

**THE BASICS OF  
STATISTICAL PROCESS CONTROL  
& PROCESS BEHAVIOUR CHARTING**

*A User's Guide to SPC*

*By David Howard*

## *The Basics of Control Charting*

*"...Shewhart perceived that control limits must serve industry in action. A. process, even in statistical control, wavers. Control limits can thus not be associated with any exact probability of looking for trouble (an assignable cause) when there is none, nor with failure to look for trouble when an assignable cause does exist. It was for such reasons that he used 3-sigma control limits. Experience of over 50 years shows how right he was."*

**W Edwards Deming (1900-1993)**

*Management-NewStyle is grateful to Mal Owen for permission to draw upon examples and illustrations from Chapter 5 of his book *SPC and Business Improvement* in preparing this booklet.*

# *The Basics of Control Charting*

## ***Contents***

1	<i>Introduction</i>	5
2	<i>Histograms &amp; Charts</i>	5
3	<i>Common and Special Causes</i>	6
4	<i>Control Limits</i>	7
5	<i>Using SPC to Drive Improvement</i>	9
6	<i>Collecting Data</i>	10
7	<i>Basis for Control Chart Limits</i>	11
8	<i>Shewhart or Otherwise?</i>	13
9	<i>Points to Note</i>	14
10	<i>Key Words</i>	14

## ***Appendix***

	<i>Calculations</i>	16
	<i>Constants</i>	16
	<i>References</i>	17

# The Basics of Control Charting



Figure 1 - W A Shewhart (1891-1967)

“How much variation should we leave to chance?”

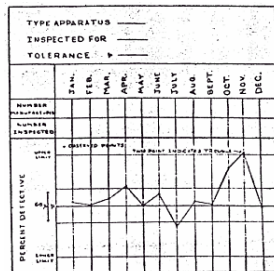


Figure 2 - The first control chart – 1924

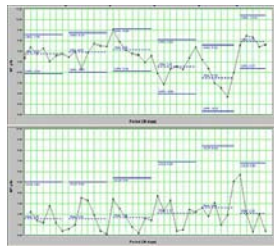


Figure 3 - Example of a FlowMap SPC-Plus control chart – 2003

## 1. INTRODUCTION

Statistical Process Control (SPC) charts offer users the chance to monitor the very heartbeat of their processes. By collecting data they can predict performance. Taking sample readings from a process seems straightforward. Of does it? Look more closely. Do we understand our process fully?

In manufacturing areas we probably do. In non-manufacturing areas we may be less confident. And who collects the data? What sample size is required? How often are samples taken? These are vital questions to those intending to daily use the control chart with a view to improving process performance, particularly in non-manufacturing, or service, areas where the techniques are new.

The control chart has been with us since 1924. It has been tried and proven, and accepted as a highly effective tool in improving processes. In view of the fact that there is currently renewed interest in Shewhart's work, it is important to consider how the control limits were originally set up.

However, at the end of the day, it is the logic and rules of collecting data and interpreting the pattern of points on the chart that is the important issue in understanding process behaviour and the discovery of insights for process improvement.

## 2. HISTOGRAMS & CHARTS

Figure 4 shows a typical set of readings obtained by collecting samples from a process. Control charting requires the mean of each sample to be used, rather than the individuals. Figure 4 also shows the calculated values of the mean  $\bar{X}$  and range R.

$x_1$	7	8	7	6	8	10	11	8	8	8	7	9	5	7	8	10	7	7	7	7
$x_2$	7	7	4	7	7	7	10	3	9	7	7	7	10	6	5	4	6	8	11	9
$x_3$	6	6	9	6	10	8	8	10	4	10	4	7	10	7	6	8	7	10	6	10
$x_4$	8	7	8	7	5	6	5	7	9	7	10	8	9	6	7	9	7	5	5	9
$x_5$	5	9	6	9	6	6	5	9	6	3	8	7	8	7	8	6	8	9	8	8
$\Sigma X$	33	37	34	35	36	37	39	37	36	35	36	38	42	33	34	37	35	39	37	43
$\bar{X}$	6.6	7.4	6.8	7.0	7.2	7.4	7.8	7.4	7.2	7.0	7.2	7.6	8.4	6.6	6.8	7.4	7.0	7.8	7.4	8.6
R	3	3	5	3	5	4	6	7	5	7	6	2	5	1	3	6	2	5	6	3

Figure 4 – Presentation of sample data

The first stage in constructing a control chart for  $\bar{X}$  requires  $\bar{X}$ , the mean, to be plotted against time as shown in Figure 5.

## The Basics of Control Charting

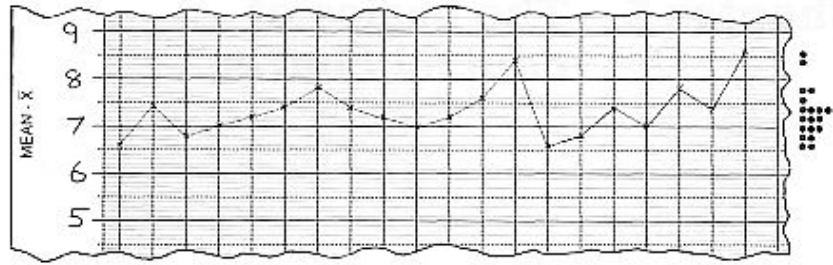


Figure 5 – Plot of the sample mean

The histogram corresponding to these  $\bar{X}$  values is also shown to the right of the chart. This pattern of points over time results in a unique profile of points – or histogram. However, the reverse is not true. For example, if all that is known is the histogram, what can we determine about the process? Very little, in fact. Figure 6 illustrates just three of the many patterns which all give the same histogram. What the histogram does not tell us, therefore, is the manner in which it was built up. We need a graph over time to determine this, and these graphs, commonly called run charts, are a first step in generating control charts.

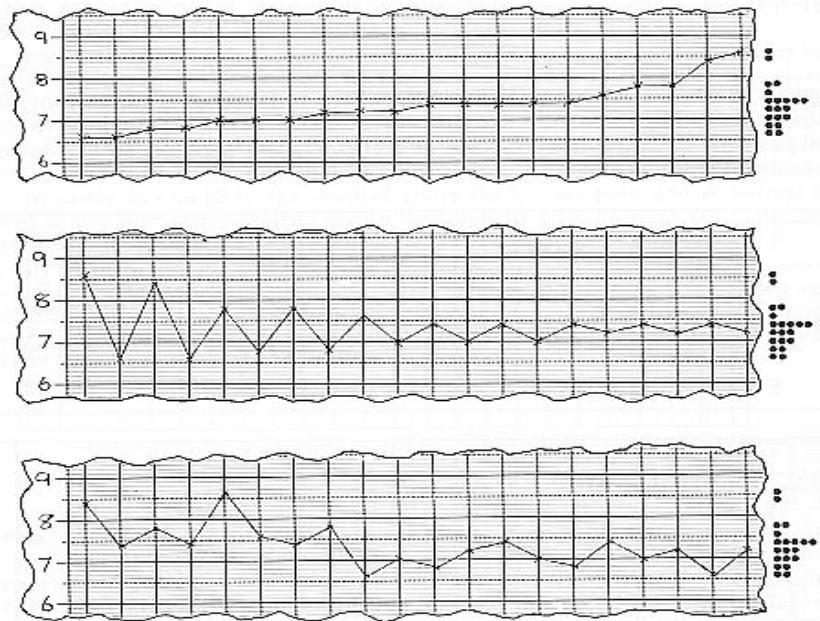


Figure 6 – Run charts from a common histogram

### 3. COMMON & SPECIAL CAUSES

The form of control chart which we use today was first generated by Shewhart. He recognised that if a process was stable it was also predictable. i.e., once the natural variation of the process has been determined, it is then possible to predict future performance. This natural variation of the process does not alter over time unless action is taken to change the system.

A process is defined as being stable if its natural variation is due to common causes. The process is then said to be under statistical control. If a process is unstable, that is because unusual factors are operating on the process. These factors, known as special causes, result in the process being out of statistical control. Shewhart recognised that we make mistakes at times, in that we take action when we should not do so.

Equally, we sometimes let things drift, assuming the process will right itself, when in fact we should react at the first sign of trouble. Shewhart was therefore aiming to devise a rule which would be sensitive enough to pick up a special cause, but not so sensitive as to react to extremes in terms of common causes.

Take the figures plotted in Figure 5 as an example. It makes sense to use a central value as a reference point. The best measure of central location is the mean value, that is the 'average' obtained by adding all one hundred readings and dividing the total by 100, giving:

$$733 / 100 = 7.33$$

In fact, this mean can be obtained much more directly. We already have the values of the 20 sample means. Hence the overall mean is given by the following calculation:

$$\sum \bar{X} / 20 = 146.6 / 20 = 7.33$$

This mean of the sample means is known as the grand mean, or  $\bar{\bar{X}}$  (X double bar), thus:

$$\bar{\bar{X}} = 7.33$$

Figure 7 shows our run chart of  $\bar{X}$  values together with a line for  $\bar{\bar{X}}$  conventionally drawn as a broken line and often called the central line. The question is now where do we draw lines on the chart which will sensibly indicate the presence of special causes?

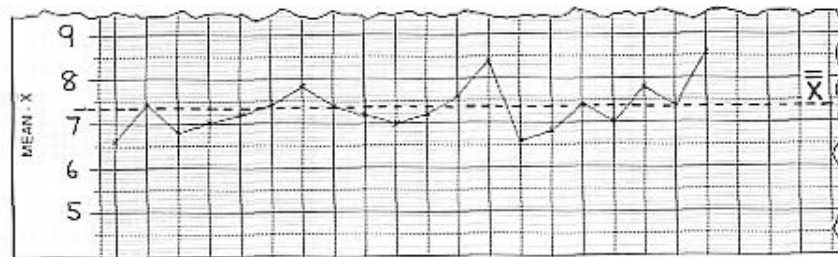


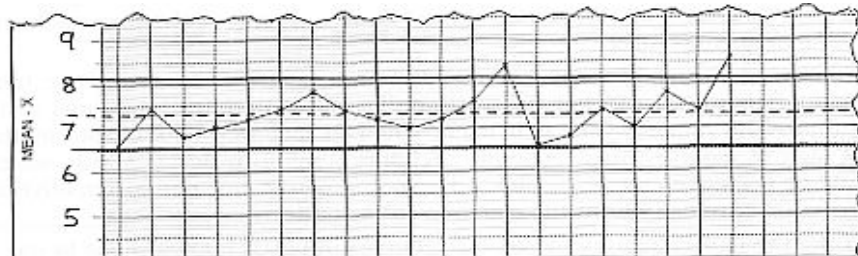
Figure 7 – Run chart for  $\bar{X}$  with central line

#### 4. CONTROL LIMITS

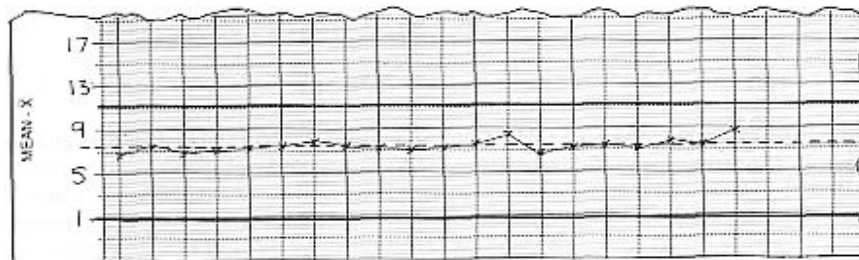
If we draw lines too close to  $\bar{\bar{X}}$  as in Fig 6 then we will be reacting too often to points which are really part of the system and not special causes. Fig 7 shows the reverse - lines set so far out that they will not pick up any change unless it is a major one and obvious. A balance between these two cases is required. Shewhart chose lines set at

*The Basics of Control Charting*

three standard deviations away from  $\bar{\tilde{x}}$  - commonly known as control limits. Why three standard deviations? Because this number has been found to be economically practical in use over the past 75 years. It keeps a balance between over, and under, reaction to process behaviour patterns. Such limits represent pragmatically powerful markers of change in process performance. They can reliably provide a predictive warning of an onset of instability.



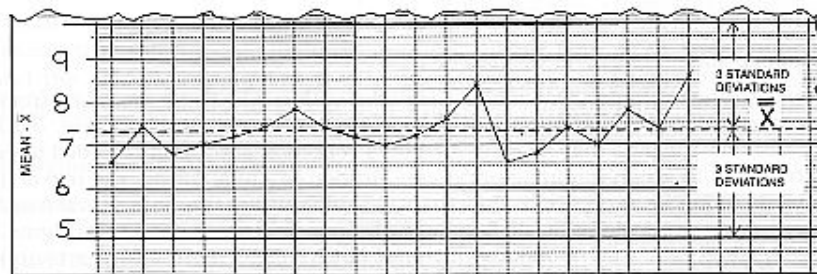
*Figure 8 – Control chart for  $\bar{X}$*



*Figure 9 – Run chart with decision lines too near to  $\bar{X}$*

There is a bit more to it than that, however, as you might expect, and the apparent simplicity of this rule brings with it some controversy. The derivation of the control limits has been a point of discussion on both sides of the Atlantic for over half a century. This is because Shewhart's original thinking has been augmented by others. (In particular we would refer the reader to the work of Don Wheeler, in the US, listed on p 24.)

In order to complete the control chart for  $\bar{X}$ , we now add a central line and upper and lower control limits (denoted by  $UCL_{\bar{X}}$  and  $LCL_{\bar{X}}$ ). This gives us the chart shown in Figure 10.



*Figure 10 – Run chart with decision lines too far from  $\bar{X}$*

Determining the position of the lines will involve some simple formulae which depend on whether we are looking at multiple of individual readings. The arithmetic steps to calculate the limits are easy to master and summarised in the Appendix.

### 5. USING SPC TO DRIVE IMPROVEMENT

Whether charting variables or attributes the approach follows the same sequence: 1) Define the process; 2) Collect the data; 3) Set up the chart; 4) Plot the results; 5) Check on control, and 6) Adjust the process. Let us now consider each in turn. In manufacturing areas a process is easy to understand and define. In non-manufacturing areas it is less so. The processes tend to be more complex and more difficult to specify clearly.

A definition of the process needs to be obtained, but it requires careful preparatory work. Flow charting is a key tool in this definition. An additional problem is that administrative processes are much more people-orientated and as a result, personalities and emotions are involved. (See the companion booklet *The Basics of Deployment FlowCharting & Process Mapping* for details of process mapping).

Process data for charting is collected by taking sample readings. Sampling, as opposed to 100% inspection, is not only easier, it is more representative and quicker. The sampling procedure will differ for variables and attributes. For variables, collecting the data involves issues of sample size, number of samples and their frequency. For attributes sampling is the exception, not the rule. Manufacturing examples generally use samples with a size of 5. These are taken from the process at regular intervals and represented on the ( $\bar{X}$ , R) chart. Figure 11 indicates how the data may be collected and organised on a collection sheet ready to be recorded on a control chart.

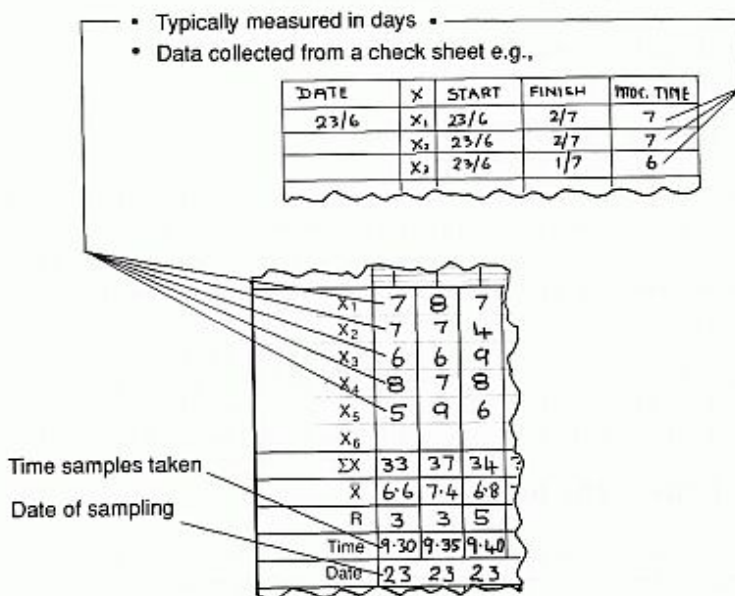


Figure 11 – Collection and recording of sample data

In non-manufacturing, the use of a sample tends to be the exception. In practice it is much more likely that we are looking at single readings - one document, one sales figure, one event. Single reading values are monitored using what we call ( $\bar{X}$ , moving R) charts. However, there is a certain logic about the sequence in which control charts are introduced.

There is a natural flow in progressing from sampled values to dealing with individual values, and there is a danger of confusing the understanding if this sequence is changed. In starting with control charts it is therefore helpful to firstly explore ( $\bar{X}$ , R) charts and then progress to ( $\bar{X}$ , moving R) charts. If we are looking at processing times, for example, then five documents could be tagged first thing in the morning and subsequently the time when each document is completed would be noted. For financial data, for example sales figures, five results would be taken in a suitable manner from the many figures which are available.

The ideal number of different samples needed to construct a chart with control limits is generally 20. There is no reason why interim calculations of process limits could not be carried out on fewer than 20 samples and then updated when 20 samples are available.

In administrative areas the frequency is typically daily, weekly, or even monthly. Judgement will dictate. The nature of the great majority of processes does not allow for sampling at hourly intervals or less, as commonly applied in manufacturing. Odd exceptions may emerge. For example, response times to a computer programme could be taken every 15 minutes. Figure 12 provides a summary of some of the issues associated with sampling.

## **6. COLLECTING THE DATA**

Whatever the organisation, manufacturing or otherwise, personnel should be responsible for monitoring their own processes. Hence, in the same way as operators in manufacturing industry collect data, their equivalents in non-manufacturing - clerical assistants, clerks, technical support staff, managers - should all collect sampled data for their administrative processes.

This may cause problems initially since it is a change from the tradition of looking for trends in data or comparing the data of one period with another. Administrative people recognise that data has traditionally been collected in order to measure levels of productivity, rather than the inherent capability of their processes.

There may be a natural reluctance to assist in an activity which may have repercussions on their own employment. This is understandable but management has the duty to re-assure those who may be confused if process change is to be achieved by the informed analysis of available data with the intention of improvement action.

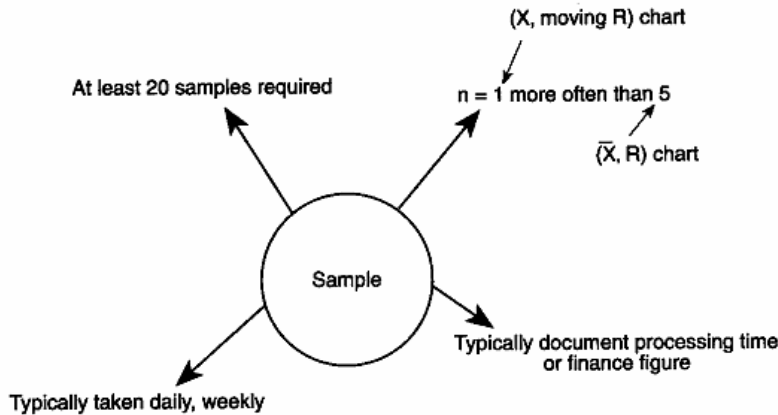


Figure 12 – Issues relating to a sample

Properly handled, however, and in an environment of trust and co-operation, many difficulties should be surmountable. The insights for improvement available from the analysis of such data are considerable and should not be overlooked by any organisation determined to improve its performance and customer service levels.

## **7. BASIS FOR CONTROL LIMITS**

Shewhart designed the control chart so that users could:

- Distinguish between common (systemic) and special (random) causes of variation within a process; and
- Provide a rule for minimising the risk of reacting to a special cause when it is in fact a common cause, and not reacting to a special cause when one is present.

He opted, after considerable investigation and experiment, to set his control limits at three standard deviations remove from the mean, central line and applied this rule to all process charts, be they monitoring variable or attribute data. It worked well, was adopted by others, and gradually became the pragmatic standard. The use of this three sigma rule makes no assumptions regarding the nature of the process.

For example, the readings may come from a normal distribution, but we do not know. It is irrelevant in the approach adopted by Shewhart, and later also advocated by Deming. An alternative approach, adopted in the UK and Europe, followed on from the publication of Shewhart's paper.

A brief summary of the normal curve is relevant here. Figure 13 illustrates how a histogram approaches a normal distribution with the properties shown in Figure 14. In particular, for sample mean  $\bar{X}$ , we have the properties shown in Figure 15. Hence, if we know the distribution is normal, it is suggested that control limits can then be set as shown in Figure 16. The chance, probability, of a point lying outside a particular control limit is then 0.00135 (nearly one part in 1,000).

## The Basics of Control Charting

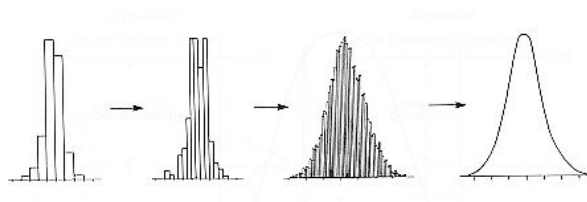


Figure 13 – Generation of a normal curve

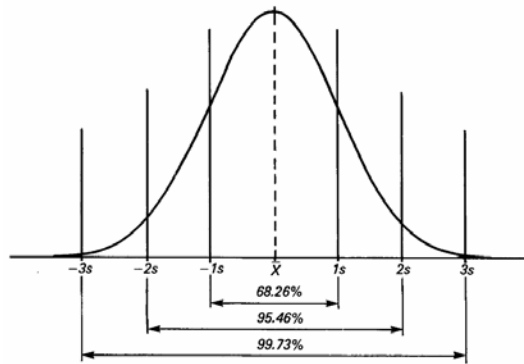


Figure 14 – Areas under the normal curve with corresponding  $\delta$  values

The British Standards Institute has preferred an operating definition of exactly one in 1,000, corresponding to 3.09 standard deviations from the mean. This subtle difference has not helped in providing a common approach to the generation of control limits and is of no practical advantage in real usage.

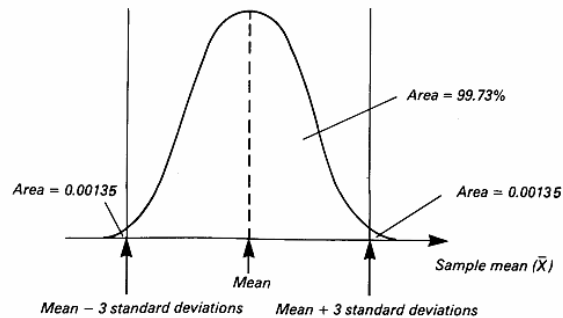


Figure 15 – Normal curve for distribution of sample mean

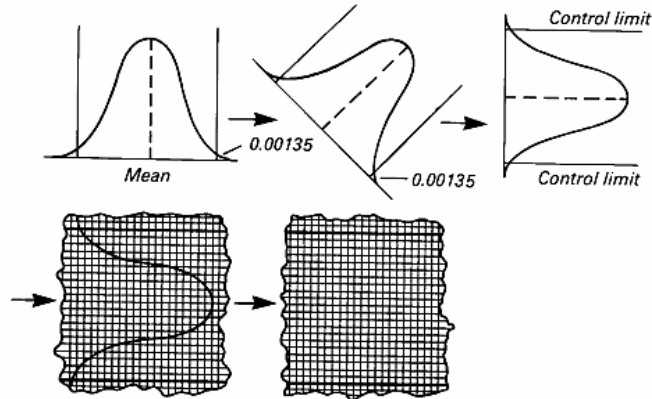


Figure 16 – Normal curve and control limits

### 8. SHEWHART OR OTHERWISE?

The Shewhart approach and the probability approach do address the question of the origin of the control limits from differing viewpoints, and to that extent they are in conflict. Figure 17 illustrates the reasoning behind the thinking. In a great many cases the underlying distribution is normal or sufficiently close to normal to be treated as such.

It is then a logical step to argue that the area in the tail is 0.00135 when the limits are at three standard deviations from the mean. The drawback is that with no prior knowledge of the process we cannot assume a normal distribution. Hence the reason for placing the emphasis on the  $3\sigma$  rather than the area in the tail.

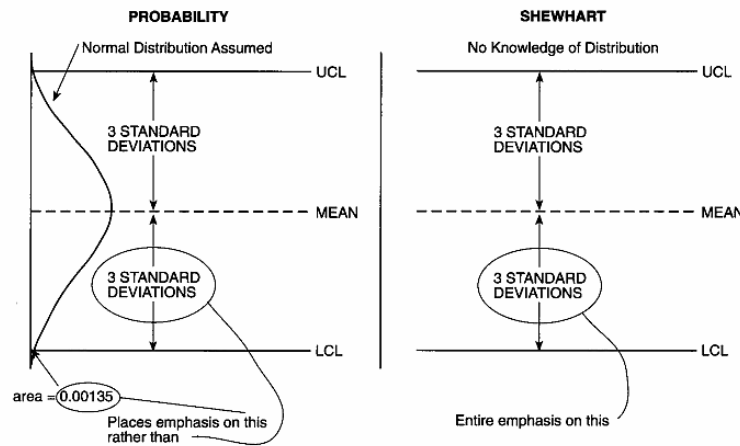


Figure 17 – Probability and Shewhart charts

The Shewhart approach is the more correct in that it makes no assumptions regarding the normal distribution, or any other distribution for that matter. It is based on the principle 'if it works in practice, let's use it'. It is too easy to be pedantic regarding one method or another. We cannot ignore the fact that there are two bodies

of opinion. For those involved in the practical application of interpreting charts in operational areas it should not be an issue. Familiarity with the actual calculating, plotting and interpreting is necessary, and the next few chapters will give you plenty of material to use.

### **9. POINTS TO NOTE**

- It might be helpful, without causing too much confusion, to mention the work of a mathematician called Tchebyshev. He showed, that if a process is stable, 89% of the time  $\bar{X}$  will fall within the limits:  $\bar{X}$  plus or minus three standard deviations, irrespective of the form of the distribution. The same result holds for control charts in general, not just  $(\bar{X}, R)$ , and provides a mathematical justification of Shewhart's approach (Figure 18).
- The Shewhart / Probability argument is less important than recognising that the limits are performance-based rather than specification-based. Performance limits which represent the process must not be confused with limits based on artificial targets. The rule for detecting a point outside a control limit is just one of four commonly adopted rules for detecting special causes.

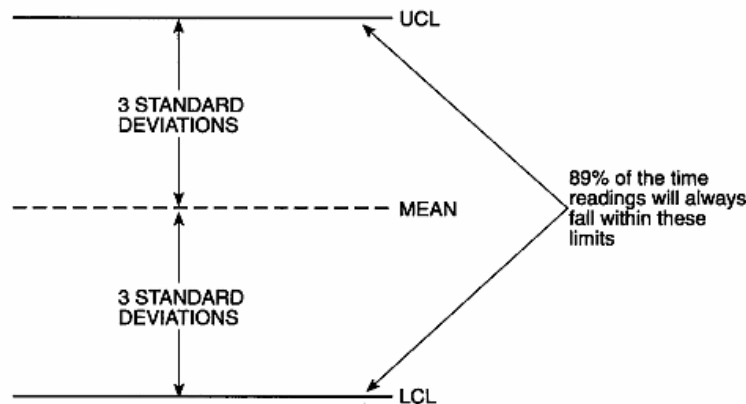


Figure 18 – Justification for limits set at  $\pm 3\sigma$

### **8. KEY WORDS**

With this general background we are able to set up control charts for both variable (measurable) and attributable (countable) data. The approach in both cases is very similar. Both types of charts require some simple calculations and, in the case of attribute charts, reference to some standard tables.

The beauty of these charts is the effective way in which they convert staid, uninteresting information into a new form. The layout is not only easier on the eye, but also offers well informed insight for corrective action. Terms, symbols and formulae introduced in this Note are listed in Table 1 and the Appendix presents the

*The Basics of Control Charting*

equations needed to calculate control limits for both chart types, sampled values ( $\bar{X}$ , R) charts and individual values (X,mR).

<b><i>Key Word</i></b>	<b><i>Meaning</i></b>
Sample	A representative group from the process.
Run Chart	A graph showing variation over time.
Stability	Condition of a process whose natural variation will not change if left to itself.
Predictable	The result of process stability.
System	The interaction of people, materials and machines which provides the environment within which we work.
Common Cause	A factor which is part of the system.
Statistical Control	Operating within natural limits of variation.
Special Cause	A factor which is outside the system. An unusual effect.
Out of Statistical Control	Not operating within natural limits of variation because of the presence of a special cause.
Grand Mean	The mean of a series of $\bar{X}$ values. Also known as $\bar{\bar{x}}$
Central Line	A line on the control chart corresponding to the average for the process.
Control Limits	The upper and lower limits of the natural variation of the process.
Normal Distribution	A symmetrical bell-shaped pattern. Readings from many processes tend towards this shape.
Probability	A numerical measure of risk.
Tchebyshev	Name of a mathematician whose work provides statistical evidence supporting Shewhart's practical use of three standard deviation limits.
$\bar{\bar{x}}$	Grand mean; i.e. $\sum \bar{X} / (\text{number of samples})$
UCL $\bar{x}$ ; LCL $\bar{x}$	Upper ; lower control limits for X.
( $\bar{X}$ , R)	Description of a control chart for dealing with samples.
( $\bar{X}$ moving R)	Description of a control chart for dealing with single readings.

*Table 1 – Meaning of Key Words and Symbols*

**APPENDIX**

**Calculations for ( $\bar{X}$ , R) Process Behaviour Charts**

*(SPC using Multiple Average Reading Values)*

R	Value of difference between lowest and highest sample
$\bar{R}$	Mean value of R; $\sum R / \text{number of values}$
$A_2$	Constant, depending upon sample size, used to calculate the control limits for $\bar{X}$ using value of R
$D_3$	Constant, depending upon sample size, used to calculate the lower control limit for R
$D_4$	Constant, depending upon sample size, used to calculate the upper control limit for R
$\bar{\tilde{x}}$	= $\sum X / \text{number of samples}$ (i.e. 20 usually)
$\bar{R}$	= $\sum R / 20$
$LCL_{\bar{X}}$	= $\bar{\tilde{x}} - A_2 \bar{R}$
$UCL_{\bar{X}}$	= $\bar{\tilde{x}} + A_2 \bar{R}$
$LCL_R$	= $D_3 \bar{R}$
$UCL_R$	= $D_4 \bar{R}$

**Calculations for ( $\bar{X}$ , mR) Process Behaviour Charts**

*(SPC using Individual Moving Range Values)*

mR	Value of difference between two adjacent samples
$\bar{X}$	Mean value of X; $\sum R / \text{number of values}$
$d_2$	Constant, depending upon sample size, used to calculate the control limits for X. For sample size = 2 the value is 1.128
$\tilde{\sigma}$	Best estimate of $\sigma$ , usually $\bar{R} / d_2 = \bar{R} / 1.128$
$LCL_X$	= $\bar{X} - 3 \tilde{\sigma}$ or $\bar{X} - 3(\bar{R} / 1.128)$
$UCL_X$	= $\bar{X} + 3 \tilde{\sigma}$ or $\bar{X} + 3(\bar{R} / 1.128)$
$LCL_R$	= $D_3 \bar{R}$
$UCL_R$	= $D_4 \bar{R} = 0$ when sample size = 2

**Table of Constants for use in Control Limit calculations**

Sample Size	$A_2$	$d_2$	$D_3$	$D_4$
2	1.880	1.128	-	3.267
3	1.023	1.693	-	2.574
4	0.729	2.059	-	2.282
5	0.577	2.326	-	2.114
6	0.483	2.534	-	2.004
7	0.419	2.704	0.076	1.924
8	0.373	2.847	0.136	1.864
9	0.337	2.970	0.184	1.816
10	0.308	3.078	0.223	1.777

**REFERENCES**

**Further Reading and Study**

*SPC and Continuous Improvement*

by Mal Owen, 1989

*SPC and Business Improvement*

by Mal Owen, 1993

*Understanding Variation*

by Donald J Wheeler, 1993

*Statistical Process Control in the Office*

by Mal Owen and John Morgan, 2000

---

This booklet is provided to users of The FlowMap System to provide general insights into business problem solving and process aligned management. The key to both of these activities is the discipline of 'thinking in systems and working on processes'.

This publication copyright (c) 2003 *Management-NewStyle*

ISBN : 0-9543866-5-5 1st Edition, October 2003

Published by *Management-NewStyle*,

Chislehurst, Kent, BR7 5NB, England [www.firstmetre.co.uk](http://www.firstmetre.co.uk)

**The FlowMap System** (c) 1990-2003 David Howard

**FlowMap Software** (c) 1994-2003 David Howard and Maurice Tomkinson

FLOWMAP is a registered trademark of *Management-NewStyle*