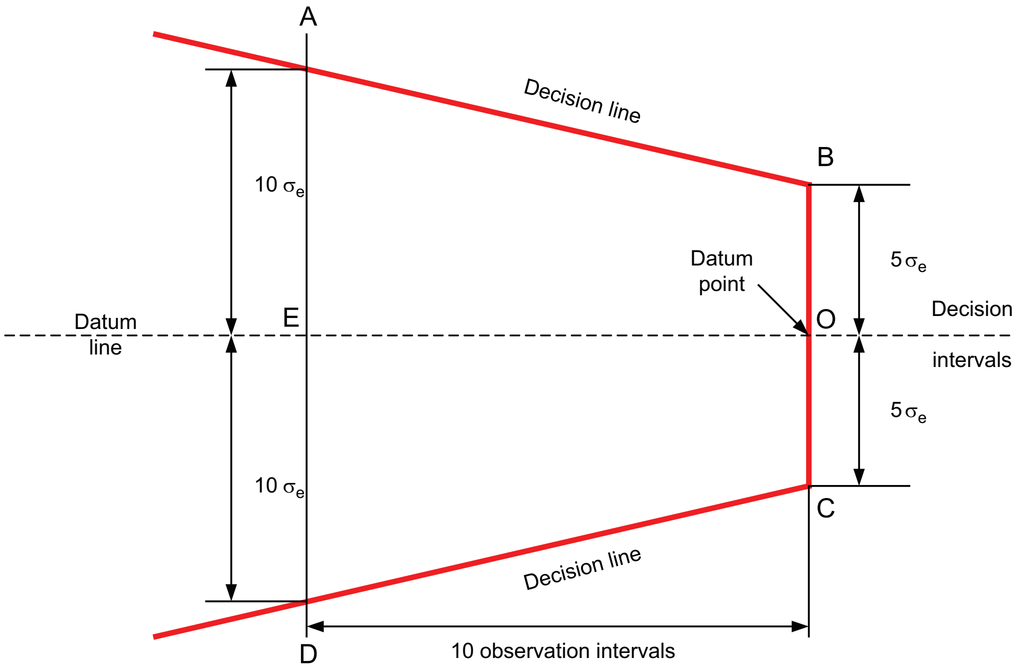


Figure 5 — Configuration and dimensions of general-purpose truncated V-mask

An actual scaled truncated V-mask is shown in Figure 6 for a process variable with a standard deviation of 0,2. The standard deviation is used here, rather than the standard error, because the particular mask is created to monitor individual observations rather than mean values.

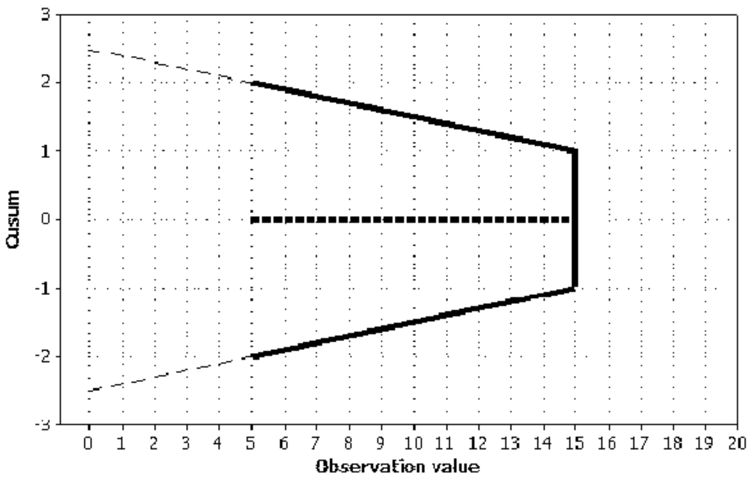


Figure 6 — Actual scaled V-mask for a process characteristic with a particular inherent variation (standard deviation = 0,2)

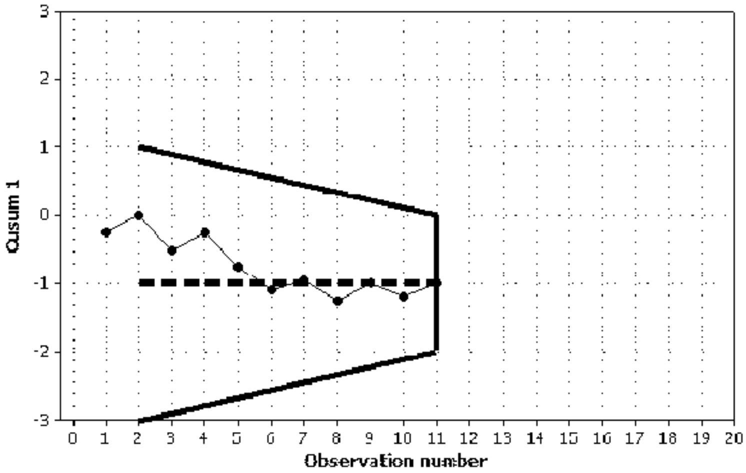
8.2.2 Application of the truncated V-mask

The mask is used by placing the datum point on a selected plotted value on the cusum chart, with the datum line aligned horizontally on the chart. In an ongoing control situation, this selected plotted value is usually the most recent point.

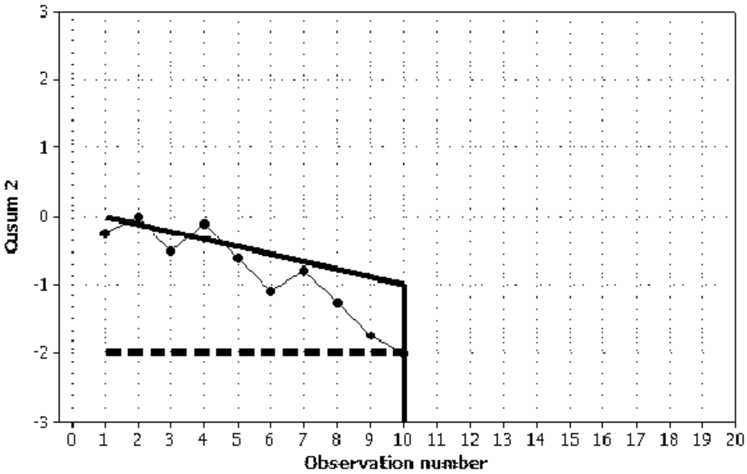
If the path of the cusum lies within the sloping arms of the mask (or their extensions beyond A and D), no significant shift in mean is indicated up to that plotted value. In a control situation the process is then said to be in a state of statistical control with respect to the target value. If, however, the path of the cusum wanders outside the sloping arms of the mask, a significant departure from the target value is signalled. In process management, the process is then said to be out-of-control.

Figure 7 illustrates an “in-control” situation where no significant departure from the target value is detected, and two “out-of-control” situations, one where there is a significant decrease in value indicated and the other where a significant increase is revealed. A standard deviation of 0,2 is used in the three illustrations in Figure 7. The target value used to construct the cusum charts is equal to the target mean for the process.

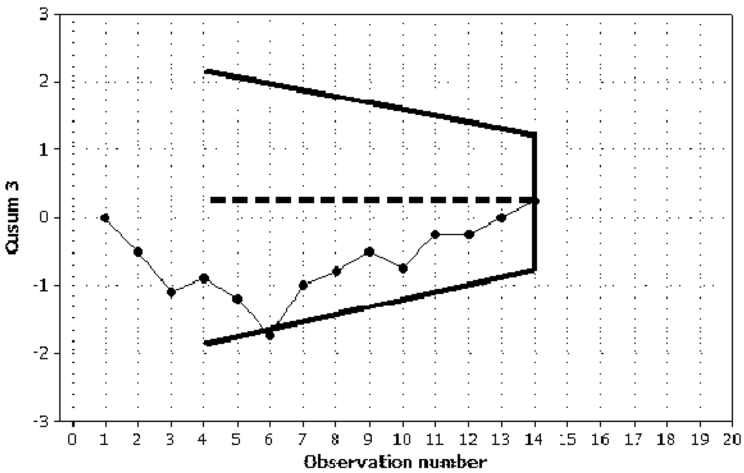
The current situation is determined by offering up the V-mask to the cusum chart progressively as data points accumulate.



a) No significant change in process mean with respect to cusum target value



b) Significant decrease in process mean with respect to cusum target value



c) Significant increase in process mean with respect to cusum target value

Figure 7 — Illustrations of use of the truncated V-mask to detect significant change in process mean

Although Figure 7 a) indicates a process mean less than the cusum target value, the V-mask does not yet register this change as a significant departure.

Figure 7 b) indicates that the process mean is significantly less than the target value. While the significant departure is not detected until observation 10, from a visual perspective the process mean appears to be running low from as early as observation 1. By noting the slope of the line through the observation points, an assessment of the actual mean of the process can be made. This will both provide a guide to the magnitude of the correction required to restore the process to its target value, and a diagnostic indicator to pinpoint what happened at observation 1 to set the process on this low level in the first place.

Figure 7 c) indicates that the process mean is significantly greater than the target value. This was not registered as significant until observation 14. It can be seen that the process appeared to be running lower than the target value until observation 6 but this was not sufficient to trigger an out-of-control condition. Then, following observation 6, the level changed to a higher value than that targeted. By measuring the line slope up to, and from, observation 6, together with its origins, both a corrective tool and a diagnostic aid are provided.

When only an upper or lower specification limit is applicable, one-sided control is appropriate. A half-mask can then be used. When monitoring against an upward/downward shift, the lower/upper portion of the mask only is required, respectively. However the full mask might still be preferred on the grounds of both simplicity and information. Any shifts in the irrelevant direction may be ignored from a specification viewpoint, or used for directing attention to significant movement in a possibly more desirable direction.

8.2.3 Average run lengths

The average run length (ARL) properties for the general-purpose truncated V-mask to the dimensions given in Figure 8 are listed in Table 4 in terms of standard deviation, or standard errors, of the plotted variable. The cusum ARL is compared with those of two decision rules associated with well-established international standard methods of control.

These rules are:

— *Shewhart Rule 1*: one point outside of, ±3 standard deviations from the centreline, namely, the action limits or control limits;

— *Shewhart Rule 2*: two consecutive points outside of, ±2 standard deviations from the centreline, namely, the warning limits.

NOTE 1 The plotted variable is assumed to be normally distributed with a standard deviation σ.

NOTE 2 The average run lengths refer to one-sided control of the average. Where two-sided control from a single target value is adopted, the ARL at the target value is halved (there will be twice the number of false alarms), but for large shifts in mean the ARL is unaffected.

NOTE 3 The standard cusum referred to has h (height of decision interval) = 5,0 and f (slope of decision line) = 0,5 as in Figure 4. The Shewhart Action Limit relates to Shewhart Rule 1 only. The Shewhart Action and Warning Limit applies to the combination of Shewhart Rules 1 and 2.

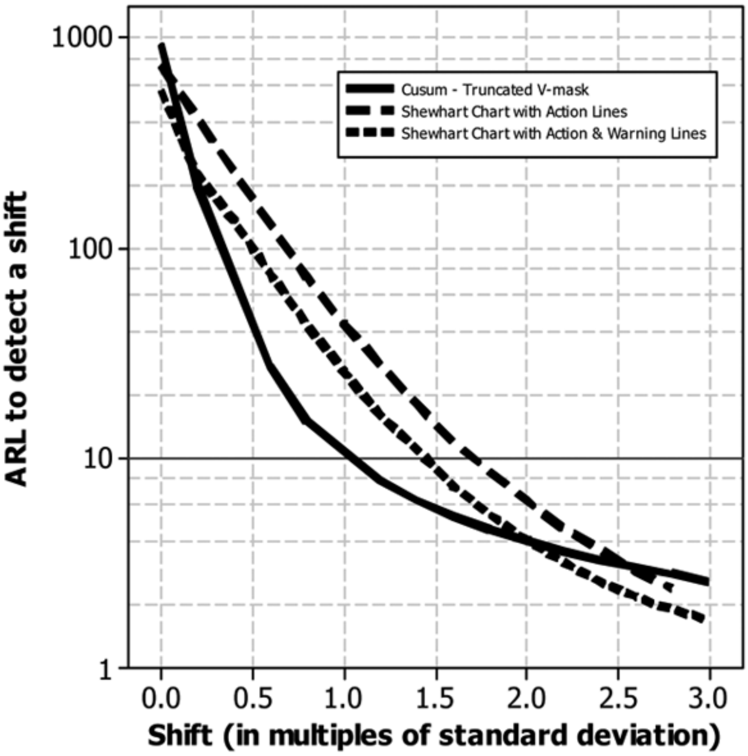


Figure 8 — Average run lengths, in terms of shift from target value, for the general-purpose truncated V-mask compared with standard Shewhart control charts

Table 4 — Average run lengths in terms of shift from target value for general-purpose truncated V-mask of Figure 5 compared with the standard Shewhart control chart using two sets of rules

Shift in process mean from target value (in units of σ_e)	**Average run length**		
	Standard truncated V-mask	**Shewhart control chart with action limits**	**Shewhart control chart with action and warning limits**
0,0	931,0	741,0	556,0
0,2	198,0	308,0	223,0
0,4	60,0	200,0	134,0
0,6	27,0	120,0	75,0
0,8	15,0	72,0	43,0
1,0	10,0	44,0	26,0
1,2	7,8	28,0	16,0
1,4	6,3	18,0	11,0
1,6	5,3	12,0	7,4
1,8	4,6	8,7	5,4
2,0	4,0	6,3	4,1
2,2	3,6	4,7	3,2
2,4	3,3	3,7	2,6
2,6	3,0	2,9	2,2
2,8	2,8	2,4	1,9
3,0	2,6	2,0	1,7

The ARL is an indicator of the effectiveness of a decision method:

- the higher the ARL at the target value, the lower the probability of false alarms;
- the lower the ARL at departures of the mean from its target value, the quicker the detection of the change.

Figure 8 and Table 4 show the following.

a) The L_0 (ARL at zero shift) value of the cusum chart is larger than that of the Shewhart chart with action limits, whereas the Shewhart chart with action and warning limits has a much smaller L_0 value. Thus out of the three charts, the cusum chart has the lowest false alarm rate, where as the Shewhart chart with action and warning limits has the highest false alarm rate.

b) For shifts up to $2\sigma_e$, the ARL of the cusum chart is less than that of either of the others indicating a quicker average response to a shift. This is especially so in the region $0,4\sigma_e$ to $1,4\sigma_e$.

c) For shifts greater than $2\sigma_e$, the Shewhart chart with action and warning rules responds more quickly than the cusum chart. For shifts greater than $2,4\sigma_e$, the Shewhart chart with action and warning limits responds more quickly than the cusum chart. However, this quicker response of the Shewhart chart is at the expense of a larger false alarm rate.

8.2.4 General comments on average run lengths

Firstly, the dimensions of the general-purpose, or standard, truncated V-mask are designed to be especially suited to detecting shifts in the region of one standard error ($1\sigma_e$). If the focus is on other shifts, then different values of h and f are used. Also, V-masks with different configurations, or shapes, to the truncated one may be chosen to improve ARL properties and hence shift detection performance. Examples are the semi-parabolic V-mask and the snub-nosed V-mask discussed in 8.4 and 8.5 respectively.

Secondly, supplementary run rules are quite often used in conjunction with the Shewhart control chart. These include “7 successive points on one side of the mean” and “7 successive plotted intervals all increasing or all decreasing”. A problem with these rules is that they significantly reduce the value of the ARL when the process mean is on target, hence drastically increasing the risk of false alarms.

Thirdly, a number of factors affect the robustness of the ARL measure. These include the shape of the underlying pattern of variation, the value of σ_e, and the independence of observed values. The ARL tables shown in Table 4 and Figure 8 are based on three assumptions:

a) observations are distributed normally;

b) the standard deviation is known exactly; and

c) successive observations are statistically independent.

The normal distribution is symmetrical. 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 will have little effect on the ARL for larger shifts in the mean. Conversely, if the shorter tail is in the direction of the potential shift, the ARL at the target level will be considerably lengthened, again, with little effect on ARL for large shifts.

The standard deviation, or standard error, is usually estimated from a selection of the same observations used for plotting the cusum. Errors of 10 % or more are not uncommon. An overestimation of σ_e increases the ARL, and an underestimation decreases it. This distortion of the ARL is most pronounced at or near target conditions but has little effect at high shifts. Table 5 is indicative of the distortion in ARLs for a 10 % error in estimation of σ_e.

Table 5 — Illustration of effect of incorrect value of standard error, σ_e, on ARL

Shift in process mean from target value, units of true σ_e	Average run length (ARL)		
	10 % overestimate of σ_e	Correct estimate of σ_e	10 % underestimate of σ_e
0,0	3 000,0	930,0	410,0
0,5	45,0	38,0	35,0
1,0	10,0	10,5	10,0
1,5	6,0	5,8	6,0
2,0	4,4	4,1	4,5

Positive autocorrelation tends to shorten the ARL and negative autocorrelation to lengthen it.

It should be noted that the effects of the three assumptions discussed here are not peculiar to the cusum chart but are also applicable to other charting methods.

8.3 Alternative design approaches

An alternative design approach, with a view to improving the performance characteristics over a wider range of shift in the mean, is to use a semi-parabolic V-mask (see 8.4), a snub-nosed V-mask (see 8.5), or a fast initial response (FIR) cusum (see 8.7).

Comparisons of the performances of these alternative designs together with the standard truncated V-mask are shown in Table 6.

Table 6 — Average run lengths (ARLs) for various cusum masks

Shift in mean from target (in σ_e units)	**Truncated V-mask** (h = 5; f = 0,5)	**Semi-parabolic V-mask** (as specified in Table 7)	**Snub-nosed V-mask** (h = 5; f = 0,5 and h = 2.05; f = 1,3)	**Truncated V-mask FIR** (h = 5; f = 0,5)
0,00	465,0	235,0	300,0	448,0
0,25	142,0	113,0	114,0	125,0
0,50	38,0	36,0	36,0	29,0
1,00	10,0	10,0	10,0	6,4
1,50	5,8	5,3	5,3	3,4
2,00	4,0	3,3	3,3	2,4
2,50	3,1	2,3	2,3	1,9
3,00	2,6	1,7	1,8	1,5
3,50	2,2	1,4	1,5	1,3
4,00	2,0	1,2	1,3	1,2

8.4 Semi-parabolic V-mask

The general-purpose, or standard, truncated V-mask was chosen with the parameters $h = 5$ and $f = 0,5$. Different values of the h and f parameters of a truncated V-mask may be selected to provide rapid response to shifts in mean of a particular size. In other cases, some industries such as the food industry, where it is required to improve the speed of detection of larger shifts in the mean while preserving the cusum superiority in signalling smaller shifts, a change of mask type is required.

One solution is the semi-parabolic mask where a curved profile is embodied into a truncated mask near its narrow end as shown in Figure 9.

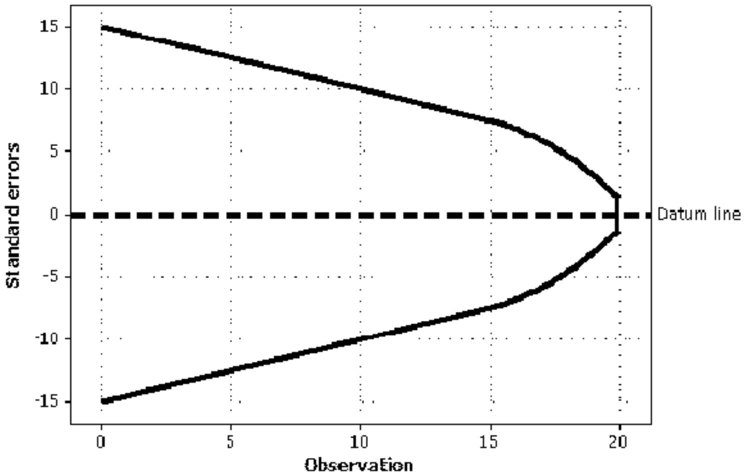


Figure 9 — Illustration of semi-parabolic mask

The basis of the semi-parabolic mask of Figure 9 is the general-purpose truncated V-mask of Figure 5. However, over the last five observation intervals at the narrow end, the mask curves into a width of $1,25\sigma_e$ rather than $5\sigma_e$. The data for the construction of the mask is given in Table 7.

Table 7 — Data for construction of semi-parabolic mask

Distance from datum, J (observation intervals, J)	0	1	2	3	4	5	10	20
Half-width of mask at J (units of σ_e)	1,25	3,10	4,65	5,90	6,85	7,50	10,00	15,00
Construction details	$Y_{\text{half-width of mask at } J} = 1,25 + 2,00J - 0,15J^2$						Linear	

The operating performance of the semi-parabolic mask is:

a) superior to that of the standard zero start truncated V-mask throughout the range of shifts in the mean. This is achieved, however, at the expense of an almost doubling in the rate of false alarms at the target value;

b) inferior to that of the standard FIR truncated V-mask both in terms of false alarm rate, and the signalling of shifts in the mean other than shifts less than 0,5 standard deviations; and

c) inferior to the snub-nosed mask shown in respect of false alarms at the target value while having a comparable performance in terms of detection of shifts in the mean.

These features are reflected in Table 6 showing the comparative ARL, in terms of shift in mean, for various cusum decision rules.

8.5 Snub-nosed V-mask

The snub-nosed V-mask is intended to achieve the same benefits as the semi-parabolic V-mask but with a simpler set-up procedure. Thus it is useful in applications where an earlier response is required for large shifts. This is achieved by superimposing two or more truncated V-masks. An illustration is shown in Figure 10 for a truncated V-mask with $h = 2{,}05$ and $f = 1{,}5$ superimposed on the standard mask with $h = 5{,}0$ and $f = 0{,}5$. Table 6 illustrates that this snub-nosed mask gives almost as good performance as the semi-parabolic one over a wider range of shifts than that achieved by the standard truncated V-mask.

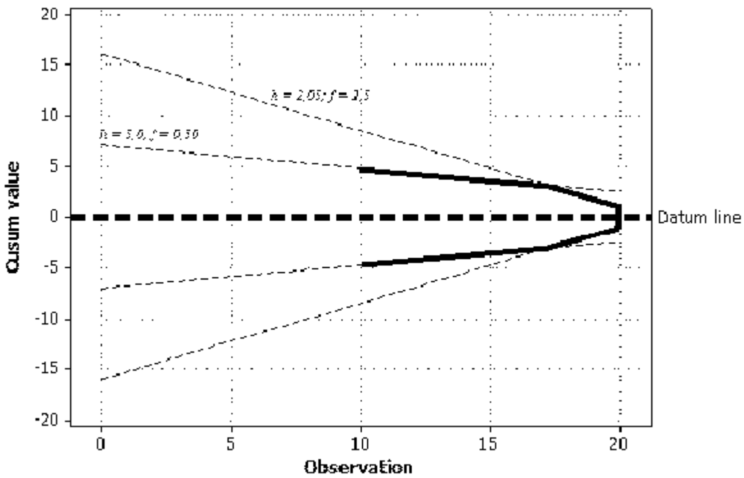


Figure 10 — Illustration of snub-nosed V-mask

8.6 Full V-mask

Decision rules may also be applied using a mask based on a complete V, as shown in Figure 11. This mask has identical performance characteristics to the truncated mask so discussion is curtailed.

Figure 11 indicates that the decision lines are brought to a vertex, O. This means that there is no longer a datum point, and the vertex is positioned on the cusum chart so that it is a distance OA ahead (to the right) of the latest observation point of interest. OA is known as the lead distance, d. For identical characteristics to the standard truncated V-mask discussed, $H = 5\sigma_e$ and $d = 10$ observation units.

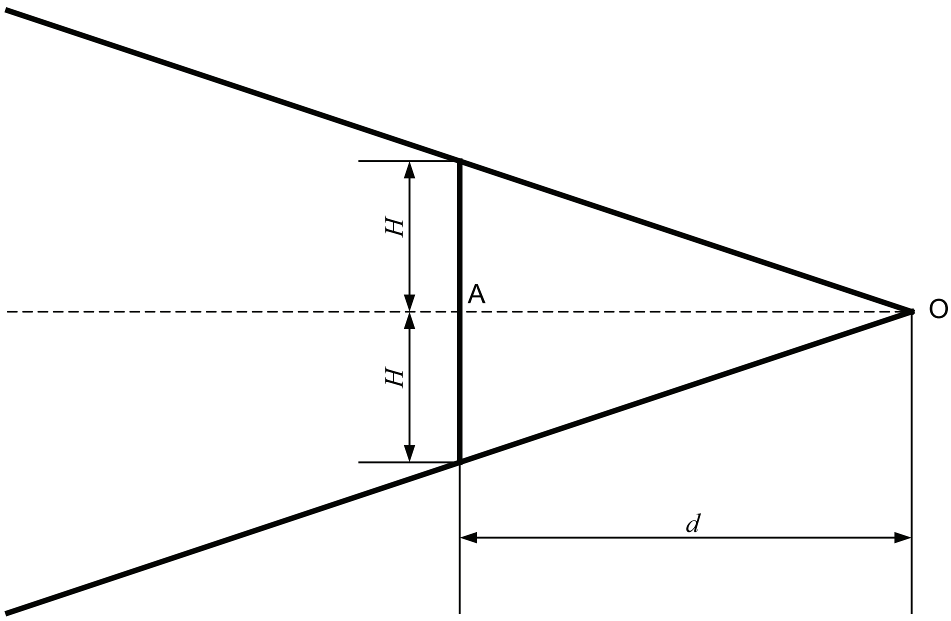


Figure 11 — Full V-mask

8.7 Fast initial response (FIR) cusum

The fast initial response (FIR) cusum is intended to reduce the ARL for shifts in the mean that it is desired to detect without significantly decreasing the on-target ARL: this with respect to the comparable ordinary decision criterion. Putting it another way, the objective is to respond more quickly to shifts while almost retaining the false alarm rate.

A comparison of the ARLs in columns 2 and 5 of Table 6 shows that the FIR scheme has a much quicker response to shifts in the mean throughout the range (29 instead of 38 at $0{,}5\sigma_e$ and 6,4 instead of 10 at $1\sigma_e$) while retaining a comparable on-target ARL, 448 compared to 465.

Table 6 also shows that the FIR scheme has a quicker response to shifts than either the semi-parabolic or snub-nosed masks throughout the range and, at the same time, has far superior on-target ARL (448 compared with 235 and 300).

With FIR, rather than accumulating from zero, a head start is given to the cusum. A convenient value for this head start is generally accepted to be half the decision interval, $h/2$.

The reasoning behind the FIR cusum is that if there is some movement before, or when, the cusum chart begins then starting the cusum part way towards the direction it is heading will hasten the signal of mean shift. On the other hand, if the process has not moved, the cusum will naturally drop back towards zero and behave just like the usual zero start cusum.

When used in conjunction with decision tabular schemes (see 8.8), the head start is often used with both the upper and lower cusums.

8.8 Tabular cusum

8.8.1 Rationale

Sometimes the primary purpose of a cusum procedure is purely to detect off-standard conditions, rather than to provide an informative visual presentation of sequential data. If so, the cusum information can be recorded purely in the form of a tabulation, as an alternative to charting. A numerical decision rule then replaces the mask used with a conventional cusum chart.

Such schemes are termed tabular cusums.

The V-mask detects changes in slope. Its decision interval, $h\sigma_e$, allows for a degree of scatter in the cusum points. The slope of the decision lines of the mask corresponds to a mean process level of "target value $\pm f\sigma_e$".

With the tabular scheme, instead of cumulating and plotting:

observation value – target value,

we separately cumulate and tabulate:

observation value – (*target value* + $f\sigma_e$),

reset the value of the cumulative sum to zero on becoming negative, for an upper cusum to detect an increase in the mean; and

accumulate and tabulate:

observation value – (*target value* – $f\sigma_e$),

reset the value of the cumulative sum to zero on becoming positive, for a lower cusum to detect a decrease in the mean.

This gives:

horizontal decision lines at "$\pm h\sigma_e$",

rather than:

decision lines with slopes "$f\sigma_e$" radiating from the datum "$h\sigma_e$", of a V-mask.

The effect, in terms of pure statistical decision-making, is precisely the same as that achieved with the comparable V-mask.

8.8.2 Tabular cusum method

The following steps are taken in setting up and interpreting a two-sided cusum decision interval scheme for a measured data characteristic having a normal distribution.

Step 1 *— Establish the cusum parameters*

a) Establish the decision interval, h.

b) Establish the slope of the decision line, f.

c) Establish the target value, T.

d) Estimate the characteristic's standard error, σ_e.

Step 2 *— Calculate the cusum criteria*

Calculate $(T + f\sigma_e)$ and $(T - f\sigma_e)$.

Step 3 *— Prepare a cusum table, for an upper tabular cusum to detect increases in the mean level, with columns*

e) Observation number.

f) Value.

g) Value – $(T + f\sigma_e)$.

h) Cusum of [Value – $(T + f\sigma_e)$].

NOTE This is a similar table to that used for a conventional cusum plot with the exception that $(T + f\sigma_e)$ replaces the T value, $f\sigma_e$, being the slope of the equivalent V-mask decision line.

Step 4 *— Prepare a cusum table, for a lower tabular cusum to detect decreases in the mean level, with columns*

As for Step 3 except:

i) Value – $(T - f\sigma_e)$.

j) Cusum of [Value – $(T - f\sigma_e)$].

***Step 5** — Enter data*

k) Enter data and perform calculations.

l) *For positive values of cusum:* Starting at zero, accumulate the column "Cusum of [Value – $(T + f\sigma_e)$]". If the cusum becomes negative at any point, reset to zero and continue at zero until the cusum again becomes positive. If the cusum touches or exceeds the decision boundary, $h\sigma_e$, an upward shift is signalled.

m) *For negative values of cusum:* Starting at zero, accumulate the column "Cusum of [Value – $(T - f\sigma_e)$]". If the cusum becomes positive at any point, reset to zero, and continue at zero until the cusum again becomes negative. If the cusum touches or falls below the decision boundary, $h\sigma_e$, a downward shift is signalled.

An example of the method is shown in Table 8 and another example of the tabular method in Annex B.

Table 8 — Example of a tabular cusum scheme

Value	Value – 11	Cusum (upper)	Value – 9	Cusum (lower)	Comments
10	–1	0	+1	0	Both cusums at zero as process is on target
10	–1	0	+1	0	
10	–1	0	+1	0	
14	+3	+3	+5	0	Process mean higher than target thence lower cusum at zero
14	+3	+6	+5	0	
3	–8	0	–6	–6	
3	–8	0	–6	–12[a]	[a] Signal of a low mean
10	–1	0	+1	–11	
10	–1	0	+1	–10	
10	–1	0	+1	–9	
10	–1	0	+1	–8	
10	–1	0	+1	–7	
17	+6	+6	+8	0	
17	+6	+12[b]	+8	0	[b] Signal of a high mean

NOTE 1 Target value = T = 10: σ_e = 2: h = 5: f = 0,5.

NOTE 2 Column 2 = Value – $(T + f\sigma_e)$ = Value – (10 + 1) = Value – 11.

NOTE 3 Column 4 = Value – $(T - f\sigma_e)$ = Value – (10 – 1) = Value – 9.

9 Cusum methods for process and quality control

9.1 The nature of the changes to be detected

9.1.1 The size of change to be detected

When designing a cusum system to monitor either a process parameter or a product characteristic, consideration should be given to the size of shift or change within the parameter or characteristic that it is important to detect. This decision will influence the shape of any "V-mask" that might be used to observe any out-of-control signals. When controlling a parameter or a characteristic, many practitioners take this as the smallest shift for which the process can be corrected. There is little point in seeking a shift smaller than this for the effect on the cusum plot is likely to create the phenomenon of "hunting" (see 9.1.5).

Changes that occur can be classified as "step", "drift" or "cyclic".

9.1.2 "Step" changes

A step change is one where the data from measurements made on a process parameter or product characteristic suddenly jump or "step" to a new level. An example of this is where a new batch of a raw material is used that differs in some way from that previously used, or where an inexperienced clerk takes over an administrative task and, until the person properly learns the tasks required, makes more errors than an experienced person. A cusum chart will identify this change by showing a significant gradient.

9.1.3 Drifting

This type of change is often associated with wear patterns of equipment or tooling but can happen where, in the human case, standards alter over time, e.g. inspection standards. The pattern would be detected by the cusum plot and depicted as an increasing (or decreasing) gradient.

9.1.4 Cyclic

A pattern that changes over time and repeats itself as a pattern is named a cyclic change. For example, it could happen in a factory where there are three work shifts and all three workers perform differently the same tasks. Since there is a given sequence of the shifts, e.g. Shift B always follows Shift A, a cyclic pattern will emerge. The cusum plots show this pattern as periods where the gradient goes in one direction followed by another where it changes back again, etc.

9.1.5 Hunting

Hunting will occur when the parameter or characteristic cannot be exactly adjusted to the desired target value and, following an out-of-control signal, the adjustment made takes the location of the parameter or characteristic to the other side of the target. The cusum plot then develops a gradient in the opposite direction and eventually the signal is received to reverse the adjustment which had previously been done.

In this way a "zigzag" pattern will be detected on the cusum plot. Clearly this would be a most unsatisfactory situation and should be avoided by careful selection of the original "target value" and the subsequent minimum adjustment to be sought. See 9.3.1 *Step 13 c)* for more counteracting hunting.

9.2 Selecting target values

9.2.1 General

The correct selection of the target value is of prime importance in the setting up of a cusum scheme.

A target value that is in between two possible values will create "hunting" as described in 9.1.5.

9.2.2 Standard (given) value as target

The simplest target value to assign is a "given" or "preassigned" value. When this option is selected, the target value is often set equal to some specification value such as a nominal or mid-tolerance value. These are found on specification documents or drawings when the application is based around engineering. If the cusum application is non-manufacturing, the given target might be some performance level such as the expected time taken to process an invoice or the budgeted monthly expenditure for a department within a company.

It is possible for the target itself to vary. For example, if the sales of ice-cream were to be monitored by a cusum chart on a monthly basis, the target value to be used each month is likely to be different according to the time of year. It may be anticipated that more ice-cream will be sold during the summer months than the winter months and so a different target may be used for each month. Failure to recognize the sales pattern and instead to use a constant value per month would lead to a misleading plot on the cusum graph paper. The cusum value is likely to rise during one period of the year and then fall in another. If the target were varied, the cusum would be better equipped to indicate whether there was any significant change in the level of sales of ice-cream with the "seasonality" removed.

Inappropriate target values can result in the phenomenon “hunting” described in 9.1.5 and careful consideration should therefore be exercised when choosing a target of this sort.

The norm is for a given value approach to be used when monitoring the mean or average level (location) for the parameter or characteristic in question. Although the same approach could be used to set the target value for monitoring the variability (dispersion) such as a subgroup standard deviation or range, this practice is not recommended within this part of ISO 7870. It is preferred to proceed using the guidance contained within 9.2.3 and later clauses.

9.2.3 Performance-based target

The target value can be set from current performance levels. This approach is compatible with performance-based control charting where the controls are set according to the recent historical performance of a process parameter or product characteristic.

For monitoring the location or dispersion, it is essential to capture data in a “trial” or “data collection” phase. Such a period should be long enough for the inherent variability to be fully observed and this will be a matter of judgement. Typically, the trial should be long enough to provide for 25 points on the cusum plot. From these data, estimates should be made of the mean value and the standard deviation.

Once determined, these target values should be used for the calculation of the cusum(s) but might require alteration at some later time if the cusum indicates a change in level. If it is not possible to make any process adjustment following such a change, or if the new level is acceptable, the only action that may be taken is to modify the target value. This is usually done after evaluating what the new level is from the most recent data and making this the new target. Thereafter the cusum will monitor the parameter or characteristic with reference to its new target value.

9.3 Cusum schemes for monitoring location

9.3.1 Standard schemes

See Figure 12.

Step 1 — *Determine the subject for cusum charting*

Determine the process parameter or product characteristic to be monitored.

NOTE 1 This might be an instruction from a customer or a key process parameter or a significant product characteristic. The subject might also be identified during a problem solving exercise.

Step 2 — *Determine the subgroup size*

The determination of a rational subgroup for the cusum chart is an identical thought process that would be used to construct any Shewhart chart.

If a process parameter is the chosen cusum subject, the most appropriate subgroup size is usually one. This is because the parameters, e.g. the temperature of a solution or the pressure in a vessel, are not likely to vary over a short time. Taking several consecutive repeat measures one after the other is unlikely to show any difference in the measurements. This would lead to technical problems when determining the standard deviation and the correct set-up of the cusum mask.

If the data are genuinely one-at-a-time, such as the sales value for a particular month, then the rational subgroup size will be one.

When a product characteristic has been chosen, the rational subgroup size is often greater than one, and typically five. Common sense should be exercised here. The subgroup size is so selected to represent the random variation in the process.

Step 3 — *Select cusum scheme*

Table 9 indicates a set of standard schemes that provide for a range of typical requirements for a cusum scheme. The table provides two basic schemes, one that gives rather long average run lengths (ARLs) at zero shift, i.e. a CS1 scheme, and another which has shorter ARL, i.e. a CS2 scheme. In other words, the CS2 scheme will detect the shift in process level quicker than the corresponding CS1 scheme, but at the expense of more “false signals”. Whoever is responsible for the selection of the scheme must determine which scenario is the more important and then select the appropriate scheme. Table 10 illustrates the differences in performance of these standard schemes.

Table 9 — Standard cusum schemes for subgroup means

Important shift in the mean to be detected[a]	CS1 schemes		CS2 schemes	
	h	f	h	f
i) < 0,75σ_e	8,0	0,25	5,0	0,25
ii) 0,75 to 1,50σ_e	5,0	0,50	3,5	0,50
iii) > 1,50σ_e	2,5	1,00	1,8	1,00

NOTE 1 CS1 schemes give average run lengths, L_0, in the region of 700 to 1 000 when the actual shift is zero.

NOTE 2 CS2 schemes give average run lengths, L_0, in the region of 140 to 200 when the actual shift is zero.

[a] For individual results (subgroup size = 1), σ_e represents the standard deviation. When the subgroup size is more than one, σ_e represents the standard error of the mean.

Once the decision has been made concerning which of CS1 or CS2 to select, the next decision is about the size of the important shift. Three typical levels of shift are provided for in the table. Depending on this selection the values for h and f can be read from the table.

If it is unclear which scheme should be selected, custom and practice indicate that a good starting scheme is to select CS1 scheme ii), i.e. h = 5,0 and f = 0,50.

Table 10 — Comparison of performance of standard cusum schemes for subgroup means

Values are average run lengths (ARLs)

Shift in the mean from target value (in units of σ_e)[a]	CS1 schemes			CS2 schemes		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
0,00	730,0	930,0	715,0	140,0	200,0	170,0
0,75	16,4	17,0	27,0	10,5	11,5	15,0
1,00	11,4	10,5	13,4	7,4	7,4	8,8
1,50	7,1	5,8	5,4	4,7	4,3	4,0

NOTE The values given are ARLs. The reader should be aware that the actual run length taken to detect an actual shift will vary and can be shorter than or longer than the ARL. When it is of particular interest, the reader should examine the distribution of run lengths for particular shifts from target to know the expected range of run lengths that might be experienced.

[a] For individual results (subgroup size = 1), σ_e represents the standard deviation. When the subgroup size is more than one, σ_e represents the standard error of the mean.

Whatever scheme is selected, the values for these parameters should be multiplied by the estimated variability, σ (or σ_e), to determine the actual size and shape of the mask. This is described in *Step 8*.

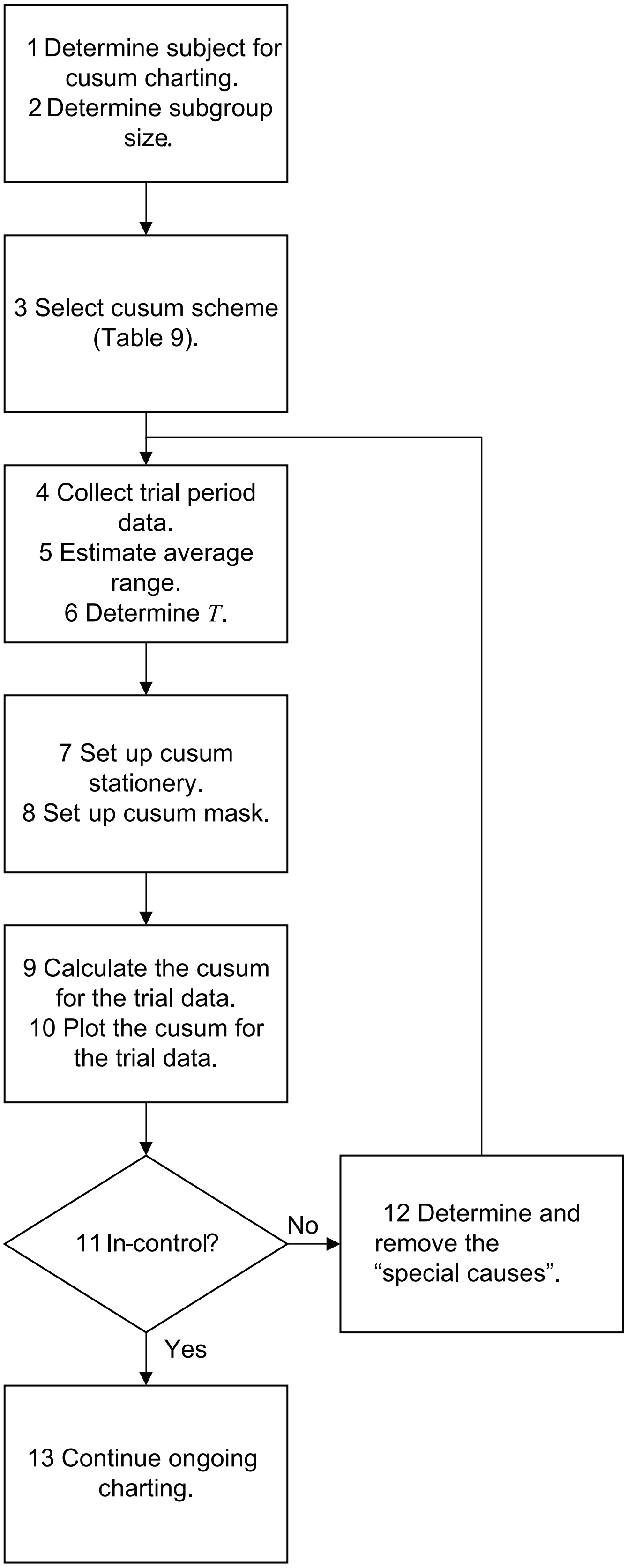


Figure 12 — Cusum set-up protocol

Step 4 — *Collect trial period data*

As indicated in 9.2.3, data should be captured that will describe the nature of the variability in the process, so that the cusum scheme can be properly “tuned”, and to assist in establishing the target value if necessary.

Determine a trial period during which all the sources of process variation will be observed. This should be long enough or the sampling frequency high enough to produce at least 25 subgroups of data.

Take care not to introduce extra sources of variation, e.g. process adjustments, during this period as this will distort the variation pattern. If there is an interruption of data collection, a decision should be made as to whether the trial period requires to be done again or whether sufficient data were generated during the shortened trial period. In general, if the number of subgroups gathered was 20 or more and if it is judged that all potential sources of variation were observed within the 20 subgroups, then this number of subgroups will be satisfactory and the trial period closed. The data from the trial should then be used to establish the levels of variability the cusum scheme will operate under. This is described in *Steps 5* and *6*.

Step 5 — *Estimate σ_e from the trial period data*

a) General

The following paragraphs outline the method for estimating σ_e. Special circumstances might occur where a different approach might be required. This different approach might require σ_e to be evaluated by looking at the standard deviation between the subgroup means.

b) Subgroup sizes more than one ($n > 1$)

i. Calculate the range (largest minus smallest observation) of each subgroup.

ii. Calculate the average range ($\overline{R}$) of all of the subgroup ranges.

iii. Estimate the within-subgroup standard deviation (σ_0) by dividing the average range by the appropriate d_2 value taken from Table 11.

iv. Estimate σ_e by dividing σ_0 by the square root of the subgroup size, i.e. $\sigma_e = \sigma_0 / \sqrt{n}$.

Table 11 — d_2 factor for estimating the within-subgroup standard deviation from within-subgroup range

Subgroup size, n [a]	d_2
2	1,128
3	1,693
4	2,059
5	2,326
6	2,534
7	2,704
8	2,847
9	2,970
10	3,078
NOTE For subgroups greater than 10 other methods may be more efficient at estimating the within-subgroup standard deviation.	
[a] Values of d_2 exist for $n > 10$. See ISO 7870-2 or other textbooks or standards.	

The within subgroup standard deviation (s) method can be used as an alternative to the subgroup range. The average subgroup standard deviation, $\overline{s}$, must be calculated instead of $\overline{R}$ and σ_0 is estimated by $\overline{s}/c_4$. Table 18 has values of c_4.

c) Subgroup size is one (n = 1)

The approach taken to estimate σ_e is to use the method of successive difference (sometimes called a moving range of two observations).

The data gathered during the trial period should be put in the sequence they were collected. The range (difference) between the first and second results should be calculated, then the range between second and third, etc. If there are k subgroups, there will be $k - 1$ ranges. Calculate the average of these ranges ($\overline{R}$).

The estimate of σ_e can then be found by dividing the average range by 1,128.

Step 6 — *Determine the target value, T*

As described in 9.2 the target value may be a given value or a performance-based value determined from data.

a) Given value

The value of the target is a specified value. It may come from a specification document or a drawing and may be a nominal size, in the case of a product characteristic, or some expected level of performance given by management in the case of a non-manufacturing process.

b) Performance-based value

Here, the target value should be determined from the data obtained during the trial period.

i. Calculate the mean value ($\overline{x}$) for each subgroup.

ii. Calculate the average $\left(\overline{\overline{x}}\right)$ of these means.

iii. Assign $\overline{\overline{x}}$ as the target, T.

Step 7 — *Set up the cusum stationery*

a) General

Clause 5 provides guidance on setting up cusum stationery.

b) Cusum table

Set up a suitable table where the cusum calculations can be written down and read from. Part of such a table is shown in Table 12.

Table 12 — Table for cusum calculations

Subgroup number	$\overline{x}$	$\overline{x} - T$	**Cusum value**, C
etc.			

If the subgroup size is one, replace $\overline{x}$ with x, the individual result in the table.

c) Cusum graph paper

Select graph paper with convenient intervals between the grid lines. The choice will be influenced by the intended use of the paper, e.g. for a wall or public display.

Select a suitable scale. The scale will be influenced by the location of the graph. For example, for graphs intended for wall or public displays the interval between the subgroup numbers on the horizontal axis might be 10 mm, whereas for a plot intended for desk use only the interval might be 5 mm.

A suitable interval for the cusum (C) axis is given by making the same interval selected for the horizontal axis approximately equal to $2\sigma_e$, rounding as appropriate. This scaling is unlikely to artificially "flatten" a significant trend or exaggerate an insignificant one.

Mark the centre point of the cusum axis 0 and draw a bold horizontal line across the graph paper at this point. Mark off the vertical cusum scale on the graph paper.

An example of such paper is shown in Figure 13.

Step 8 *— Set up the cusum mask*

Clause 8.2.1 describes the geometry of the standard cusum mask and Figure 5 illustrates the components of the mask and how it is to be scaled.

The values of h, f and σ_e should be determined as described in this subclause.

a) Calculate $H = h\sigma_e$

b) Calculate $F = f\sigma_e$

The mask should be drawn according to the scale selected for the cusum paper. This is essential if the mask is to be used correctly to make decisions about whether a change of the predetermined size has happened.

NOTE 2 Some masks are made from a see-through material such as acetate. The mask outline can be traced on using indelible ink. Sometimes the mask can be cut out from a piece of card, again with the values of H and F marked out using the scale of the cusum paper.

NOTE 3 Computer programs exist which will display a cusum plot with the mask drawn over it, all automatically scaled.

Step 9 *— Calculate the cusum for the trial data*

Using the target value determined in *Step 6* and a table similar to that shown in Table 12, calculate the cusum values for the trial data.

Step 10 *— Plot the cusum for the trial data*

The tabulated cusum values generated as mentioned above should be plotted onto suitable graph paper similar to that shown in Figure 13, with the plot beginning at the left and extending in a rightwards direction. Join up all plotted points as this makes any trend easier to see and later, when the mask is superimposed, it helps identify out-of-control signals.

Step 11 *— Review the cusum plot of the trial data for out-of-control*

Superimpose the mask over the cusum plot.

This is done by locating the "lead point" indicated in Figure 7 a) over the last plotted cusum value, taking care to maintain the centreline of the mask parallel with the zero axis on the paper. This ensures that the mask is correctly orientated.

Any point going outside the arms (decision lines) of the mask indicates the presence of an out-of-control process, even if the offending point is not the last plotted point and even if later plotted points return within the arms of the mask. See Figure 7 b).

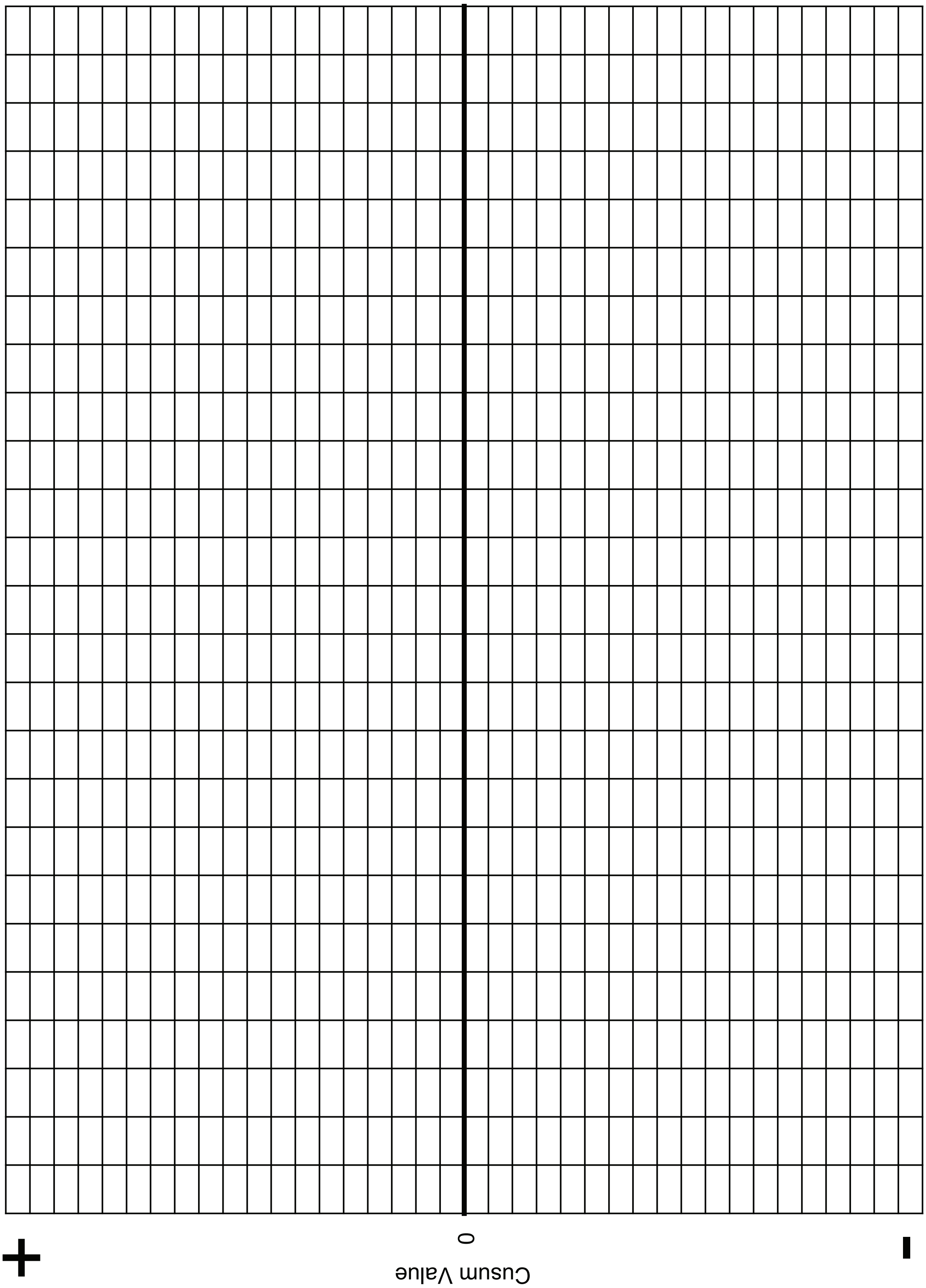


Figure 13 — Example cusum paper

Step 12 — *Identify and remove "special causes"*

a) General

It is essential to investigate any out-of-control points on the cusum plot and identify the "special cause".

b) "Special cause(s)" identified and prevented from reoccurrence

Once the special cause has been identified and steps have been taken to prevent such a future event, the values for the target and the standard error (or standard deviation) might require revision. If only one out-of-control point was observed and the cause has been satisfactorily handled, then the values previously assigned to the target and the standard error or standard deviation may be revised using the original trial period data less the data for the out-of-control subgroup. Revise the calculations for the scaling of the cusum graph paper and those for the dimensions of the mask and rescale the paper and the mask as needed.

If there are several out-of-control points in the trial data, it indicates rather more problems with the process and it is recommended the process be reviewed, corrected and then a fresh trial period be initiated and the cusum set-up protocol repeated with these new data.

c) "Special cause(s)" identified but not prevented from reoccurrence

There are occasions when the special cause is not preventable in the future due to uneconomic circumstances or technical considerations.

In such circumstances, the cusum parameters are based on all of the trial data and used for ongoing monitoring. In other words, these special causes are to be regarded as part of the random variation of the process.

d) "Special cause(s)" unidentified

Some "special causes" might remain unidentified. This is always very unsatisfactory as it prevents process improvement. Every effort should be made to investigate the special causes and the use of other statistical and problem solving techniques should be used to do this. Techniques such as the statistical design of experiments are particularly powerful in this regard.

If the special cause(s) remain unidentified, the steps to take are as those contained in c) above.

Step 13 — *Continue ongoing charting*

a) General

If the trial data provided an "in-control" situation or when the new data are collected after the satisfactory resolution of "special causes", the cusum is ready for ongoing monitoring of the process parameter or product characteristic. The scaling of the paper, the mask parameters and the mask scaling are now used to monitor the data from future subgroups.

If future out-of-control signals appear, it is essential to investigate and to decide what actions to take on the process. These can range from a process adjustment, such as on a tool, to the adoption of a new target value if the process has moved to a more desired location.

b) Process adjustments

The amount of adjustment required can be determined from the cusum plot.

i. Determine the value of the cusum for the last plotted point.

ii. Determine the value of the cusum at the place where the cusum plot went out of the action arms on the mask. In the case of multiple violations, take the most recent, i.e. the out-of-control point located nearest the lead point on the mask.

iii. Calculate the difference in the cusum value between these two points.

iv. Count the number of plotted points between the lead point and the out-of-control point.

v. Divide the difference calculated in iii) by the number of plotted points counted in iv), i.e. calculate the local gradient of the cusum plot. This is the estimate of the shift from the target value that the process has undergone.

vi. An adjustment of the same magnitude may be made to the process. Record on the cusum plot the changes made.

vii. Return the cusum value to zero and continue monitoring.

NOTE 4 Rather than carry out the calculations described here, it is possible, due to the geometric nature of the mask and cusum paper, to generate “look-up” graphs to read the amount of shift from target. An example of this can be seen in Figure 14.

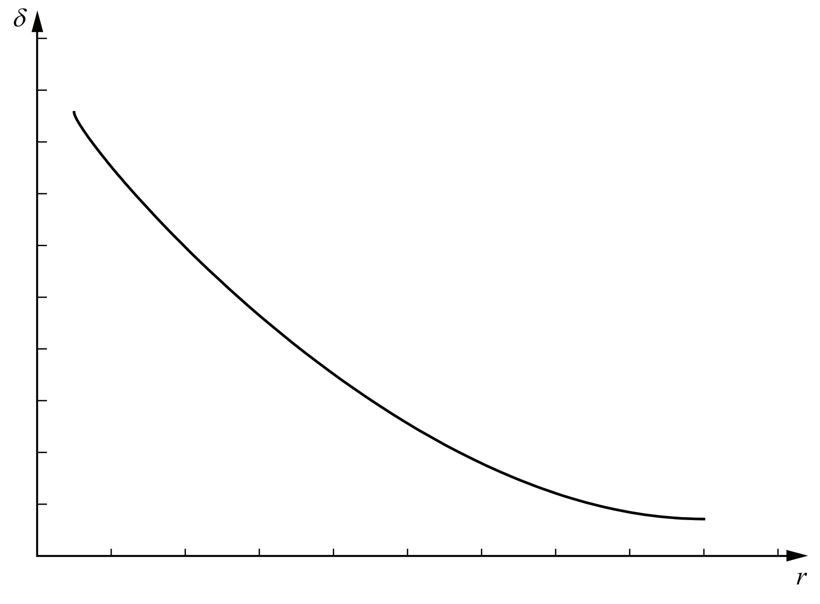


Figure 14 — Schematic example of a look-up graph for the shift from target

c) Process adjustments — anti-hunting

The concept of "hunting" was introduced in 9.1.5. The application of cusum methods suggests that adjusting a process for 100 % of the calculated shift from target can introduce hunting. Following such an adjustment the cusum later indicates the need to further adjust the process, but in the opposite direction, i.e. as if the adjustment were too much. Should the process now be adjusted back by an amount equal to the latest indicated shift there can later be yet more adjustments in the opposite direction. This phenomenon is known as "hunting".

It can be caused by the short-term cusum gradient exaggerating the real shift that might be cured by a modification to the mask design, or by the actual adjustment needed being impossible to achieve.

One practical solution is to make a process adjustment of less than 100 % of the implied shift. This lesser percentage is known as an "anti-hunting factor". This lessens the tendency of the cusum to "hunt". Custom and practice together with some research have indicated that an anti-hunting factor of 75 % works very well. Therefore it is recommended, at least in the beginning, that only 75 % of the implied shift be adjusted for.

Graphs such as that shown in Figure 14 can be drawn in a manner that incorporates the anti-hunting factor. This can be desirable if numerous people are involved with operating the cusum system.

An alternative anti-hunting factor is $r/(r + 1)$, where r is the number of points between the lead point and the out-of-control point. Using this factor, the size of process adjustment, φ, is calculated as:

$$\varphi = \frac{C_r}{r+1}$$

where C_r is the difference in the cusum value between the lead point and the out-of-control point. When a step change occurs, this method has a good performance to prevent hunting.

9.3.2 Standard schemes — Limitations

The basic cusum schemes described in 9.3.1 provide good starting positions for most applications and in many cases will not need any further alteration. For a few applications though, it might be noticed after some time that the basic scheme selected could be improved either because the ARL to detect a shift of an important size is too large or the frequency of "false alarms" is too high.

9.3.3 "Tailored" cusum schemes

NOTE The design of a specific cusum scheme requires more knowledge and input than the basic schemes described in 9.3.1. Anyone requiring a cusum scheme might wish to consult with a specialist to help with the design.

a) Determine the size of the (important) shift from the target to be detected, δ.

b) Estimate the standard error, σ_e, (or the standard deviation if the subgroup size is one) as described earlier.

c) Specify the intended ARL when the shift is of the size given in 9.3.1 *Step 3*, L_δ.

d) Specify the intended ARL when the shift is zero, i.e. the frequency of "false alarms", L_0.

e) Calculate the standardized shift, $\Delta = \delta/\sigma_e$.

f) Enter the graph in Figure 15 and read off h for the calculated value of Δ and taking account of the values for L_δ and L_0.

g) The value for f may also be read from the graph corresponding with the calculated value for Δ.

h) Modify the cusum mask according to the new values for h and f as indicated earlier.

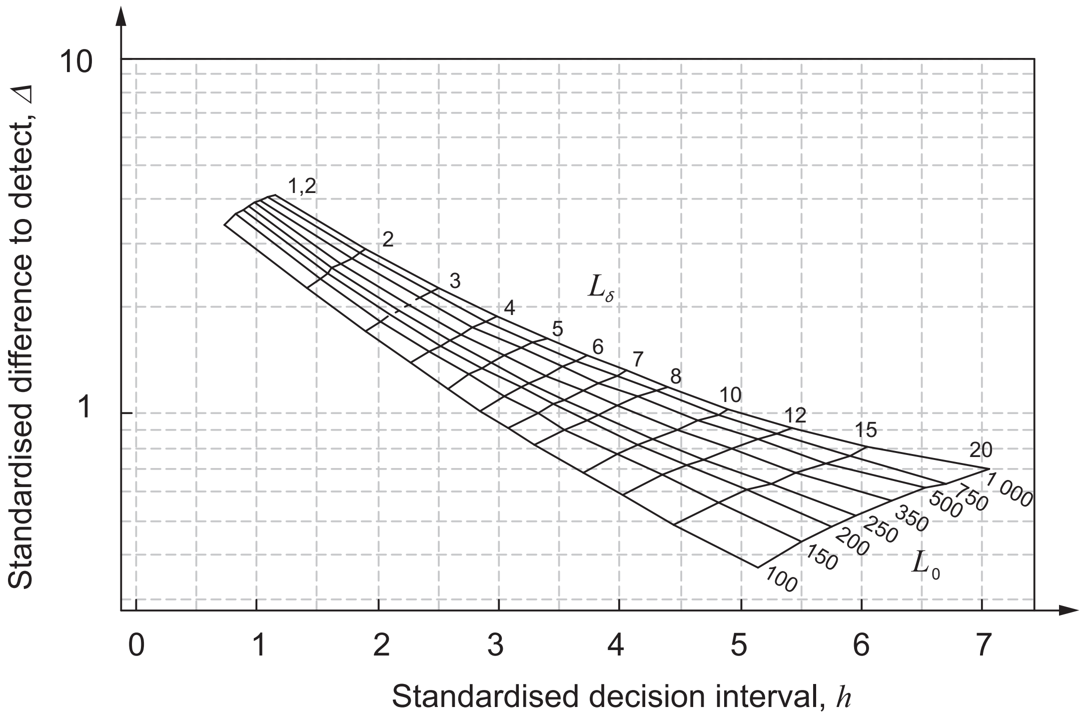


Figure 15 — Intercept chart for mask parameters (assuming a normal variable)

9.4 Cusum schemes for monitoring variation

9.4.1 General

In addition to monitoring the process location, it is essential to monitor the process variation, which will in most cases be the short-term variation.

The two most appropriate measures of variation to use are the within-subgroup range and the within-subgroup standard deviation. A choice should be made as to which will be selected. That decision will depend on the ease of calculation and the level of comprehension of the measures on the part of those involved with the calculation. Many involved with operating control charts select the range as the preferred measure due to its ease of calculation and its implicit simplicity, and for the subgroup size often selected, e.g. five samples, the efficiency of the range is nearly as good as the standard deviation.

If the subgroup size is one, the measure used should be the range based upon the difference between successive results.

9.4.2 Cusum schemes for subgroup ranges

The following steps should be followed to establish a suitable scheme to monitor process variation using the within-subgroup range. Some of the steps will have been completed if the cusum scheme for monitoring the mean value has been done.

Step 1 *— Determine the subject for cusum charting*

Refer to 9.3.1 *Step 1*.

Step 2 — *Determine the subgroup size*

Refer to 9.3.1 *Step 2*.

Step 3 — *Select cusum scheme for range*

Table 13 specifies a set of standard schemes that provide for typical requirements for a cusum scheme for range. As described in 9.3.1, the table provides two basic schemes, one that gives rather long average run lengths (ARLs) at the expected level of variation, i.e. a CS1 scheme, and another which has shorter ARLs, i.e. a CS2 scheme. The CS2 scheme will detect the shift in process level quicker than the corresponding CS1 scheme, but at the expense of more "false signals". Table 14 illustrates the differences in performance of these standard schemes.

Table 13 — Standard cusum schemes for subgroup ranges

Subgroup size	CS1 scheme		CS2 scheme	
	h	f	h	f
2	2,50	0,85	2,50	0,55
3	1,75	0,55	1,75	0,35
4	1,25	0,50	1,25	0,30
5	1,00	0,45	1,00	0,30
6	0,85	0,45	0,85	0,30
7	0,70	0,45	0,70	0,30
8	0,55	0,40	0,55	0,25
9	0,55	0,40	0,55	0,25
10	0,50	0,35	0,50	0,25

NOTE 1 CS1 schemes give average run lengths, L_0, in the region of 600 to 1 000 when the process operates at the expected level of variability.

NOTE 2 CS2 schemes give average run lengths, L_0, in the region of 150 to 210 when the process operates at the expected level of variability.

Select either the CS1 or CS2 scheme. The same selection criteria used in the selection of the scheme for subgroup means should be applied. If a longer ARL is required when no change has occurred, select a CS1 scheme; otherwise, select the CS2 scheme.

Whatever scheme is selected, the values for these parameters should be multiplied by the estimated variability $\overline{R}$ to determine the actual size and shape of the mask. This is described in *Step 8*.

Step 4 — *Collect trial period data*

The instructions given for location also apply here.

Step 5 — *Estimate $\overline{R}$ from the trial period data*

Calculate $\overline{R}$ using one of the methods described in 9.3.1 *Step 5*.

Table 14 — Comparison of performance (ARL) of standard cusum schemes for subgroup ranges

Subgroup size	Actual process variability level	CS1 scheme	CS2 scheme
2	$\bar{R}$	779,0	170,0
	$2\bar{R}$	7,2	5,5
	$4\bar{R}$	2,3	2,1
3	$\bar{R}$	893,0	196,0
	$2\bar{R}$	4,5	3,6
	$4\bar{R}$	1,6	1,5
4	$\bar{R}$	918,0	157,0
	$2\bar{R}$	3,3	2,7
	$4\bar{R}$	1,3	1,2
5	$\bar{R}$	771,0	179,0
	$2\bar{R}$	2,7	2,3
	$4\bar{R}$	1,2	1,1
6	$\bar{R}$	942,0	204,0
	$2\bar{R}$	2,4	2,0
	$4\bar{R}$	1,1	1,1
8	$\bar{R}$	893,0	162,0
	$2\bar{R}$	2,0	1,7
	$4\bar{R}$	1,0	1,0
10	$\bar{R}$	635,0	184,0
	$2\bar{R}$	1,7	1,5
	$4\bar{R}$	1,0	1,0
NOTE The values given are ARLs. The reader should be aware that the actual run length taken to detect an actual change will vary and might be shorter than or longer than the ARL. When it is of particular interest, the reader should examine the distribution of run lengths for particular shifts from target to know the expected range of run lengths that might be experienced.			

***Step 6 —** Determine the target value, T*

a) Given value

In statistical quality or process control, the most common method of setting the target range is as described in b) below. However, there can be occasions when it is preferred to set the target from some assumed level. If this is so, the target range is made equal to a given value for range.

If the variability is described through a given standard deviation, the target range may be calculated as $T = d_2\sigma$ where d_2 may be taken from Table 11, the value dependent on the subgroup size to be used.

b) Performance-based value

From the data obtained during the trial period set the target range equal to $\bar{R}$.

***Step 7** — Set up the cusum stationery*

Set up a cusum table (or add to the existing cusum table) and develop cusum graph paper as described in 9.3.1 *Step 7*.

The cusum graph paper on which the cusum for range is to be plotted might require a different scale to that chosen to monitor the mean value. A suitable scaling can be obtained using the following calculation, rounding up or down to the nearest convenient value.

The cusum scale interval for range is $a\overline{R}$, where a is taken from Table 15.

Table 15 — Cusum graph paper for range scale coefficients

Subgroup size	a
2	1,50
3	1,00
4	0,85
5	0,75
6	0,65
8	0,55
10	0,50

***Step 8** — Set up the cusum mask*

Using the values of h and f chosen in *Step 3*, calculate:

a) $H = h\overline{R}$; and

b) $F = f\overline{R}$.

Construct the mask using the calculated values of H and F and scaling the mask according to the scale selected for the cusum graph paper.

***Step 9** — Calculate the cusum for the trial data*

Using the target value determined in *Step 6* and using a table similar to that shown in Table 12, calculate the cusum values for range for the trial data.

***Step 10** — Plot the cusum for the trial data*

Plot the cusum for range onto the cusum graph paper for range as described in 9.3.1 *Step 7* and *Step 10*.

***Step 11** — Review the cusum plot of the trial data for out-of-control*

Review the cusum plot as described in 9.3.1 *Step 11*.

***Step 12** — Identify and remove "special causes"*

a) General

It is essential to investigate any out-of-control points on the cusum plot and identify the "special cause".

If it becomes necessary to revise the target range in accordance with one of the following subclauses, then it will also become necessary to revise the mask and possibly the cusum graph paper for the control of the mean value.

b) *"Special cause(s)" identified and prevented from reoccurrence*

Once the special cause has been identified and steps have been taken to prevent such a future event, the value for the target range might require revision. If only one out-of-control point was observed and has been satisfactorily handled as described in a), then the values previously assigned to the target may be revised using the original trial period data by eliminating the data for the out-of-control subgroup. Revise the calculations for the scaling of the cusum graph paper and those for the dimensions of the mask and rescale the paper and the mask as needed.

If there are several out-of-control points in the trial data, it indicates rather more problems with the process. It is recommended that the process be reviewed, corrected and then a fresh trial period be initiated and the cusum set-up protocol repeated with this new data.

c) *"Special cause(s)" identified but not prevented from reoccurrence*

There are occasions when the special cause is not preventable in the future due to uneconomic circumstances or technical considerations.

In such circumstances, the cusum parameters are based on all of the trial data and used for ongoing monitoring. In other words, these special causes are to be regarded as part of the random variation of the process.

d) *"Special cause(s)" unidentified*

If the special cause(s) remain unidentified, the steps in c) should be followed.

This is always very unsatisfactory as it inhibits process improvement. Every effort should be made to investigate the special causes.

Step 13 **—** *Continue ongoing charting*

a) *General*

Continue charting as prescribed in 9.3.1 *Step 13*.

b) *Process actions*

As in the case of monitoring the mean value, if an out-of-control signal is observed, then the amount of change occurring may be estimated from the cusum gradient. In this case, the interpretation is of how much the variability, here expressed in terms of $\bar{R}$, has altered.

If the direction of the cusum indicates an increase in range, the reaction, in the case of some equipment or machinery, might be to call for maintenance engineers to repair the equipment. If this proves successful, the actions taken should be recorded, the value of the cusum reset to zero and the process allowed to continue. If the process has returned to its previous level of variability, the cusum will now operate "in-control".

If the direction of the cusum indicates a reduction in range, it will usually be regarded as a good event and the special cause should be identified and steps taken to preserve it. If this is successful, the masks for both mean and range (and possibly the graph papers) should be adjusted to reflect the new situation. The target range should also be reassessed to a new lower value. The cusum for range should then be rezeroed before plotting is continued.

It would be unnecessary to rezero the cusum for the mean, but, from a review made using the revised mask for the mean of the previously plotted points for the period, it is now known that the range was really lower. New out-of-control points can now be observed for the mean plot.

c) *Variation estimate — anti-hunting*

As in the case of the cusum for the location, if it is considered necessary to adopt anti-hunting measures, then the value of 75 % of the response is recommended from practice. Therefore, only 75 % of the indicated change in the target range should be taken.

9.4.3 Cusum schemes for subgroup standard deviations

The procedure for setting up a cusum scheme for monitoring standard deviations is very similar to that for monitoring subgroup ranges. Consequently, only changes from 9.4.2 are given in this subclause and it should be read in association with the whole content of 9.4.2.

The schemes detailed in this clause for monitoring subgroup standard deviations depend on there being more than one observation per subgroup. A range-based method of successive difference is preferred to monitor variation if the data to be collected are one-at-a-time values, e.g. monthly sales figures.

Step 3 — *Select cusum scheme for standard deviation*

Table 16 specifies a set of standard schemes that provide for typical requirements for a cusum scheme for monitoring the standard deviation. The table provides two basic schemes, one that gives rather long average run lengths (ARLs) at the expected level of variation, i.e. a CS1 scheme, and another which has shorter ARLs, i.e. a CS2 scheme. A CS2 scheme will give a few more “false signals” than a CS1 scheme but will detect an important change a little more quickly than the corresponding CS1 scheme. Table 17 illustrates the differences in performance of these standard schemes.

Table 16 — Standard cusum schemes for subgroup standard deviations

Subgroup size	CS1 scheme		CS2 scheme	
	h	f	h	f
2	2,00	0,50	2,00	0,25
3	1,60	0,35	1,60	0,15
4	1,15	0,35	1,15	0,20
4	1,15	0,35	1,15	0,20
5	0,90	0,35	0,90	0,20
6	0,80	0,32	0,80	0,20
7	0,70	0,30	0,70	0,20
8	0,60	0,30	0,60	0,20
9	0,55	0,30	0,55	0,20
10	0,50	0,30	0,50	0,20
12	0,40	0,30	0,40	0,20
15	0,35	0,27	0,35	0,18
20	0,30	0,23	0,30	0,16

NOTE 1 CS1 schemes give average run lengths, L_0, in the region of 700 to 1 000 when the process operates at the expected level of variability.

NOTE 2 CS2 schemes give average run lengths, L_0, in the region of 150 to 200 when the process operates at the expected level of variability.

Select either the CS1 or CS2 scheme. The same selection criteria used in the selection of the scheme for subgroup means should be applied. If a longer ARL is required when no change has occurred, select a CS1 scheme; otherwise, select the CS2 scheme.

Whatever scheme is selected, the values for these parameters should be multiplied by the estimated variability ($\hat{\sigma}_0$) to determine the actual size and shape of the mask. This is described later in *Step 5*.

Table 17 — Comparison of performance (ARL) of standard cusum schemes for subgroup standard deviations

Subgroup size	Actual process variability level	CS1 scheme	CS2 scheme
2	σ_0	920,0	185,0
	$2\sigma_0$	7,4	5,6
	$4\sigma_0$	2,3	2,1
3	σ_0	920,0	155,0
	$2\sigma_0$	4,4	3,7
	$4\sigma_0$	1,6	1,5
4	σ_0	840,0	180,0
	$2\sigma_0$	3,2	2,6
	$4\sigma_0$	1,3	1,2
5	σ_0	820,0	155,0
	$2\sigma_0$	2,6	2,2
	$4\sigma_0$	1,1	1,1
6	σ_0	850,0	190,0
	$2\sigma_0$	2,2	1,9
	$4\sigma_0$	$< 1,1$	$< 1,1$
8	σ_0	720,0	180,0
	$2\sigma_0$	1,7	1,6
	$4\sigma_0$	1,0	1,0
10	σ_0	930,0	200,0
	$2\sigma_0$	1,5	1,4
	$4\sigma_0$	1,0	1,0
12	σ_0	840,0	170,0
	$2\sigma_0$	1,3	1,2
	$4\sigma_0$	1,0	1,0
15	σ_0	860,0	170,0
	$2\sigma_0$	1,2	1,1
	$4\sigma_0$	1,0	1,0
The values given are ARLs. The reader should be aware that the actual run length taken to detect an actual change will vary and might be shorter than or longer than the ARL. When it is of particular interest, the reader should examine the distribution of run lengths for particular shifts from target to know the expected range of run lengths that might be experienced.			

Step 5 — *Estimate σ_0 from the trial period data*

a) For each subgroup calculate the within-subgroup standard deviation (s).

b) Calculate the average within-subgroup standard deviation ($\overline{s}$).

Estimate the within-subgroup standard deviation, $\hat{\sigma}_0 = \overline{s}/c_4$ where c_4 can be read from Table 18.

Table 18 — c_4 factor for estimating the within-subgroup standard deviation

Subgroup size, n [a]	c_4
2	0,797 9
3	0,886 2
4	0,921 3
5	0,940 0
6	0,951 5
7	0,959 4
8	0,965 0
9	0,969 3
10	0,972 7
12	0,977 6
15	0,982 3
20	0,986 9

[a] Values of c_4 exist for n $>$ 20. See ISO 7870-2 or other textbooks or standards.

Step 6 — *Determine the target value, T*

a) Given value

In statistical quality or process control, the most likely method of setting the target within-subgroup standard deviation is as described in b) below. However, there might be occasions when it is preferred to set the target from some assumed level of σ_0. If this is so, the target within-subgroup standard deviation is calculated as $T = c_4\sigma_0$, where c_4 is taken from Table 18.

b) Performance-based value

From the data obtained during the trial period, set the target within-subgroup standard deviation equal to $\overline{s}$.

Step 7 — *Set up the cusum stationery*

Set up a cusum table (or add to the existing cusum table) and develop cusum graph paper.

The cusum graph paper on which the cusum for within-subgroup standard deviation is to be plotted might require a different scale to that chosen to monitor the mean value. A suitable scaling may be obtained using the following calculation, rounding up or down to the nearest convenient value.

The cusum scale interval for within-subgroup standard deviation is $a\sigma_0$, where a is taken from Table 19.

Table 19 — Cusum graph paper scale coefficients for within-subgroup standard deviation cusum

Subgroup size	a
2	1,50
3	1,00
4	0,85
5	0,75
6	0,65
8	0,55
10	0,50
15	0,40
20	0,35

Step 8 — *Set up the cusum mask*

Using the values of h and f chosen in *Step 3*, calculate:

a) $H = h \times \hat{\sigma}_0$; and

b) $F = f \times \hat{\sigma}_0$.

Construct the mask using the calculated values of H and F and scaling the mask according to the scale selected for the cusum graph paper.

9.5 Special situations

9.5.1 Large between-subgroup variation

In certain circumstances, it might become important to allow some of the between-subgroup variation of means to be considered as part of the random variation. An example of this is small fluctuations in the mean detected by the cusum chart, but there are no plans to eliminate them. To prevent the cusum continuously showing out-of-control, these small fluctuations should be included in the estimate of the variability.

Calculate the between-subgroup mean standard deviation (known as the standard error of the mean), $s_{\bar{x}}$. These may be from the trial period data or some other period of data which is representative of the variability. Use the value $s_{\bar{x}}$ in scaling the cusum paper and the mask, instead of σ_e used earlier to set up for means. Annex A contains a method that helps to identify the suitability of this approach.

This procedure should have the desired effect of reducing the number of small and possibly “spurious” out-of-control signals on the cusum plot and make for more appropriate quality control.

9.5.2 “One-at-a-time” data

Some subjects for cusum monitoring will generate data that by their nature will occur one-at-a-time and any notion of subgrouping such data will make no sense. Examples, given earlier, would be monthly sales figures, or the temperature of a tank of a chemical used in a manufacturing process where to record several repeat temperatures taken at approximately the same time would not show any variation between the observations. In such a circumstance, the within-subgroup variation becomes zero and so no mask could be drawn.

Another example of this is the manner in which a golf score is determined. Some holes have a different expected number of strokes and a golfer measures performance against the expected number for each hole. The aggregate of the differences becomes the cusum.

Subgroup sizes of one also occur when taking and/or analysing samples is very expensive.

The approach taken should be to set the subgroup size of one and then follow the steps outlined in 9.3 and those following, taking the subgroup size to be one, i.e. n = 1. Thus, the location (mean level) is monitored by the individual results themselves while the level of the variation is monitored by the range shown between successive results.

The target for the mean value should be $T = \bar{x}$ determined from the trial period data, or the target should be a given value. The target for the range should be either $\bar{R}$ determined from the successive differences in the trial data, or set to T = 1,128σ if the standard deviation is a given value. Although the subgroup size is one, the effective subgroup size as far as the range is concerned is two.

9.5.3 Serial dependence between observations

The basis for a cusum chart, as for any control chart, is that of independence between the plotted points. There will be certain processes or sets of data where this might be untrue, e.g. processes where there is some closed-loop controller at work, such as a thermostat passing information to a heating device, or processes where seasonality is expected, such as in sales data.

The effect of this on a cusum plot could be serious and dramatically affect its performance, in some instances leading to false signals and in others to missing changes of an important amount.

Statistical tests exist that will indicate whether the data are serially dependent and whether the association is “positive” or “negative”. A straightforward method is to measure the correlation between the original data in their sequence of creation against the same data displaced by one, i.e. the first result in the original data is compared with the second, the second compared with the third, etc. If the calculated correlation coefficient is much greater than zero, it indicates a positive association between the observations, i.e. the results “move” generally in the same direction. If the correlation coefficient is much less than zero, then a negative association is indicated, suggesting that if a result is higher than its predecessor then the subsequent result will be in the opposite direction, a feature common to processes experiencing overcorrection. Such correlations are easily calculated using spreadsheet programs or other statistical computer packages and such an analysis is recommended.

Correlation coefficients range from −1 to +1 and the threshold for a value significantly different from zero will depend on how many data points are used in any study. If the data set is small, a seemingly large numerical value for the correlation coefficient can nevertheless be statistically non-significant, whereas with a large data set, a correlation coefficient quite close to zero might be interpreted as statistically significant. Table 20 provides approximate guidance on significant correlation coefficients.

Table 20 — Critical intervals for the correlation coefficient

Number of “paired” data points	Critical interval for correlation coefficient[a]
10	±0,45
15	±0,37
20	±0,33
25	±0,30

[a] At the 0,05 level of significance (two-sided).

If the calculated value of the correlation coefficient lies inside the interval given in Table 20, there is no reason to suppose that there is serial dependence. It is possible that there is some weak serial dependence that the sample size was too small to detect.

If serial dependence is discovered, specialist assistance might be required to determine the best way forward. Actions might involve a deeper process analysis as to the reason for the dependence. A solution to the problem can be simple or complex. If the dependence is caused by seasonality, it could be overcome by altering the target for every time period. The cusum value is made independent of the seasonality in this way.

9.5.4 Outliers

Cusums require protection against outliers. If an outlier occurs, its influence on the cusum value might be great and might lead to a spurious out-of-control signal. The following is a simple but effective method for protecting the cusum against outliers.

a) A result is deemed to be an outlier if:

   1) a subgroup mean is more than ±3σe from the target value; or

   2) an individual result is more than ±3σ from the target value.

Record the result as an outlier but do not add the value into the cusum calculation unless the subsequent result is outside the suspicion limits.

b) A result is a suspected outlier if:

   1) subgroup mean is more than $\pm 2\sigma_e$ from the target value; or

   2) an individual result is more than $\pm 2\sigma$ from the target value.

If two consecutive results are beyond the suspicion limits, include both of the results in the cusum calculation. This will almost always result in an out-of-control signal.

NOTE More rigorous methods (ISO 5727-5) to detect outliers exist but these are not very useable in “real time”. The method described here is a simple and practical way to proceed.

9.6 Cusum schemes for discrete data

9.6.1 Event count — Poisson data

9.6.1.1 General

Countable data relate to counts of events where each item of data is the count of the number of particular events per given time period or quantity of product. Instances are: number of accidents or absentees per month, number of operations or sorties per day, number of incoming telephone calls per minute, or number of non-conformities per unit or batch.

The Poisson distribution has two principal parts to play in cusum analysis:

a) as an approximation to the more cumbersome binomial (see 9.6.2) when

   n is large and p is small, say $n > 20$ and $p < 0{,}1$; and

b) as a distribution in its own right, when events occur randomly in time or space and the observation is made of the number of events in a given interval.

The validity of the Poisson model hinges on the independence of events and their occurrence at an average rate that is assumed to be stable (in the absence of special causes).

Due to the general lack of symmetry associated with the Poisson (and the binomial) model, different decision rules should be used for evaluating shifts in the upward and the downward directions. Therefore, if a truncated V-mask is used, the mask will not be symmetrical as before. There will be different values for the slope and decision interval for the upper and lower halves.

As a further apparent complication, but introduced for ease of calculation, certain distributions are sometimes approximated by others. For example, under certain conditions the Poisson or normal distribution serves as an approximation to the binomial, and in others the normal for the Poisson.

In 9.3 and 9.4, the ARLs for normally distributed data were determined simply from the ARLs of a standardized normal distribution, having a mean of 0 and a standard deviation of 1. Discrete distributions do not possess this feature. Each parameter should be individually calculated. Hence, tabulations for discrete cusum design purposes have been, of necessity, restricted to selected combinations for movement in the upward direction only. More recently, the ready availability of software routines has significantly increased the choice of cusum designs for discrete data.

9.6.1.2 General cusum decision rules for discrete data

A cusum scheme for discrete data is uniquely specified in terms of the type of distribution of the data and two parameters, K the datum value, and H the decision interval. Key mental markers in the choice of the parameters are the following.

a) The design of a decision scheme cusum is essentially a two stage process:

   1) selection of a K and H combination to give the desired in-control ARL; and

   2) determination of the swiftness of the signal response at various appropriate shifts in the mean.

b) The datum value K should be chosen on the basis of the specified shift in mean for which a response is to be signalled. A convenient K is at a value between the in-control mean and the out-of-control mean for which the cusum should have maximum sensitivity. The particular value of K is dependent on the type of distribution of the data and the definition of what is an acceptable value for the mean.

9.6.1.3 Cusum schemes for count data

Step 1 — Determine the actual mean rate of occurrence, m, and the standard deviation, σ_e.

Step 2 — Select a reference or target rate of occurrence, T_m. Frequently this will take the value m.

Step 3 — Decide on the most appropriate decision rule by selecting which scheme to apply. A preferred option is either a CS1 scheme with an ARL at target level of at least 1 000, or a CS2 scheme with an ARL at target level of at least 200. Refer to Table 21.

Step 4 — Determine H and K values thus:

a) for T_m ($0{,}1 \leqslant T_m \leqslant 25{,}0$) enter Table 21 at the nearest value to T_m. Use linear interpolation between values of T_m from 10,0 to 25,0; or

b) if $T_m > 25{,}0$ refer to the appropriate tables in 9.3 relating to the normal distribution, using $\sigma_e = \sqrt{T_m}$ as the normal approximation to the Poisson is now appropriate.

As an example, suppose T_m = 25. For a normal variable with a mean of 25, standard deviation of 5, H = 24 and K = 28; the corresponding ARL is about 1 500. The true ARL for a Poisson variable with H = 24 and K = 28 is 1 085. This shortening of the ARL arises from the skewness and discreteness of Poisson distributions.

***Step 5* —** Construct and apply a V-mask or tabulate:

a) for charting: plot the sum of differences ($X - T_m$) and use a truncated V-mask with decision interval, H, and slope, F (= $X - T_m$); or

b) for tabular cusum: form the sum of differences ($X - K$), reverting to zero whenever the accumulation becomes negative. Test against H for a signal of shift.

***Step 6* —** Assess the ARL performance of the chosen scheme at shifts of interest from the nominal using Table 22.

EXAMPLE

Step 2: Reference mean rate, T_m = 4.

Step 3: Use a CS1 scheme.

Step 4: Enter Table 21 at T_m = 4,0. Hence, H = 8 and K = 6.

Step 5 a): Plot the cusum and construct and apply V-mask (H = 8, F = 2).

Step 5 b): Tabulate and construct a tabular cusum (H = 8, K = 6).

Step 6: The performance of the scheme is shown below. If the process was operating at the target level, the ARL (L_0) is 1 736. However, the ARL will fall to 10 if the rate increases to 6,60.

H	K	T_m	L_0
8,0	6,00	4,000	1 736

ARL	1 000	500	200	100	50	20	10	5	2
m	4,160	4,380	4,710	5,000	5,300	5,90	6,60	7,80	11,50

NOTE Data extracted from Table 22.

9.6.2 Two classes data — binomial data

9.6.2.1 General

With classified data, each item of data is classified as belonging to a number of categories. Frequently, the number of categories is two, namely a binomial situation where, for instance, the outcome is usually expressed as 0 and 1, or as pass/fail, profit/loss, in/out, or presence/absence of a particular characteristic.

Data having two classes are termed “binomial” data. A measure can be inherently binomial, e.g. was a profit or loss made, is someone in or out? Sometimes it is arrived at indirectly by categorizing some other numerical measure. Take, for instance, the case where telephone calls are classified on whether they last more than 10 min or, perhaps, whether they are answered within 6 rings.

Table 21 — CS1 and CS2 schemes for count (Poisson) data in terms of T_m, H and K

Target event rate	CS1 scheme		CS2 scheme	
T_m	H	K	H	K
0,100	1,5	0,75	2,0	0,25
0,125	2,5	0,50	2,5	0,25
0,160	3,0	0,50	2,0	0,50
0,200	3,5	0,50	2,5	0,50
0,250	4,0	0,50	3,0	0,50
0,320	3,0	1,00	4,0	0,50
0,400	2,5	1,50	3,0	1,00
0,500	3,0	1,50	2,0	1,50
0,640	3,5[a] or 4,0	1,50	2,0	2,00
0,800	5,0	1,50	3,5	1,50
1,000	5,0	2,00	5,0	1,50
1,250	4,0	3,00	5,0	2,00
1,600	5,0	3,00	4,0	3,00
2,000	7,0[a] or 8,0	3,00	5,0	3,00
2,500	7,0	4,00	5,0	4,00
3,200	7,0	5,00	5,0	5,00
4,000	8,0	6,00	6,0	6,00
5,000	9,0	7,00	7,0	7,00
6,400	9,0	9,00	9,0	8,00
8,000	9,0	11,00	9,0	10,00
10,000	11,0	13,00	11,0	12,00
15,000	16,0	18,00	11,0	18,00
20,000	20,0	23,00	14,0	23,00
25,000	24,0	28,00	17,0	28,00

NOTE 1 CS1 schemes give ARLs at target generally between 1 000 and 2 000 observations. CS2 schemes give ARLs at target generally between 200 and 400 observations.

NOTE 2 For $T_m < 1$, a scaling is recommended as a reminder that the individual observations contain limited information in isolation. Determine the average number of observations required to yield one event, that is $1/T_m$. Round this value up to a convenient integer for plotting and adopt it as the horizontal interval for the cusum chart. Mark off the vertical cusum scale in intervals of the same length as the horizontal scale, and label these, as consecutive even integers above and below zero, i.e. 0, +2, +4, etc., –2, –4, etc.

NOTE 3 The choice of values of T_m up to 10 is based on the R 10 series of preferred numbers giving ten approximately equal ratios between successive entries within each decade.

NOTE 4 For T_m from 10 to 25, equal spaced values are given to facilitate interpolation. Intermediate schemes in this region can be obtained by linear interpolation in both H and K and rounding to integer values. It is preferable to round both H and K in the same sense.

[a] The lower value of H gives L_0 slightly below 1 000; the higher value gives L_0 nearly 2 000.

Table 22 — ARL characteristics for cusum schemes in terms of H and K for count (Poisson) data

Parameters		CS1 scheme		CS2 scheme		Mean rates of occurrence at stated values of ARL								
H	K	T_m	L_0	T_m	L_0	1 000	500	200	100	50	20	10	5	2
2,0	0,25	—	—	0,100	212	0,057	0,072	0,102	0,135	0,179	0,29	0,43	0,74	1,99
2,5	0,25	—	—	0,125	227	0,078	0,097	0,131	0,166	0,220	0,33	0,49	0,82	2,12
2,0	0,50	—	—	0,160	230	0,095	0,121	0,168	0,220	0,280	0,42	0,59	0,91	2,09
1,5	0,75	0,100	1 033	—	—	0,120	0,130	0,181	0,240	0,320	0,46	0,66	0,99	2,11
2,5	0,50	0,125	1 371	0,200	278	0,138	0,167	0,220	0,280	0,350	0,49	0,68	1,05	2,32
3,0	0,50	0,160	1 609	0,250	264	0,179	0,210	0,270	0,330	0,400	0,56	0,77	1,12	2,74
3,5	0,50	0,200	1 461	—	—	0,220	0,250	0,310	0,370	0,440	0,60	0,84	1,31	3,02
4,0	0,50	0,250	966	0,320	271	0,250	0,280	0,340	0,400	0,470	0,65	0,91	1,41	3,37
3,0	1,00	0,320	1 174	0,400	446	0,330	0,390	0,480	0,570	0,690	0,91	1,17	1,63	3,08
2,0	1,50	—	—	0,500	260	0,360	0,420	0,540	0,640	0,780	1,04	1,32	1,78	3,17
2,5	1,50	0,400	1 103	—	—	0,410	0,490	0,610	0,730	0,890	1,15	1,46	1,93	3,37
2,0	2,00	—	—	0,640	221	0,420	0,510	0,650	0,790	0,970	1,27	1,58	2,09	3,44
3,0	1,50	0,500	1 475	—	—	0,540	0,620	0,740	0,860	1,010	1,28	1,60	2,12	3,85
3,5	1,50	0,640	833	0,800	249	0,620	0,700	0,830	0,950	1,100	1,38	1,70	2,26	3,74
4,0	1,50	0,640	1 843	—	—	0,700	0,790	0,920	1,040	1,190	1,47	1,81	2,38	4,44
5,0	1,50	0,800	1 439	1,000	274	0,840	0,920	1,040	1,160	1,310	1,60	1,95	2,64	5,25
5,0	2,00	1,000	1 904	1,250	259	1,090	1,190	1,350	1,500	1,680	2,00	2,37	3,09	5,90
4,0	3,00	1,250	1 867	1,600	354	1,380	1,530	1,750	1,950	2,200	2,61	3,04	3,76	6,35
5,0	3,00	1,600	1 118	2,000	188	1,640	1,770	1,940	2,180	2,420	2,83	3,29	4,07	6,60
7,0	3,00	2,000	894	—	—	1,980	2,110	2,310	2,490	2,710	3,09	3,57	4,59	7,55
8,0	3,00	2,000	1 927	—	—	2,110	2,330	2,430	2,600	2,810	3,23	3,78	4,80	8,40
5,0	4,00	—	—	2,500	300	2,170	2,350	2,600	2,870	3,160	3,63	4,16	5,00	7,60
7,0	4,00	2,500	1 761	—	—	2,620	2,800	3,050	3,260	3,450	3,99	4,53	5,60	8,85
5,0	5,00	—	—	3,200	245	2,730	2,940	3,270	3,560	3,890	4,45	5,00	6,00	8,50
7,0	5,00	3,200	1 318	—	—	3,280	3,480	3,780	4,030	4,320	4,88	5,50	6,50	9,80
6,0	6,00	—	—	4,000	373	3,640	3,880	4,240	5,550	4,930	5,50	6,20	7,20	10,40
8,0	6,00	4,000	1 736	—	—	4,160	4,380	4,710	5,000	5,300	5,90	6,60	7,80	11,50
7,0	7,00	—	—	5,000	348	4,600	4,840	5,200	5,600	5,900	6,60	7,30	8,50	11,60
9,0	7,00	5,000	1 268	—	—	5,100	5,300	5,700	6,000	6,400	6,90	7,70	9,10	13,50
9,0	8,00	—	—	6,400	226	5,800	6,100	6,500	6,800	7,200	7,90	8,60	10,00	14,20
9,0	9,00	6,400	1 351	—	—	6,500	6,800	7,200	7,600	8,100	8,80	9,60	11,10	15,20
9,0	10,00	—	—	8,000	213	7,200	7,600	8,000	8,400	8,900	9,70	10,50	11,90	16,20
9,0	11,00	8,000	946	—	—	8,000	8,300	8,800	9,300	9,800	10,50	11,40	13,00	16,40
11,0	12,00	—	—	10,000	234	9,300	9,600	10,100	10,500	11,000	11,90	12,80	14,60	19,80
11,0	13,00	10,000	1 052	—	—	10,000	10,400	11,000	11,400	11,900	12,70	13,70	15,50	20,30
11,0	18,00	—	—	15,000	214	13,900	14,300	15,100	15,600	16,300	17,40	18,50	20,40	25,90
16,0	18,00	15,000	1 289	—	—	15,100	15,500	16,100	16,500	17,200	18,20	19,40	21,70	29,10
14,0	23,00	—	—	20,000	215	18,800	19,300	20,100	20,700	21,400	22,60	24,00	26,20	32,90
20,0	23,00	20,000	1 140	—	—	20,100	20,500	21,100	21,700	22,300	23,50	24,90	27,60	36,80
17,0	28,00	—	—	25,000	222	23,700	24,300	25,100	25,800	26,500	27,80	29,30	31,90	40,00
24,0	28,00	25,000	1 085	—	—	25,100	25,500	26,200	26,700	27,400	28,70	30,40	33,50	44,90

NOTE The table relates to upward movement of the mean.

The conditions that should be satisfied for a binomial distribution are the following.

a) There is a fixed number of trials, n.

b) Only two possible outcomes are possible at each trial.

c) The trials are independent.

d) There is a constant probability of “success”, p, in each trial.

e) The variable is the total number of “successes” in n trials.

The binomial distribution is very cumbersome to calculate, so, when handling ARLs and decision criteria, other distributions may be used as approximations to the binomial.

Because of the wide range of possible values of the two binomial distribution parameters, n (often corresponding to sample size) and p (the proportion of items having the attribute of interest), it is impracticable to provide comprehensive tables for all combinations. However, approximate procedures can be applicable in a wide range of situations. These procedures are given below.

— Situation 1: Where $T_p < 0{,}1$ (i.e. the target, or reference, proportion is below 10 %), use the appropriate scheme for a Poisson variable with $T_m = np$.

— Situation 2: Where $T_m > 20$ (i.e. the average number of “events” per sample under target conditions exceeds 20), use the appropriate scheme for a normal distribution.

9.6.2.2 Situation 1: $T_p < 0{,}1$ — Poisson-based scheme

Although, here, the binomial distribution is appropriate, it is very unwieldy to use and may be approximated by the Poisson distribution. The step-by-step method described in 9.6.1.3 is used but with $T_m = np$. Poisson will always give an ARL, at target, shorter than that for the binomial case, at target, but the ARLs at appreciable shifts from the target conditions will closely match those for the binomial for the same average rate of occurrence of events.

EXAMPLE

Step 2: Say n = 20 and p = 0,025, thus np = 0,5, and so reference value, T_m = 0,5.

Step 3: Use a CS1 scheme.

Step 4: Enter Table 21 at T_m = 0,5. Hence, H = 3 and K = 1,5.

Step 5 a): Plot cusum and construct and apply V-mask (H = 3, F = 1,0).

Step 5 b): Tabulate and construct a tabular cusum (H = 3, K = 1,5).

Step 6: The performance of the scheme is shown below. If the process was operating at the target level, the ARL (L_0) is 1 475. However, the ARL will fall to 10 if the rate increases to 1,60, i.e. p = 0,080.

H	K	T_m	L_0
3,0	1,50	0,500	1 475

ARL	1 000	500	200	100	50	20	10	5	2
m	0,540	0,620	0,740	0,860	1,010	1,28	1,60	2,12	3,85

NOTE Data extracted from Table 22.

9.6.2.3 Situation 2: $T_m > 20$ — Normal-based scheme

In situation 2, choose a suitable pair of h, f parameters, e.g. 5, 0,5, for the corresponding normal variable. The cusum parameters for the binomial are then obtained as:

$H = h \times \sqrt{nT_p(1-T_p)}$, rounding H to the nearest integer.

$K = nT_p + \left[f \times \sqrt{nT_p\left(1-T_p\right)} \right]$, rounding K to the nearest integer.

$F = f \times \sqrt{nT_p(1-T_p)}$, rounding F to the nearest integer.

EXAMPLE

Step 2: Say n = 80 and T_p = 0,3, h = 5 and f = 0,5.

Step 4: $H = h \times \sqrt{nT_p(1-T_p)} = 5 \times \sqrt{80 \times 0,3(1-0,3)} \approx 20$

$K = nT_p + \left[f \times \sqrt{nT_p\left(1-T_p\right)} \right] = (80 \times 0,3) + \left[0,5 \times \sqrt{80 \times 0,3\left(1-0,3\right)} \right] \approx 26$

$F = f \times \sqrt{nT_p(1-T_p)} = 0,5 \times \sqrt{80 \times 0,3(1-0,3)} \approx 2$

Step 5 a): Plot cusum and construct and apply V-mask (H = 20, F = 2).

Step 5 b): Tabulate and construct a tabular cusum (H = 20, K = 26).

Step 6: If the process was operating at the target level the ARL (L_0) is approximately 930. However, the ARL will fall to approximately 10 if the proportion (p) increases to 0,35.

Annex A
(informative)

Von Neumann method

When setting up a cusum chart for data presentation purposes, the choice of procedure for measuring the variability may be made largely on the 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 performed, 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.

1. Compute $s_{\overline{x}}$.
2. Compute $\sum_{i=2}^{k} w_i^2$.
3. Count the number of subgroups, k.
4. Compute $\dfrac{\sum_{i=2}^{k} w_i^2}{2(k-1)s_{\overline{x}}^2}$.
5. Compute $1 \pm \dfrac{2}{\sqrt{(k+2)}}$.

If the value calculated in step 4 lies above the upper bound in step 5, the implication is of negative serial correlation, e.g. overcontrol or alternation. A value below the lower bound may occur due to cycling or other forms of positive serial correlation, e.g. lag effects, or changes of mean level within the sequence of observations, whether regular or irregular step changes, drifting or trends.

Alternatively, the calculated value in step 4 may be assessed from Table A.1 of 0,05 (two-tail) probability points.

Table A.1 — Probability points (two-tail) for the von Neumann test

Number of subgroups	Lower critical value	Upper critical value
20	0,58	1,42
30	0,65	1,35
50	0,73	1,27
75	0,78	1,22
150	0,84	1,16
200	0,86	1,14

Annex B
(informative)

Example of tabular cusum

A historical mean of 35 and a standard deviation of 6 have been established and have been used as the parameters for a Shewhart chart and also a cusum chart. The target value has been set to be 35.

Twenty-four days of data have been collected: 25,8, 33,4, 31,6, 26,0, 36,4, 33,0, 35,8, 41,8, 44,2, 37,2, 35,0, 41,8, 33,4, 38,4, 30,2, 33,8, 42,6, 39,6, 32,0, 48,4, 44,6, 43,0, 40,8 and 50,6.

A Shewhart chart of the data is shown in Figure B.1. No signals are identified using the standard Shewhart chart tests.

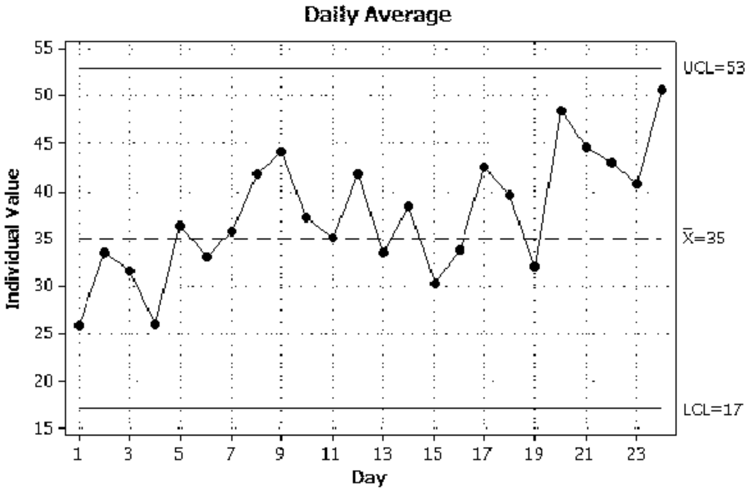


Figure B.1 — Shewhart chart of daily average value

A cusum plot of the data is shown in Figure B.2. It shows a "signal" at day 24. The Shewhart chart in Figure B.1 does not detect this change.

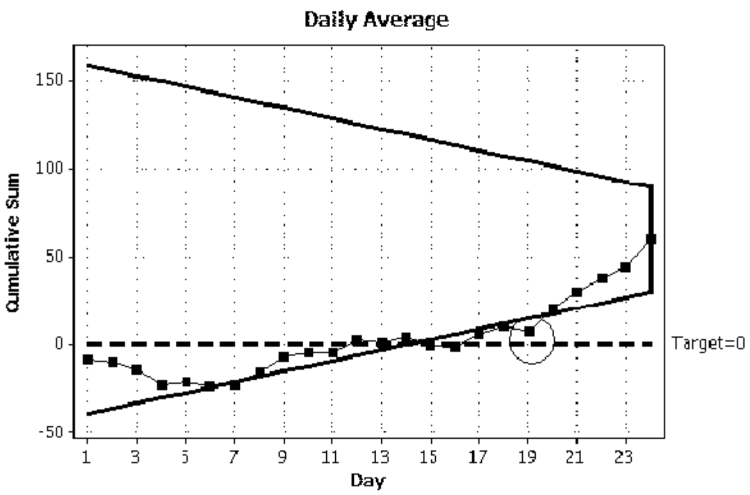


Figure B.2 — Cusum plot and V-mask of daily average

A tabular version of this cusum is shown below in Table B.1.

Table B.1 — Tabular cusum of daily average data

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Day	**Daily average** (X)	$(X - T - F)$	**Sum "Hi"**	**Number "Hi"**	$(X - T + F)$	**Sum "Lo"**	**Number "Lo"**
0			15,0	0		–15,0	0
1	25,8	–12,2	2,8	1	–6,2	–21,2	1
2	33,4	–4,6	0,0	0	1,4	–19,8	2
3	31,6	–6,4	0,0	0	–0,4	–20,2	3
4	26,0	–12,0	0,0	0	–6,0	–26,2	4
5	36,4	–1,6	0,0	0	4,4	–21,8	5
6	33,0	–5,0	0,0	0	1,0	–20,8	6
7	35,8	–2,2	0,0	0	3,8	–17,0	7
8	41,8	3,8	3,8	1	9,8	–7,2	8
9	44,2	6,2	10,0	2	12,2	0,0	0
10	37,2	–0,8	9,2	3	5,2	0,0	0
11	35,0	–3,0	6,2	4	3,0	0,0	0
12	41,8	3,8	10,0	5	9,8	0,0	0
13	33,4	–4,6	5,4	6	1,4	0,0	0
14	38,4	0,4	5,8	7	6,4	0,0	0
15	30,2	–7,8	0,0	0	–1,8	–1,8	1
16	33,8	–4,2	0,0	0	1,8	0,0	0
17	42,6	4,6	4,6	1	10,6	0,0	0
18	39,6	1,6	6,2	2	7,6	0,0	0
19	32,0	–6,0	0,2	3	0,0	0,0	0
20	48,4	10,4	10,6	4	16,4	0,0	0
21	44,6	6,6	17,2	5	12,6	0,0	0
22	43,0	5,0	22,2	6	11,0	0,0	0
23	40,8	2,8	25,0	7	8,8	0,0	0
24	50,6	12,6	37,6	8	18,6	0,0	0

NOTE $T = 35$; $f = 0,5s$; $h = 5s$.

This cusum uses a "fast initial response" (FIR), so a starting value is placed at "Day 0" in the "Sum Hi" and "Sum Lo" columns. This value is equal to +2,5s and –2,5s respectively, or +15,0 and –15,0. Likewise, 0 has been placed in the "Number Hi" and "Number Lo" columns at Day 0. Since $h = 5s$, the limits for a cusum signal are +30 and –30 respectively.

The value for "Day 1" is 25,8. The calculations for column (3) are as follows:

$$25,8 - 35 - 3 = -12,2$$

This value, −12,2, is added to the previous “Sum Hi” value of 15,0, resulting in 2,8. Since “2,8” is positive, a “1” is added in the “Number Hi” column to the previous value, i.e. 0. The new “Number Hi” count is 1.

The calculations for column (6) are:

$$25,8 - 35 + 3 = -6,2$$

This value of −6,2 is added to the previous “Sum Lo” value of −15,0 resulting in −21,2. Since “−21.2” is negative, a “1” is added in the “Number Lo” column to the previous value, i.e. 0. The new “Number Lo” is 1.

These calculations may seem tedious, but it is imagined that a computer will be used to perform them.

The value for “Day 2” is 33,4. The calculations for column (3) are as follows:

$$33,4 - 35 - 3 = -4,6$$

This value, −4,6, is added to the previous “Sum Hi” value of 2,8 resulting in −1,8. Since this is negative, the “Sum Hi” value is changed to 0 and the “Number Hi” count is also changed to 0, i.e. for this side of the cusum, only the “positive” cusum values are “counted”.

The calculations for column (6) are:

$$33,4 - 35 + 3 = 1,4$$

This value of 1,4 is added to the previous “Sum Lo” value of −21,2 resulting in −19,8. As before, since “−19,8” is negative, a “1” is added in the “Number Lo” column to its previous value, i.e. the new “Number Lo” is 2.

This procedure continues until either a “Sum Hi” or a “Sum Lo” exceeds the h value (30 or −30, in this example). This occurs on “Day 24” when the “Sum Hi” value is 37,6.

An estimate of when the “shift” occurred and the amount of the “shift” can be obtained from the tabular table. Note from Table B.1 that the effect of the FIR value in both “Sum Hi” and “Sum Lo” decreases relatively quickly if the process is in the vicinity of the target. If it is not, then usually a “Sum Hi” or “Sum Lo” signal will be obtained.

In this example, when the “Sum Hi” signal is obtained, the “Number Hi” is 8. This suggests that the change occurred between “Day 16” and “Day 17”. The estimated shift, assuming that FIR has “died out”, is:

$$F + \frac{\text{Sum Hi}}{\text{Number Hi}}$$

If the indicated “shift” was negative, the estimate is given by:

$$-F - \frac{\text{Sum Lo}}{\text{Number Lo}}$$

Since the “Sum Hi” exceeded the h value of 30, the estimated shift is:

$$F + \frac{\text{Sum Hi}}{\text{Number Hi}} = 3 + \frac{37,6}{8} = 7,70$$

NOTE The last eight values of the original data were shifted upwards by approximately one standard deviation, i.e. 6 units. The cusum has identified the correct place of the shift and an estimated shift of 7,7 that is not significantly different to 6.

Figure B.3 shows a graph of the tabular cusum.

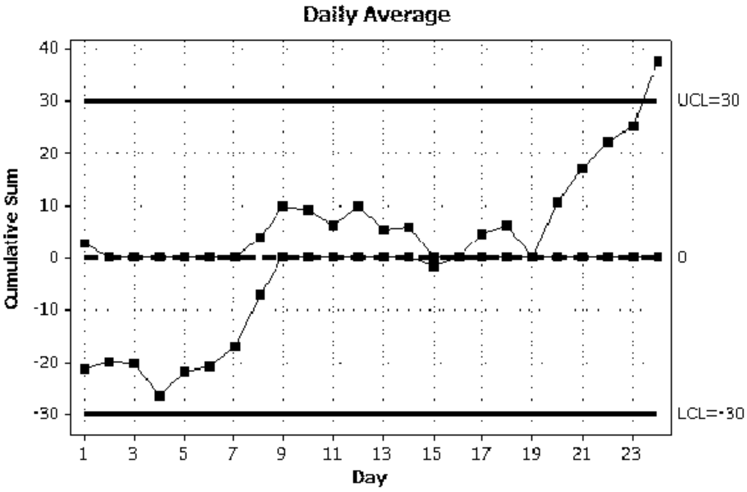


Figure B.3 — Plot of tabular cusum for daily average

Annex C
(informative)

Estimation of the change point when a step change occurs

Whether a process is in control or not can be judged from control charts by using specific criteria, but control charts cannot identify the assignable cause. However, the behaviour of the past plots on the control chart may introduce information for finding assignable causes. This is an important role that the estimation of the change point fills.

In the case of step change in the process mean $\mu(t)$, the change point is τ as shown in the equation:

$$\mu(t) = \begin{cases} T & (t = 1,\ldots,\tau) \\ T + \delta & (t = \tau + 1,\ldots) \end{cases}$$

A cusum chart has a good performance for the estimator of a change point. The estimator of the change point is the time point when the path of the cusum wanders outside the arms of the V-mask as shown in Figure C.1.

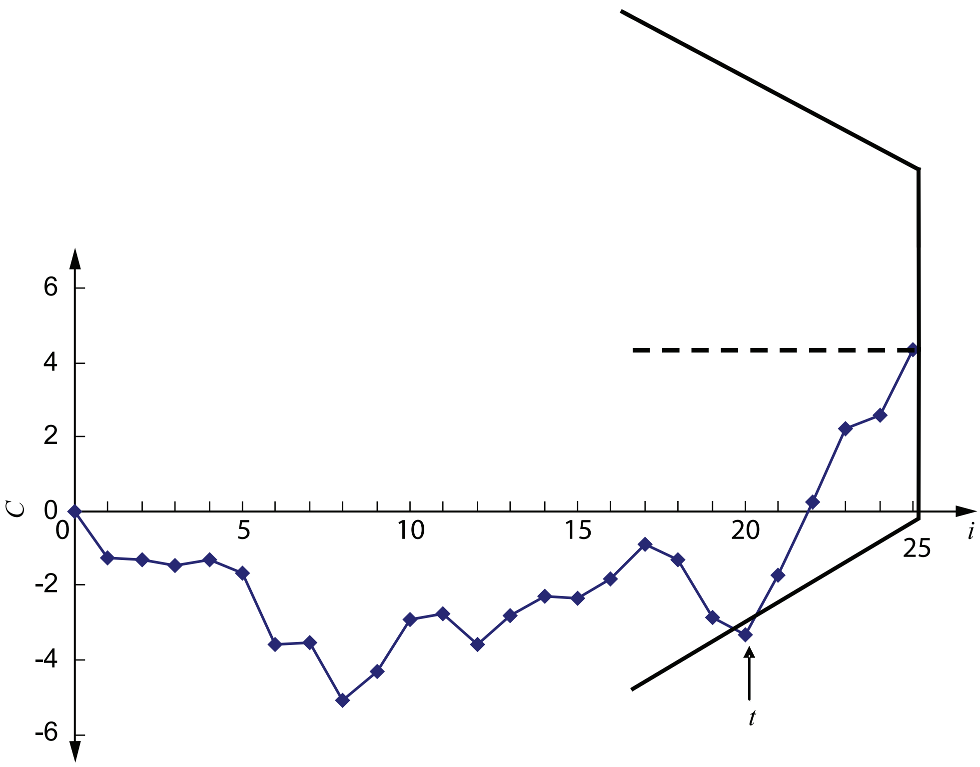
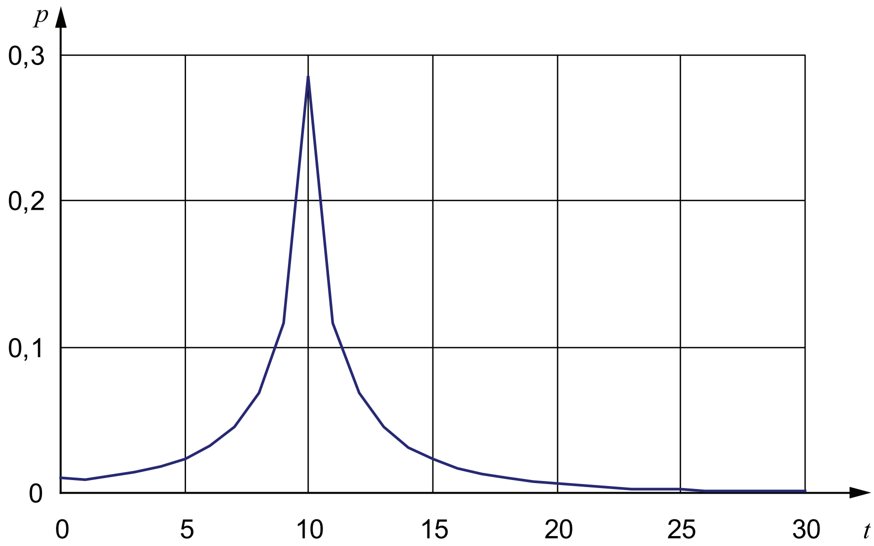


Figure C.1 — Estimate of the change point using the truncated V-mask

Figure C.2 shows a distribution of the estimator from a cusum chart, in the case of a CS1 scheme (h = 0,5, f = 0,5). The type 1 error should be taken into consideration because the change point is estimated only after the chart indicates a signal. Therefore, the distribution in Figure C.2 is derived on the condition that the type 1 error does not occur.

The distribution is unimodal and its mode is identical to the true change point, τ = 10. Although the size of shift is not so large in the case of Figure C.2, that is $\delta = 1{,}0\sigma_e = \Delta$, this approach is preferred.



NOTE For a truncated V-mask with parameters h = 5,0, f = 0,5, Δ = 1,0 and τ = 10.

Figure C.2 — Distribution of the estimator of the change point by using cusum

Bibliography

[1] NISHINA, K. Estimation of the change-point from cumulative sum tests. *Reports of Statistical Application Research*, *JUSE*, **33**(4), 1986

[2] NISHINA, K. Estimation of the amount of shift using cumulative sum tests. *Reports of Statistical Application Research*, *JUSE*, **35**(3), 1988

[3] NISHINA, K. A comparison of control charts from the viewpoint of change-point estimation. *Quality and Reliability Engineering International*, **8**, 1992

[4] ISO 5725-5, *Accuracy (trueness and precision) of measurement methods and results — Part 5: Alternative methods for the determination of the precision of a standard measurement method*

[5] ISO 7870-1, *Control charts — Part 1: General guidelines*

[6] ISO 7870-2[1)], *Control charts — Part 2: Shewhart control charts*

[7] ISO 9000, *Quality management systems — Fundamentals and vocabulary*

[8] BS 5703, *Guide to data analysis, quality control and improvement using cusum techniques*

1) Under preparation.

ICS 03.120.30

Price based on 63 pages