INTERNATIONAL STANDARD

ISO 7870-4

Second edition
2021-09

Control charts —

Part 4: Cumulative sum charts

Cartes de contrôle —

Partie 4: Cartes de contrôle à somme cumulée

Reference number
ISO 7870-4:2021(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

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

This second edition of ISO 7870-4 cancels and replaces the first edition (ISO 7870-4: 2011), which has been technical revised.

The main changes compared to the previous edition are as follows:

— Manhattan diagram removed (former 6.7);

— V-mask types in Types of CUSUM decision schemes reduced to one V-mask;

— von Neumann method removed (former Annex A).

A list of all parts in the ISO 7870 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

This document 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 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.

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.

The intention of this document 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 detects 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 require 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 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 7870-6.

Control charts —

Part 4:
Cumulative sum charts

1 Scope

This document describes 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 documents are referred to in the text in such a way that some or all of their content constitutes requirements 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.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

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 to entry: With a charted CUSUM, the deviations from the target value are cumulated.

Note 2 to entry: Using a 'V' mask, the target value is often referred to as the reference value or the nominal control value. If so, it needs be acknowledged that it is not necessarily the most desirable or preferred value, as can appear in other standards. It is simply a convenient target value for constructing a CUSUM chart.

3.1.2
representative out of control value
⟨tabulated CUSUM⟩ value which controls the sensitivity of the procedure

Note 1 to entry: The upper out of control value is $T + f\sigma_e$, for monitoring an upward shift. The lower control 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 the *representative out of control value* (3.1.2)

Note 1 to entry: It is necessary to distinguish between *f*, that relates to a standardized reference shift, and *F*, that relates to an observed reference shift; $F = f\sigma_e$. It plays a crucial role for constructing the tabular form of the CUSUM chart.

3.1.4
decision interval
H, *h*
⟨tabulated CUSUM⟩ cumulative sum of deviations from a *representative out of control value* (3.1.2) required to yield a signal

Note 1 to entry: It is necessary to distinguish between *h*, that relates to a standardized decision interval, and *H*, that relates to an observed decision interval; $H = h\sigma_e$.

3.1.5
average run length
ARL
average number of samples taken up to the point at which a signal occurs

Note 1 to entry: The average run length (ARL) is usually related to a process level, in which case it carries an appropriate subscript as, for example, ARL_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
FIR	fast initial response
LCL	lower control limit
RL	run length
SPC	statistical process control
UCL	upper control limit

3.3 Symbols

a	scale coefficient
C	CUSUM value
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

K reference value, equal to sum of target T and observed reference shift F

H observed decision interval

h standardized decision interval

ARL_δ average run length at δ shift

ARL_0 average run length at no shift μ population mean value

n subgroup size

m number of subgroups within a preliminary study

p probability of ‘success’

$\overline{R}$ mean subgroup range

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 m subgroups

T target value

τ true change point

$\hat{\tau}$ estimated change point

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 at 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; these are 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 can 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 A 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 generally provides 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 — 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 gives 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 can 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 Process

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

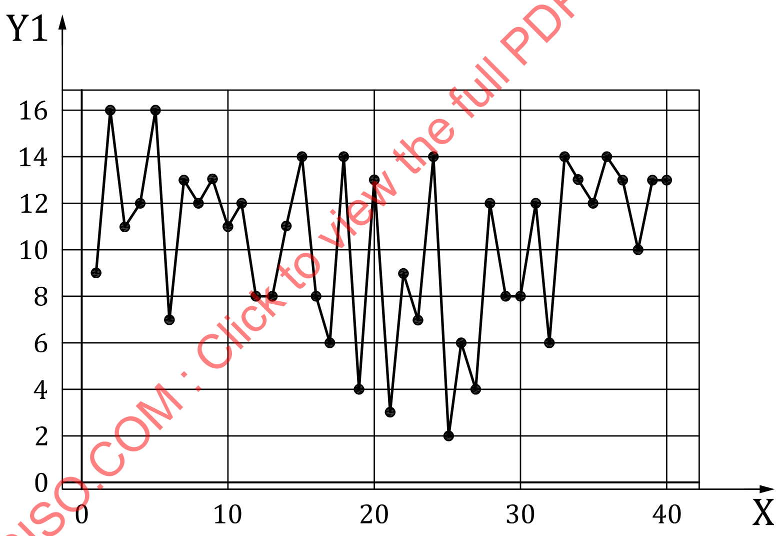
9, 16, 11, 12, 16, 7, 13, 12, 13, 11, 12, 8, 8, 11, 14, 8, 6, 14, 4, 13, 3, 9, 7, 14, 2, 6, 4, 12, 8, 8, 12, 6, 14, 13, 12, 14, 13, 10, 13, 13.

The reference or target voltage value is 10 V.

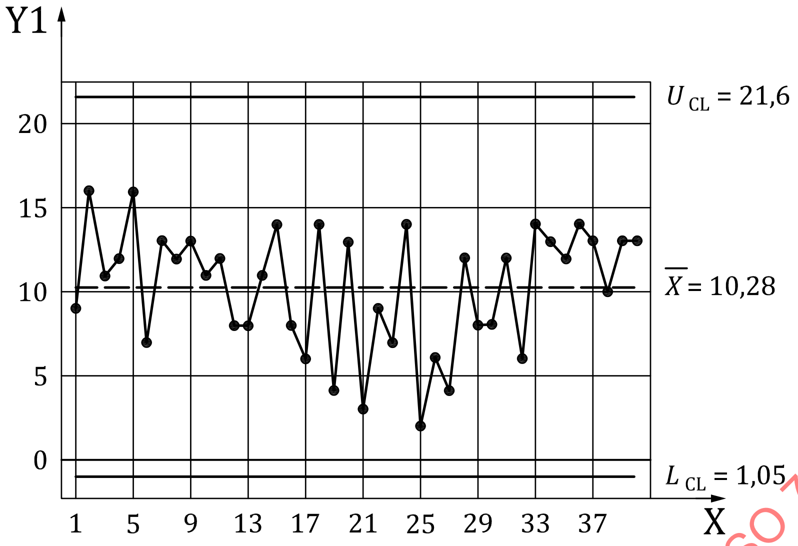
6.2 Simple plot of results

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

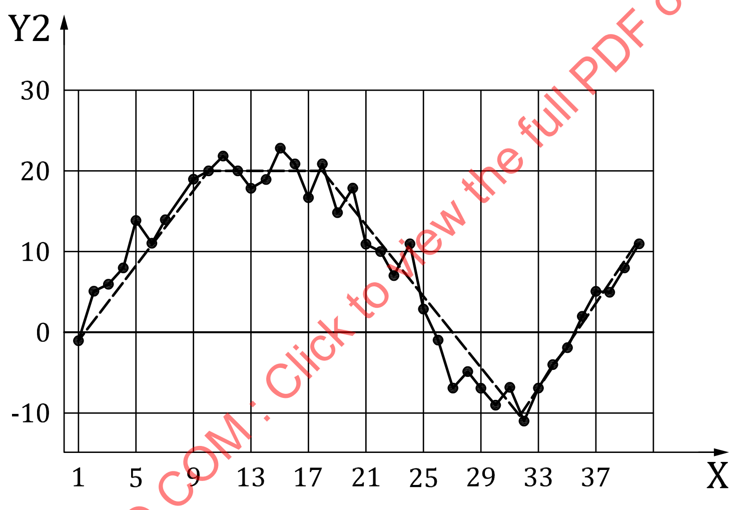
Apart from indicating a general drop away in the middle portion from a high start and with an equally high finish, Figure 1 a) is not very revealing because of the extremely noisy, or spiky, data throughout.



a) Simple plot of motor voltages



b) Standard control chart for individuals



c) CUSUM chart

Key

X motor number

Y1 voltage

Y2 cumulative sum

Figure 1 — Motor voltage example

6.3 Standard control chart for individual results

The next level of sophistication is 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 criterion to test for process stability and control is just “no points lying above the upper control limit (UCL) or below the lower control limit (LCL)”. All points reside within these limits. Hence, one can 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 the 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 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, 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 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 target value, T. 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	-4
15	14	+4	+23	35	12	+2	−2
16	8	−2	+21	36	14	+4	+2
17	6	−4	+17	37	13	+3	+5
18	14	+4	+21	38	10	0	+5
19	4	−6	+15	39	13	+3	+8
20	13	+3	+18	40	13	+3	+11

7 Fundamentals of making CUSUM-based decisions

7.1 Need for decision rules

Decision rules are 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 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 measurement taken every quarter of an hour to make up a subgroup of four every hour, gives 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.

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 can be too much local variation to provide a realistic basis for assessing whether 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 tends 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.

7.3 Measuring the effectiveness of a decision rule

7.3.1 Basic concepts

The ideal performance of a decision rule is 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 (usually denoted as run length, RL) 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, in effect, increases variability. Not taking appropriate action when the process has changed gives rise to 'under control'. This also, in effect, increases variability and results in increasing the cost of production.

The actual RL 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 occasionally. Generally, however, the ARL is looked upon as a reasonable measure of effectiveness of a decision rule. The aim is summarized in Table 2.

Table 2 — ARL patterns and process conditions

True process condition	Required CUSUM response	Ideal response
At or near target	Long ARL (few false alarms)	ARL → infinity
Substantial departure from target	Short ARL (rapid detection)	ARL is 1

7.3.2 Example of 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. It is well-known that about 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, one expects, on average, to see a value beyond the upper control limit only once in every 741 observation intervals.

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 resulting in 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 roughly 2,28 % 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 takes 44 observation intervals to signal a shift in the mean of one standard deviation.

When two-sided limits are considered, 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.

In practice, other signalling rules such as the addition of warning limits, runs above and below the mean, and so on, 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 to achieve a more rapid detection of shift while retaining long ARLs when on target. The CUSUM method is well suited to this.

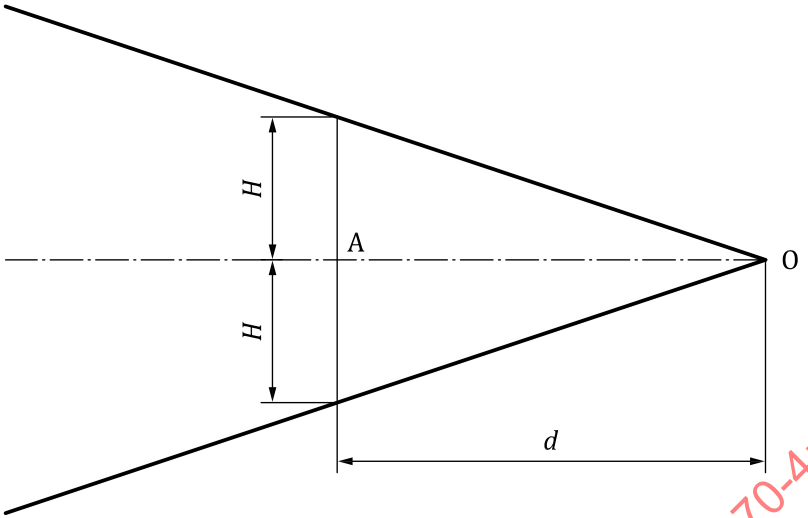
8 Types of CUSUM decision schemes

8.1 V-mask

Plotting the cumulated deviations from target *T* and applying a V-mask constitutes one way of executing the CUSUM control chart.

8.1.1 Configuration and dimensions

A V-mask is illustrated in Figure 2. The apex O is placed in distance *d* from the last CUSUM value which is placed consequently at A. The V-mask is then fully specified by either giving the distance *H* (= 5 σ_e, for example) to the limbs or by setting the angle between the horizontal line and the upper limb.



Key

A recent CUSUM value

H decision interval

d distance between apex and recent CUSUM value

O apex of V-mask

Figure 2 — Configuration and dimensions of a V-mask

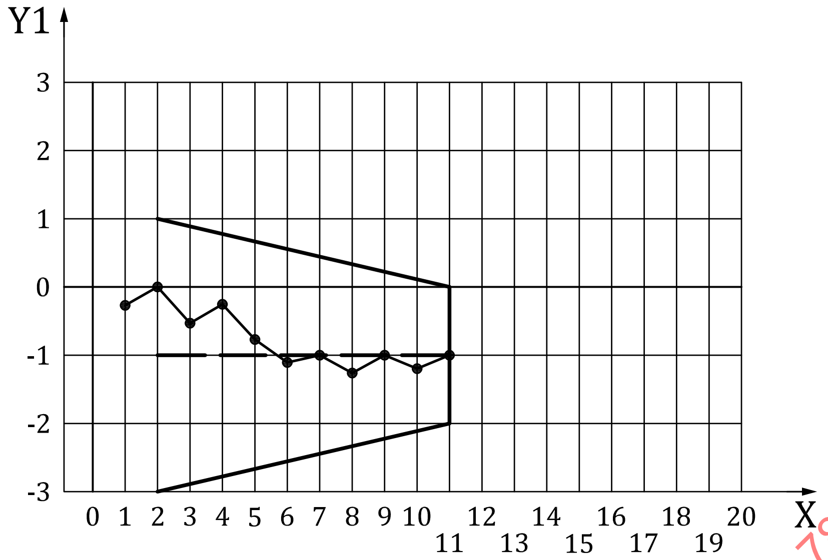
8.1.2 Application of the V-mask

The mask is used by placing point *A* (see Figure 2) 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.

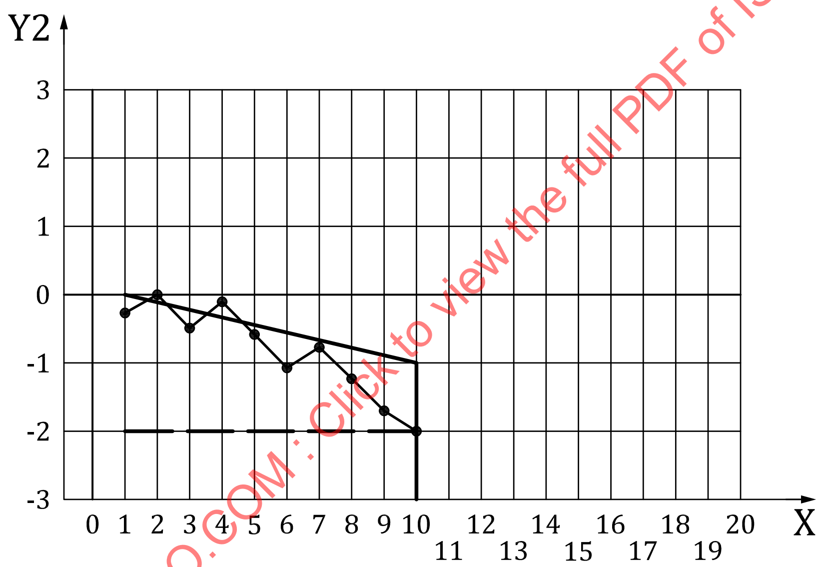
If the path of the CUSUM lies within the sloping arms of the mask (or their extensions), no significant shift in mean is indicated up to that plotted value. In a control situation, the process is then said to be in a state of statistical control with respect to the target value. If, however, the path 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 3 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 3. 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**

Key

X observation value

Y1 CUSUM 1

Y2 CUSUM 2

Y3 CUSUM 3

Figure 3 — Illustrations of use of the V-mask to detect significant change in process mean

Although Figure 3 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 3 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 provides both 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 3 c) indicates that the process mean is significantly greater than the target value. This was not registered as significant until observation 14. 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 can be ignored from a specification viewpoint, or used for directing attention to significant movement in a possibly more desirable direction.

8.1.3 Average run lengths

The average run length (ARL) properties for the V-mask to the dimensions given in Figure 4 are listed in Table 3 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.

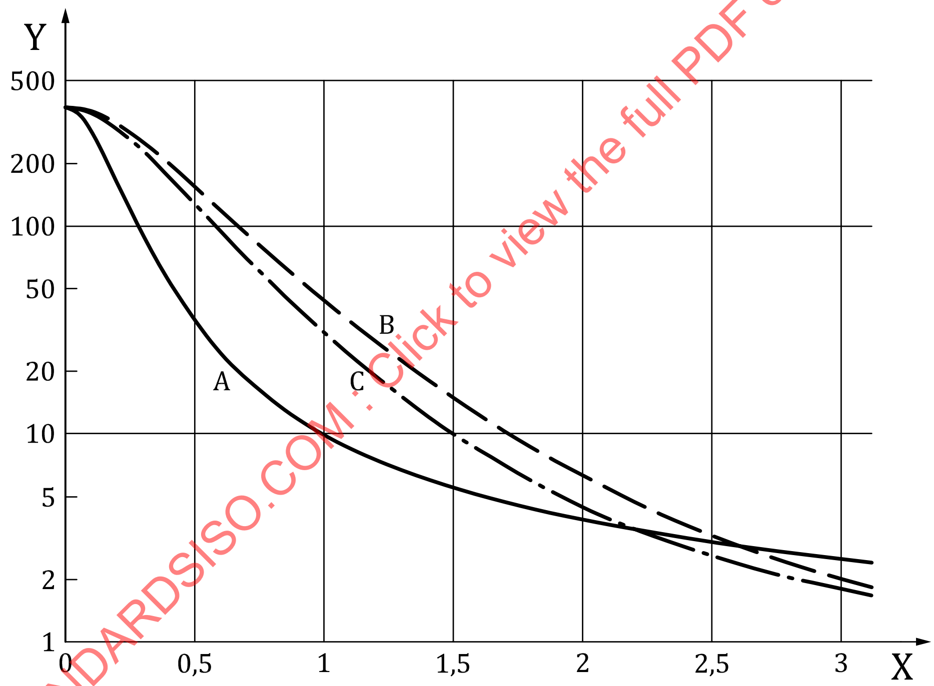
These rules are:

— rule 1: one point outside of, ± 3 standard deviations from the centreline, namely, the action limits or control limits — this is the common Shewhart alarm rule;

— rule 2: two consecutive points outside of, ± 2 standard deviations from the centreline, namely, the warning limits — very simple runs rule which already improves the detection performance for small shifts.

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

NOTE 2 The standard CUSUM referred to has h (height of decision interval) = 4,77 and f (slope of decision line) = 0,5. The Shewhart chart relates to rule 1 only. The Shewhart chart with 2 of 2 runs rule applies to the combination of Shewhart rules 1 and 2. For the latter combination, the values 2 (warning) and 3 (action) limits are increased by a factor of 1,03 to achieve as well 370,4.



Key

X shift (in multiples of standard deviation)

Y ARL to detect a shift

A CUSUM (solid line)

B shewhart chart (dashed line)

C shewhart chart with 2 of 2 runs rule (dash-dotted line)

Figure 4 — Average run lengths, in terms of shift from target value, for the CUSUM chart compared with standard Shewhart control charts

Table 3 — Average run lengths in terms of shift from target value for the CUSUM chart (f=0,5, h=4,77) 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		
	CUSUM	Shewhart control chart	Shewhart control chart with 2 of 2 runs rule
0,0	370,4	370,4	370,4
0,2	163,6	308,4	292,7
0,4	54,5	200,1	172,1
0,6	24,6	119,7	94,3
0,8	14,4	71,6	30,7
1,0	9,9	43,9	30,7
1,2	7,5	27,8	18,9
1,4	6,1	18,2	12,2
1,6	5,1	12,4	8,3
1,8	4,4	8,7	6,0
2,0	3,9	6,3	4,5
2,2	3,5	4,7	3,5
2,4	3,1	3,6	2,8
2,6	2,9	2,9	2,4
2,8	2,7	2,4	2,0
3,0	2,5	2,0	1,8

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 4 and Table 3 show the following.

a) For shifts up to 2σ_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 0,2σ_e to 1,6σ_e.

b) For shifts greater than 2σ_e, the Shewhart chart with runs rules responds more quickly than the CUSUM chart. For shifts greater than 2,6σ_e, the Shewhart chart alone responds more quickly than the CUSUM chart.

8.1.4 General comments on average run lengths

Firstly, the dimensions of the V-mask are designed to be especially suited to detecting shifts in the region of one standard error (1σ_e). If the focus is on other shifts, then different values of h and f are used. Secondly, many 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 3 and Figure 4 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, results in shortening the ARL at target, but has 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 is 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 4 is indicative of the distortion in ARLs for a 10 % error in estimation of σ_e.

Table 4 — 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	946,3	370,4	172,3
0,5	51,6	35,3	25,8
1,0	11,8	9,9	8,5
1,5	6,3	5,5	4,9
2,0	4,3	3,9	3,5

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

The effects of the three assumptions discussed are not peculiar to the CUSUM chart but are also applicable to other charting methods.

8.2 Fast-initial response (FIR) CUSUM

The fast-initial response (FIR) CUSUM is intended to reduce the ARL for early 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.

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, $h/2$. Note that the FIR concept is easier to understand within the tabular CUSUM framework described in 8.3.

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 hastens the signal of mean shift. On the other hand, if the process has not moved, the CUSUM naturally drops back towards zero and behaves just like the usual zero start CUSUM.

When used in conjunction with decision tabular schemes (see 8.3), the head start is often used with both the upper and lower CUSUMs.

8.3 Tabular CUSUM

8.3.1 Rationale

Often 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 ± $f\sigma_e$'.

With the tabular scheme, instead of cumulating and plotting:

observation value − target value,

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 '± $h\sigma_e$',

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.

8.3.2 Deployment

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 5 and another example of the tabular method in Annex A.

Table 5 — 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 Nature of the changes to be detected

9.1.1 Size of the changes 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 influences the shape of any ‘V-mask’ that can 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 identifies 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 is 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 can 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 emerges. 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 occurs 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 is detected on the CUSUM plot. Clearly, this is 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 creates ‘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 monthly, the target value to be used each month is likely to be different according to the time of year. It can be anticipated that more ice-cream is sold during the summer months than the winter months and so a different target can 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.

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 is 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 can 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 can 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 monitors 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 5.

Step 1 — Determine the subject for CUSUM charting

Determine the process parameter or product characteristic to be monitored.

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

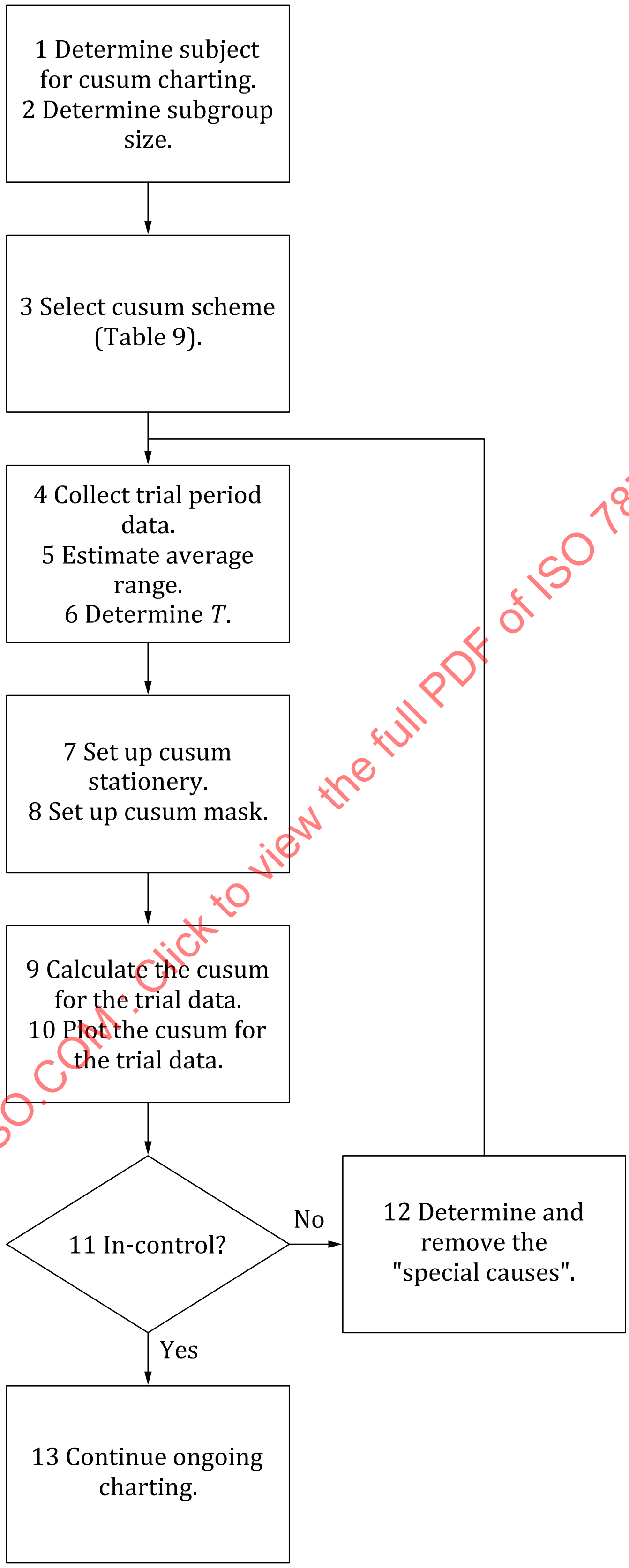


Figure 5 — CUSUM set-up protocol

Step 2 — Determine the subgroup size

The determination of a rational subgroup for the CUSUM chart is an identical thought process that can 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 leads 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 month, then the rational subgroup size is 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. The subgroup size is so selected to represent the random variation in the process.

Step 3 — Select CUSUM scheme

Table 6 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 detects 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 should determine which scenario is the more important and then select the appropriate scheme. Table 7 illustrates the differences in performance of these standard schemes (the ARL statements refer to upper CUSUM charts).

Table 6 — Standard CUSUM schemes for subgroup means

Important shift in the mean to be detected[a]	CS1 schemes		CS2 schemes	
	h	f	h	f
i) < $0{,}75\sigma_e$	8,0	0,25	5,0	0,25
ii) 0,75 to $1{,}50\sigma_e$	5,0	0,50	3,5	0,50
iii) > $1{,}50\sigma_e$	2,5	1,00	1,8	1,00
NOTE 1 CS1 schemes give average run lengths, ARL_0, in the region of 700 to 1 000 when the actual shift is zero.				
NOTE 2 CS2 schemes give average run lengths, ARL_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 Table 7 (ARL values are given again for a one-sided CUSUM). 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 7 — Comparison of performance of standard CUSUM schemes for subgroup means

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	737,0	931,0	716,0	142,0	200,0	172,0
0,75	16,4	17,0	27,3	10,4	11,5	15,3
1,00	11,4	10,4	13,4	7,4	7,4	8,8
1,50	7,1	5,7	5,4	4,7	4,2	4,1

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

Step 4 — Collect trial period data

As indicated in 9.2.3, data should be captured that describes 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 are 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 distorts 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 is satisfactory and the trial period closed. The data from the trial should then be used to establish the levels of variability the CUSUM scheme operates 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 can occur where a different approach can be required. This different approach requires 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 ($\bar{R}$) of all 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 8.

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 8 — 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 can 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, $\bar{s}$, should be calculated instead of $\bar{R}$ and σ_0 is estimated by $\bar{s}/c_4$. Table 15 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 m subgroups, there are $m - 1$ ranges. Calculate the average of these ranges ($\bar{R}$).

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 is a given value or a performance-based value determined from data.

a) Given value

The value of the target is a specified value. It comes from a specification document or a drawing and is 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

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

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

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

iii. Assign $\bar{\bar{x}}$ 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 9.

Table 9 — Table for CUSUM calculations

Subgroup number	$\bar{x}$	$\bar{x}$ −T	**CUSUM value,** C
etc.			

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

c) CUSUM graph paper

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

Select a suitable scale. The scale is 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 should be 10 mm, whereas for a plot intended for desk use only the interval is 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 6.

Step 8 — Set up the CUSUM mask

Subclause 8.1.1 describes the geometry of the standard CUSUM mask and Figure 2 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 $F = f\sigma_e$

NOTE Computer programs exist which 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 9 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 like that shown in Figure 6, 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 3 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 3 b).

Key

Y CUSUM value

Figure 6 — 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) 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 can 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 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’ can 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.

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, do not need any further alteration. For a few applications though, it can be noticed after some time that the basic scheme selected can be improved either because the ARL to detect a shift of an important size is too large, or because the frequency of ‘false alarms’ is too high.

9.3.3 ‘Tailored’ CUSUM schemes

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 can 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, ARL_δ.

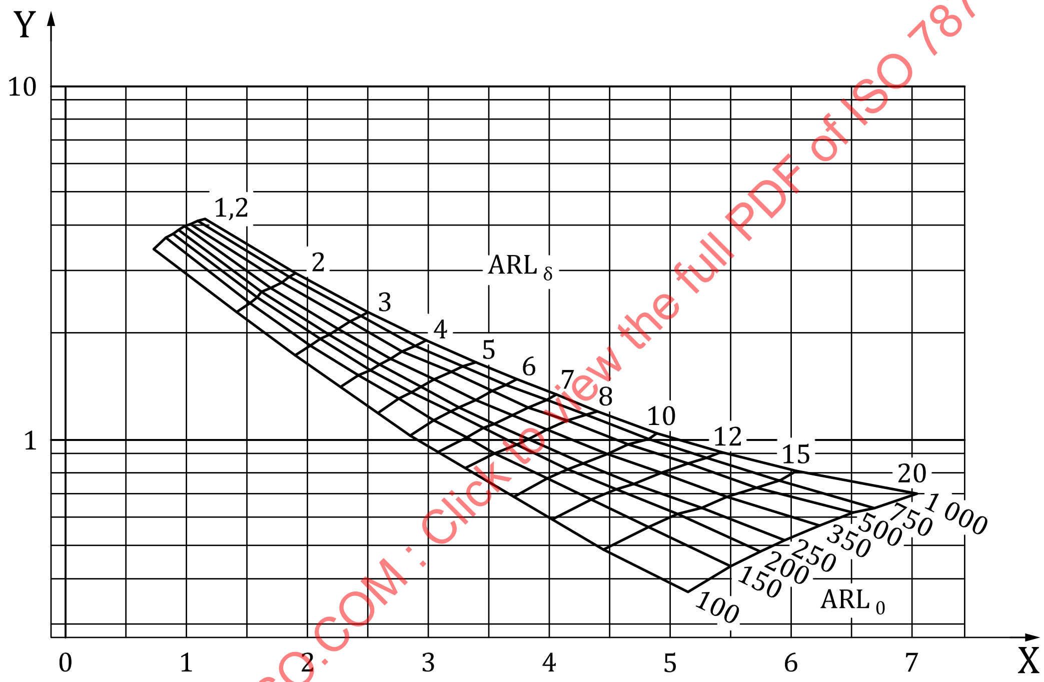
d) Specify the intended ARL when the shift is zero, i.e. the frequency of 'false alarms', ARL_0.

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

f) Enter the graph in Figure 7 and read off h for the calculated value of Δ and taking account of the values for ARL_δ and ARL_0. It is highly recommended to look for software solutions (many commercial and free packages offer simple functions to determine h and ARL_δ for given Δ and ARL_0).

g) The value for f can 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.



Key

X standardized decision interval, h

Y standardized difference to detect, Δ

Figure 7 — Nomogram for V-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 in most cases is 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 is selected. That decision depends 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 are 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 10 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 detects the shift in process level quicker than the corresponding CS1 scheme, but at the expense of more 'false signals'. Table 11 illustrates the differences in performance of these standard schemes.

Table 10 — 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, ARL_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, ARL_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 $\bar{R}$ 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.

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

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

Table 11 — 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 varies 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 is be calculated as $T = d_2\sigma$ where d_2 is taken from Table 8, 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 $a\bar{R}$, where a is taken from Table 12.

Table 12 — 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) $H = h\bar{R}$; and

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

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 like 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 also becomes 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 requires 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 can 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 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 is estimated from the CUSUM gradient. In this case, the interpretation is of how much the variability, 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, is 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 now operates 'in-control'.

If the direction of the CUSUM indicates a reduction in range, it is usually 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 re-zeroed before plotting is continued.

It is unnecessary to re-zero 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 like 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 13 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 gives a few more 'false signals' than a CS1 scheme but detects an important change a little more quickly than the corresponding CS1 scheme. Table 14 illustrates the differences in performance of these standard schemes.

Table 13 — 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
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, ARL_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, ARL_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 (σ_0) to determine the actual size and shape of the mask. This is described later in Step 5.

Table 14 — 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

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

Table 14 *(continued)*

Subgroup size	Actual process variability level	CS1 scheme	CS2 scheme
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 varies and can be shorter than or longer than the ARL. When it is of interest, the reader can examine the distribution of run lengths for particular shifts from target to know the expected range of run lengths that can 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 15.

Table 15 — 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
[a] Values of c_4 exist for n > 20. See ISO 7870-2 or other textbooks or standards.	

Table 15 *(continued)*

Subgroup size, n [a]	c_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 are 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 15.

b) Performance-based value

From the data obtained during the trial period, set the target within-subgroup standard deviation equal to 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 requires a different scale to that chosen to monitor the mean value. A suitable scaling is 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 16.

Table 16 — 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 becomes 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 can 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.

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 generate data that by their nature occur one-at-a-time and any notion of subgrouping such data makes no sense. Examples, given earlier, are 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 does not show any variation between the observations. In such a circumstance, the within-subgroup variation becomes zero and so no mask can be drawn.

Another example of this is the way 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 $T = 1{,}128\sigma$ 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 the classical setup of a CUSUM chart, as for any control chart, is that of independence between the plotted points. There are certain processes or sets of data where this can 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 can 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 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 is 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 depends 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 17 provides approximate guidance on significant correlation coefficients.

Table 17 — 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 17, 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 is required to determine the best way forward. Actions involves 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, overcomes by altering the target for each time. The CUSUM value is made independent of the seasonality in this way. For more details we refer to ISO 7870-9.

9.5.4 Outliers

CUSUMs require protection against outliers. If an outlier occurs, its influence on the CUSUM value can be great and can 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 ± $3,5\sigma_e$ from the target value; or

   2) an individual result is more than ± 3,5σ 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 ± $2\sigma_e$ from the target value; or

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

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

NOTE More rigorous methods (ISO 5725-5) to detect outliers exist but these are not very useable in 'real time'. The method described 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 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, when events occur randomly in time or space and the observation is made of the number of events in each 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 is not symmetrical as before. There are 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, the reference value K derived from a representative out of control 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 reference value K should be chosen based on 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 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, μ, and the standard deviation, σ_e.

Step 2 — Select a reference or target rate of occurrence, *T*. Frequently this is labelled as μ_0.

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 18

Step 4 — Determine *H* and *K* values thus:

a) for *T* (0,1 ≤ *T* ≤ 25,0) enter Table 18 at the nearest value to *T*. Use linear interpolation between values of *T* from 10,0 to 25,0; or

b) if *T* > 25,0 refer to the appropriate tables in 9.3 relating to the normal distribution, using $\sigma_e = \sqrt{T}$

c) as the normal approximation to the Poisson is now appropriate.

As an example, suppose *T* = 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*) and use a V-mask with decision interval, *H*, and slope, *F* (= *K* – *T*); 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 20.

EXAMPLE

Step 2: Reference mean rate, *T* = 4.

Step 3: Use a CS1 scheme.

Step 4: Enter Table 21 at *T* = 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 in Table 18. If the process was operating at the target level, the ARL (ARL_0) is 1 736. However, the ARL falls to 10 if the rate increases to 6,60.

Table 18 — Excerpt from Table 20 needed for example

H	*K*	*T*	ARL_0	ARL_δ	**1 000**	**500**	**200**	**100**	**50**	**20**	**10**	**5**	**2**
8,0	6,00	4,000	1 736	m=T+ δ	4,160	4,380	4,710	5,000	5,300	5,90	6,60	7,80	11,50

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 many 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 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 19 — CS1 and CS2 schemes for count (Poisson) data in terms of *T*, *H* and *K*

Target event rate	CS1 scheme		CS2 scheme	
T	*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 The choice of values of *T* 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 3 For *T* 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 ARL_0 slightly below 1 000; the higher value gives ARL_0 nearly 2 000.