

BS ISO 7870-4:2011
BS 5703:2011

Incorporating corrigendum February 2012



BSI Standards Publication

Control charts

Part 4: Cumulative sum charts



National foreword

This British Standard is the UK implementation of ISO 7870-4:2011. It supersedes BS 5703-1:2003, BS 5703-2:2003, BS 5703-3:2003, and BS 5703-4:2003, which are withdrawn.

This international standard has been produced by compressing the four parts of BS 5703 into a single part of ISO 7870. The resultant standard contains everything that was already in BS 5703 but it is written in a more concise manner. The content of this international standard has been further enhanced and augmented by the inclusion of text and examples that have been contributed by other nations during the drafting phase of the project.

The UK participation in its preparation was entrusted to Technical Committee SS/4, Statistical Process Management.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2012

Published by BSI Standards Limited 2012

ISBN 978 0 580 77134 7

ICS 03.120.30

Compliance with a British Standard cannot confer immunity from legal obligations.

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 30 September 2011.

Amendments issued since publication

Date	Text affected
29 February 2012	This corrigendum renumbers BS ISO 7870-4:2011 as dual numbered standard BS ISO 7870-4:2011/BS 5703:2011

Control charts —

**Part 4:
Cumulative sum charts**

Cartes de contrôle —

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



Reference number
ISO 7870-4:2011(E)



Documents

Document Sharing Hub
© ISO 2011



COPYRIGHT PROTECTED DOCUMENT

© ISO 2011

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword	iv
Introduction.....	v
1 Scope	1
2 Normative references	1
3 Terms and definitions, abbreviated terms and symbols.....	1
3.1 Terms and definitions	1
3.2 Abbreviated terms	2
3.3 Symbols.....	3
4 Principal features of cumulative sum (cusum) charts.....	4
5 Basic steps in the construction of cusum charts — Graphical representation.....	5
6 Example of a cusum plot — Motor voltages.....	5
6.1 The process	5
6.2 Simple plot of results	6
6.3 Standard control chart for individual results	7
6.4 Cusum chart — Overall perspective.....	7
6.5 Cusum chart construction.....	8
6.6 Cusum chart interpretation	9
6.7 Manhattan diagram.....	12
7 Fundamentals of making cusum-based decisions	12
7.1 The need for decision rules.....	12
7.2 The basis for making decisions	13
7.3 Measuring the effectiveness of a decision rule.....	14
8 Types of cusum decision schemes	16
8.1 V-mask types	16
8.2 Truncated V-mask	16
8.3 Alternative design approaches	22
8.4 Semi-parabolic V-mask	23
8.5 Snub-nosed V-mask	24
8.6 Full V-mask	24
8.7 Fast initial response (FIR) cusum	25
8.8 Tabular cusum	25
9 Cusum methods for process and quality control	27
9.1 The nature of the changes to be detected	27
9.2 Selecting target values	28
9.3 Cusum schemes for monitoring location	29
9.4 Cusum schemes for monitoring variation	39
9.5 Special situations	47
9.6 Cusum schemes for discrete data.....	49
Annex A (informative) Von Neumann method.....	56
Annex B (informative) Example of tabular cusum	57
Annex C (informative) Estimation of the change point when a step change occurs.....	61
Bibliography.....	63

Foreword

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

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

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

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

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

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

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

- *Part 1: General guidelines*
- *Part 3: Acceptance control charts*
- *Part 4: Cumulative sum charts*

The following part is under preparation:

- *Part 2: Shewhart control charts*

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

Introduction

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

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

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

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

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

Control charts —

Part 4: Cumulative sum charts

1 Scope

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

2 Normative references

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

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

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

3 Terms and definitions, abbreviated terms and symbols

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

3.1 Terms and definitions

3.1.1

target value

T

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

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

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

3.1.2

datum value

⟨tabulated cusum⟩ value from which differences are calculated

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

3.1.3

reference shift

F, f

⟨tabulated cusum⟩ difference between the **target value** (3.1.1) and **datum value** (3.1.2)

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

3.1.4

reference shift

F, f

⟨truncated V-mask⟩ slope of the arm of the mask (tangent of the mask angle)

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

3.1.5

decision interval

H, h

⟨tabulated cusum⟩ cumulative sum of deviations from a **datum value** (3.1.2) required to yield a signal

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

3.1.6

decision interval

H, h

⟨truncated V-mask⟩ half-height at the datum of the mask

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

3.1.7

average run length

L

average number of samples taken up to the point at which a signal occurs

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

3.2 Abbreviated terms

ARL average run length

CS1 cusum scheme with a long ARL at zero shift

CS2 cusum scheme with a shorter ARL at zero shift

DI decision interval

EWMA exponentially weighted moving average

FIR fast initial response

LCL lower control limit

RV reference value

UCL upper control limit

3.3 Symbols

a	scale coefficient
C	cusum value
C_r	difference in the cusum value between the lead point and the out-of-control point
c_4	factor for estimating the within-subgroup standard deviation
δ	amount of change to be detected
Δ	standardized amount of change to be detected
d	lead distance
d_2	factor for estimating the within-subgroup standard deviation from within-subgroup range
F	observed reference shift
f	standardized reference shift
H	observed decision interval
h	standardized decision interval
J	index number
φ	size of process adjustment
K	cusum datum value for discrete data
k	number of subgroups
L_0	average run length at zero shift
L_δ	average run length at δ shift
μ	population mean value
m	mean count number
n	subgroup size
p	probability of “success”
$\bar{R}$	mean subgroup range
r	number of plotted points between the lead point and the out-of-control point
σ	process standard deviation
σ_0	within-subgroup standard deviation
$\hat{\sigma}_0$	estimated within-subgroup standard deviation

σ_e	standard error
s	observed within-subgroup standard deviation
$\bar{s}$	average subgroup standard deviation
s_x^-	realized standard error of the mean from k subgroups
T	target value
T_m	reference or target rate of occurrence
T_p	reference or target proportion
τ	true change point
t	observed change point
V_{avg}	average voltage
$\hat{V}_{\text{avg}}$	estimated average voltage
w	difference between successive subgroup mean values
x	individual result
$\bar{x}$	arithmetic mean value (of a subgroup)
$\bar{\bar{x}}$	mean of subgroup means

4 Principal features of cumulative sum (cusum) charts

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

- a) It is sensitive in detecting changes in the mean.
- b) Any change in the mean, and the extent of the change, is indicated visually by a change in the slope of the graph:
 - 1) a horizontal graph indicates an “on-target” or reference value;
 - 2) a downward slope indicates a mean less than the reference or target value: the steeper the slope, the bigger the difference;
 - 3) an upward slope indicates a mean more than the reference or target value: the steeper the slope, the bigger the difference.
- c) It can be used retrospectively for investigative purposes, on a running basis for control, and for prediction of performance in the immediate future.

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

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

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

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

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

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

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

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

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

- a) choose a convenient plotting interval for the horizontal axis and make the same interval on the vertical axis equal to 2σ (or $2\sigma_e$ if a cusum of means is to be charted), rounding off as appropriate; or
- b) where it is required to detect a known change, say δ , choose a vertical scale such that the ratio of the scale unit on the vertical scale divided by the scale unit on the horizontal scale is between δ and 2δ , rounding off as appropriate.

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

6 Example of a cusum plot — Motor voltages

6.1 The process

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

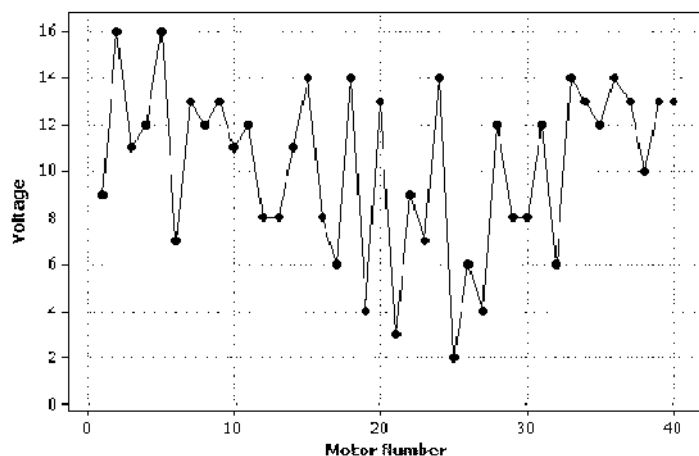
9, 16, 11, 12, 16, 7, 13, 12, 13, 11, 12, 8, 8, 11, 14, 8, 6, 14, 4, 13, 3, 9, 7, 14, 2, 6, 4, 12, 8, 8, 12, 6, 14, 13, 12, 14, 13, 10, 13, 13.

The reference or target voltage value is 10 V.

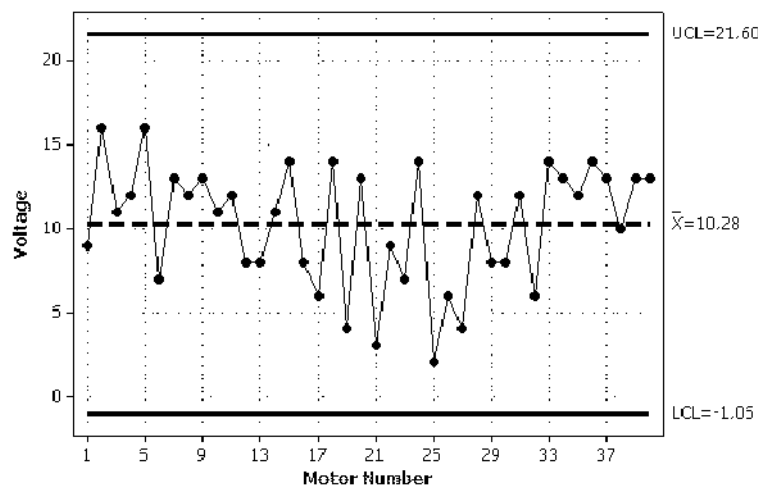
6.2 Simple plot of results

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

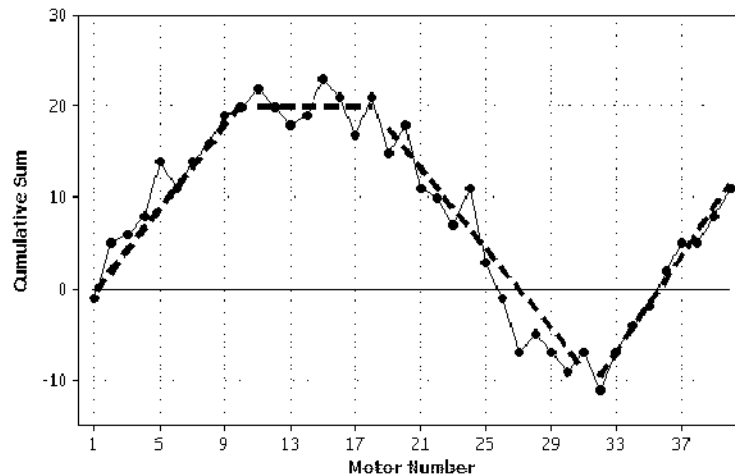
Apart from indicating a general drop away in the middle portion from a high start and with an equally high finish, Figure 1 a) is not very revealing because of the extremely noisy, or spiky, data throughout.



a) Simple plot of motor voltages



b) Standard control chart for individuals



c) Cusum chart

Figure 1 — Motor voltage example

6.3 Standard control chart for individual results

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

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

- a) no points lying above the upper control limit (UCL) or below the lower control limit (LCL),
- b) no runs of seven or more intervals upwards or downwards,
- c) no runs of seven points above or below the centreline.

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

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

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

6.4 Cusum chart — Overall perspective

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

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

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

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

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

6.5 Cusum chart construction

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

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

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

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

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

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

6.6 Cusum chart interpretation

6.6.1 Introduction

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

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

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

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

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

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

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

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

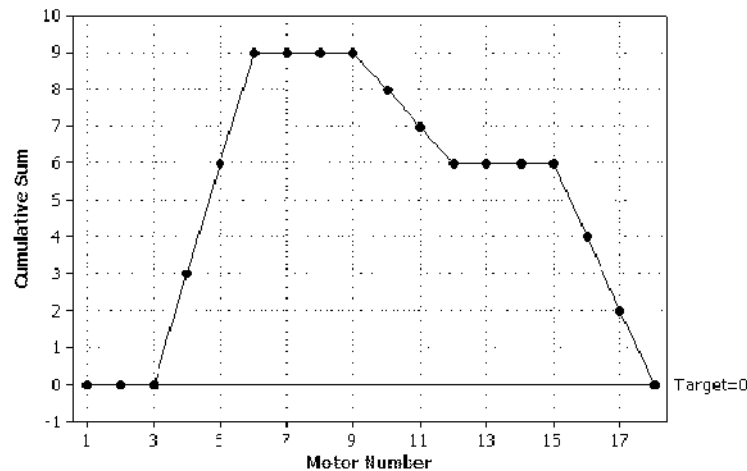


Figure 2 — Cusum chart of imaginary motor voltage data to illustrate its interpretation

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

- motors 1 to 3, 7 to 9 and 13 to 15 were all at the reference value of 10 V and that these are all represented by horizontal lines in the cusum chart. It will also be noted that the positions of the horizontal lines with respect to the vertical scale are not related to these actual motors but rather to previous performances;
- motors 4 to 6 were at a value higher than the reference value, namely 13 V, and that these motors are represented by an upward slope on the cusum chart. This is obvious here as there is no variability in voltage between the motors to confuse the issue. If there were noise then the equation to calculate the average value over the period from the particular slope is:

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

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

- similarly for motors 9 to 12:

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

- and for motors 16 to 18:

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

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

- 1 to 3, 7 to 9 and 13 to 15, the voltage remained constant at a value of 10;
- 4 to 6, the voltage also remained constant but at a value of 13;
- 10 to 12, the voltage remained constant at a value of 9; and
- 16 to 18, the voltage remained constant at a value of 8.

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

6.6.3 Interpretation using “actual” data

The cusum chart of Figure 1 c) shows:

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

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

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

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

- c) for motors 19 to 31:

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

- d) for motors 32 to 40:

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

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

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

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

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

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

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

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

6.7 Manhattan diagram

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

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

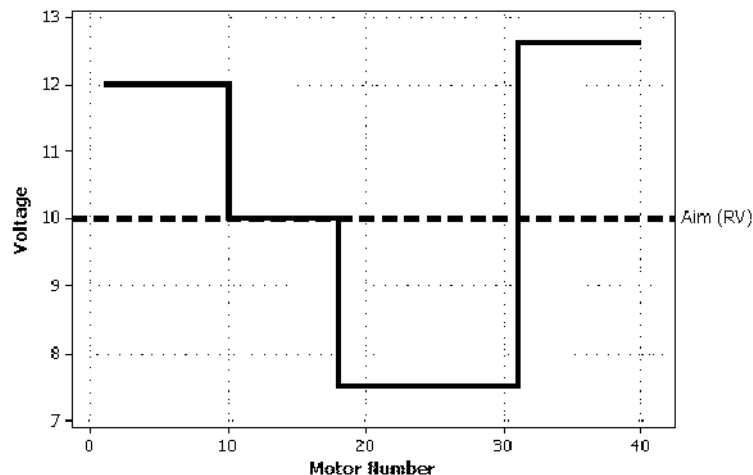


Figure 3 — Manhattan plot of motor data

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

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

7 Fundamentals of making cusum-based decisions

7.1 The need for decision rules

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

- a) *for in-process control:* adjustment of process conditions;
- b) *in an improvement context:* investigation of the underlying cause of the change; and
- c) *in a forecasting mode:* analysis of and, if necessary, adjustment to the forecasting model or its parameters.

7.2 The basis for making decisions

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

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

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

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

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

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

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

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

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

- a) an industrial process is controlled by a thermostat or other automatic control device;
- b) the quality of raw material input can be subject to minor variations without violating specification; and
- c) in monitoring a patient's response to treatment, there might be minor metabolic changes connected with meals, hospital or domestic routine, etc., but any effect of treatment should be judged against the overall typical variation.

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

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

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

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

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

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

7.3 Measuring the effectiveness of a decision rule

7.3.1 Basic concepts

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

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

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

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

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

7.3.2 Example of the calculation of ARL

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

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

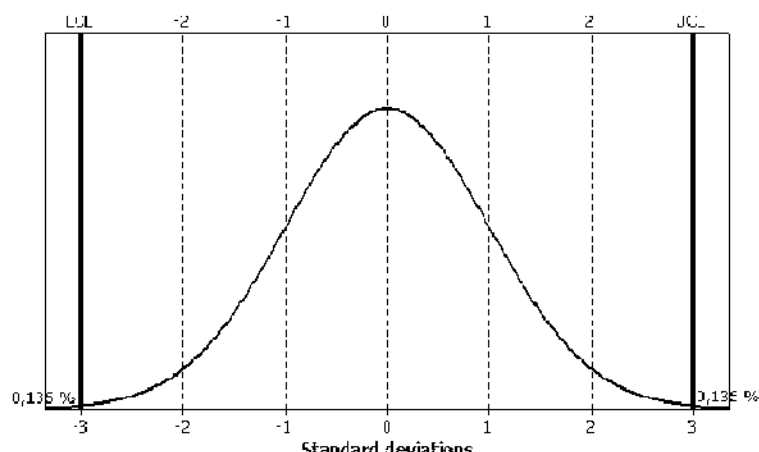


Figure 4 — Plot of standardized normal distribution

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

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

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

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

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

Summarizing:

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

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

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